بسم الله الرحمن الرحیم

انتشارات دانشگاه فردوسی (مشهد)، شمارهٔ ۱۲۲

آمـــار

برای دانشجویان رشته‌های اقتصاد و بازرگانی

نوشته:

سیدعلی‌اکبر رزمی

عضو هیأت علمی دانشگاه فردوسی

فهرستنویس پیش از انتشار توسط کتابخانه مرکزی و مرکز اسناد.

رزمی، علی اکبر
آمار برای دانشجویان رشته‌های اقتصاد و بازرگانی / نوشته علی اکبر رزمی. ـ مشهد: دانشگاه فردوسی مشهد، ۱۳۷۰.
۲۶۳ ص.: جدول، نمودار. ـ (انتشارات دانشگاه فردوسی مشهد، ۱۲۲)

۱. آمار. ۲. آمار بازرگانی. الف. عنوان

HA ۲۹/ ۵۱۹/۵

مشخصات:

نام کتاب: آمار برای دانشجویان رشته‌های اقتصاد و بازرگانی

نوشته: سیّد علی‌اکبر رزمی

ناشر: انتشارات دانشگاه فردوسی (مشهد)

تیراژ: ۲۰۰۰ نسخه

تاریخ انتشار: اردیبهشت‌ماه ۱۳۷۱

چاپ و صحافی: مؤسسهٔ چاپ و انتشارات دانشگاه فردوسی (مشهد)

قیمت: ۱۶۰۰ ریال

فهرست

فصل دوم : مشخصه‌های تمرکز و پراکندگی

فصل سوم : توزیع احتمال

بدینوسیله از همکاری صمیمانه مسئولین محترم دانشگاه امام صادق (ع) که با حمایت بیدریغ و همه‌جانبهٔ خود تألیف این کتاب را امکانپذیر نمودند قدردانی نموده و توفیق آنها را از خداوند بزرگ مسئلت می‌نمایم.

همچنین زحمات همکار ارجمند جناب آقای سیدمهدی مصطفوی عضو هیأت علمی دانشکده علوم اداری و اقتصادی دانشگاه فردوسی (مشهد) که ویرایش کتاب را انجام داده‌اند و نیز همکاری کلیه دست‌اندرکاران حوزه معاونت پژوهشی این دانشگاه را ارج نهاده و موفقیت روزافزون آنها را از خداوند قادر متعال آرزومندم.

مقدمه

اهمیت آمار

علم آمار یکی از علومی است که امروزه برای اداره امور اجتماعات بشری شدیداً" مورد نیاز بوده و شدت این نیاز با پیشرفت و توسعه اقتصادی این جوامع و پیچیده‌ترشدن بافت سیاسی، اجتماعی آنها بطور روزافزونی روبه افزایش است به طوری که اکنون کمتر رشته تحصیلی دانشگاهی را می‌توان یافت که درس آمار به‌عنوان یکی از دروس ضروری آن رشته قلمداد نشده و در کنار سایر دروس ضروری، به‌تدریس آن پرداخته نشود. علت‌این امر نیز واضح است. تصمیم‌گیری در هریک‌از زمینه‌های اقتصادی، اجتماعی، سیاسی و غیره به‌مفهوم وسیع هریک از این کلمات وقتی‌میسراست که اطلاعات لازم بصورت نسبتاً" واقع نما و دقیقی در اختیارمسئولین تصمیم‌گیر ذیربط قرارداشته باشد. مقدار زیادی از این اطلاعات‌ضروری تنها ازطریق روشهای آماری قابل حصول هستند. همانطورکه گفته شد کاربرد روشهای آماری مسئله‌ای عام بوده و زمینه‌های بسیار زیاد و متنوعی را در بر می‌گیرد. به‌عنوان مثال وقتی مسئولین فرهنگی یک‌کشور می‌خواهندسیاست خودرا درزمینه تألیف و نشر کتاب پایه‌گذاری کنند باید از وضعیت مطالعه کتاب توسط عموم مردم از نظر کمی و کیفی با خبر باشند تا بتوانند سیاستهایی متناسب با احتیاجات و تمایلات جامعه اتخاذ نمایند. برای کسب این اطلاعات لازم‌است ابتدا به‌مردم مراجعه‌شده و از میزان و نوع مطالعات آنها پرسش به عمل آید. ازآنجا که معمولاً" تعداد افراد جامعه مورد مطالعه زیاد است و مراجعه و پرسش ازهمه آنها بسیارمشکل و حتی در بسیاری موارد غیر ممکن است برای سهولت کار در بیشتر اوقات بجای مراجعه به‌همه افراد جامعه، نمونه‌هایی متناسب ازآن افرادانتخاب شده و فقط‌اعضاء این نمونه‌ها ازجهت خصوصیات ویژه‌ای مورد‌پرسش قرار می‌گیرند. در مرحله بعدی‌اطلاعات خام بدست آمده در مرحله قبل را طبقه‌بندی نموده و پارامترهای بخصوصی مثل میانگین تعداد صفحات مورد مطالعه، میانگین کیفیت مطالعات و غیره را بدست می‌آورند. این

پارامترها هر مقدار و هر معنایی که داشته باشند فقط می‌توانند وضعیت نمونه انتخابی را نشان دهند درحالی‌که هدف مسئولین اطلاع از وضعیت مطالعه‌عموم افراد جامعه‌است. برای حل این مشکل‌از روشهای خاصی‌استفاده می‌کنندکه درنتیجه آنها می‌توان بر مبنای‌اطلاعات حاصله‌ازنمونه،اطلاعاتی‌هرچند بطورتقریبی درمورد جامعه‌بدست‌آورد. پس ازآنکه پارامترهای مربوط به جامعه بطور تقریبی و یا دقیق مشخص شدند زمان آن فرا می‌رسد که با تفسیر این اطلاعات نتایجی را در مورد مطالعه افراد جامعه استنتاج نموده و سپس سیاستهای فرهنگی مورد نظر را متناسب با وضعیت جامعه اتخاذ نمایند. به‌همین ترتیب وقتی که موءسسه استاندارد می‌خواهد در مورد کیفیت محصولات یک کارخانه خاص مانند شیشه همدان‌اظهار نظر کند ابتدا نمونه‌هایی از محصولات این موءسسه را انتخاب نموده و سپس خصوصیات مورد نظر از قبیل میزان مقاومت آنها در مقابل حرارت و ضربه، ضخامت‌شیشه و غیره را اندازه‌گیری نموده و پس از آن پارامترهای مهم آماری از قبیل میانگین و انحراف معیار این خصوصیات را در مورد نمونه و یا نمونه‌های ماخوذه محاسبه نموده و این نتایج را بطریقی به‌جامعه اصلی یعنی کلیه تولیدات کارخانه شیشه همدان تعمیم داده یا بر مبنای نتایج نمونه، پارامترهای مربوط به جامعه را بصورت تقریبی بدست می‌آورد. با تفسیر و تحلیل‌این نتایج اظهارنظر در مورد محصولات این‌کارخانه که هدف موءسسه استاندارد است‌قابل‌انجام خواهد بود. شیوه عمل در موارد بسیار دیگری از قبیل میزان درآمد افراد جامعه و توزیع‌آن، وضعیت سلامت‌افراد جامعه و هزاران هزارمورد دیگر در زمینه‌های گوناگون‌از قبیل اقتصادی بهداشتی، فرهنگی،صنعتی و غیره همانند دو مثال فوق است.

کلیه عملیاتی که در دو مثال مزبور و تمامی موارد مشابه دیگر انجام می‌گیرند به‌نحوی با روشهای آماری مرتبطند. در ابتدای کار یعنی هنگام نمونه‌گیری، برای‌آنکه نمونه‌های اخذ شده مناسب بوده و هدف محقق را تأمین کنند باید از تئوری نمونه‌گیری[1] استفاده نمود که یک تئوری کاملاً" آماری است. محاسبه پارامترهای موردنظر از قبیل میانگین و واریانس نیز اگرچه ماهیتی ریاضی دارد اما ازمسائل موردشمول علم آمار است. انتقال‌نتایج ماخوذه در مورد نمونه به جامعه کاری صرفاً" آماری بوده و بر مبنای تئوریها و نظریات آماری قابل انجام است. بدین ترتیب ملاحظه می‌گردد که اداره امور اجتماعات بشری به نحو مطلوب، با بافت پیچیده و مسائل گوناگونی که دارند با علم آمار ارتباط‌تنگاتنگ و نزدیکی دارد و به‌همین دلیل علم آمار امروزه در میان اکثریت قریب به اتفاق علوم جایگاه ویژه‌ای دارد.

حال که اهمیت علم آمار و روشهای آماری مشخص گردید لازم است قبل از ورود

---

1 - Sampling Theory

به مباحث اصلی تصویری هرچند کلی و مختصر از آنچه در درس آمار مورد بحث قرار می‌گیرد ارائه گردد تا در هنگام مطالعه مباحث آماری، خواننده چهار چوب کلی بحث را در نظر داشته و جایگاه هریک از موضوعاتی را که در طول این کتاب مورد بحث قرار می‌گیرند بصورت منطقی بداند. یعنی بداند که بعنوان مثال چرا او برای تعیین میانگین درآمد افراد یک جامعه و یا بسیاری از کارهای آماری دیگر محتاج دانستن تئوری احتمال و نقش آن در عملیات آماری می‌باشد. برای تحقق هدف فوق یعنی ارائه یک تصویر کلی از مسائل مورد بحث در علم آمار، ذیلا" ابتدا هدف مشخص هریک از عملیات آماری ذکر شده و سپس روش تحقق این اهداف مختصرا" توضیح داده می‌شود.

اهداف عملیات آماری

بطور کلی عملیات آماری را از نظر نتیجه‌ای که هریک از آنها مشخص می‌کنند می‌توان به سه دسته تقسیم کرد.

الف ــ عملیاتی که در نتیجه آنها پارامترهایی از جامعه از قبیل میانگین، واریانس و انحراف معیار و .... بطور دقیق و یا تقریبی تعیین می‌گردند. مهمترین مبحث آماری که به این عملیات مربوط می‌شود، مبحث تخمین نقطه‌ای و فاصله‌ای و به عبارت بهتر تئوری تخمین[1] و در مواردی روشهای آزمون فرضیه‌ها بوده و مهمترین پارامترهایی که معمولا" در این گونه عملیات محاسبه می‌گردند میانگین، واریانس و انحراف معیار می‌باشند.

ب: عملیاتی که در نتیجه آنها درمورد تساوی مقدار دو یا چند پارامتر از چند جامعه و بخصوص از چند نمونه می‌توان قضاوت نمود. آزمون تساوی میانگین دو یا چند جامعه، واریانس دو یا چند جامعه و غیره نمونه‌هایی از این موارد هستند که عمدتا" در دو مبحث آزمون فرضیه‌ها[2] و تجزیه واریانس[3] مورد آزمون قرار می‌گیرند.

ج ــ بالاخره دسته سوم عملیاتی هستند که در آنها تأثیر یک عامل بر عامل دیگر اعم از آنکه این تأثیر، از نوع علت و معلولی بوده و یا نوع دیگری باشد مورد بررسی و آزمون قرار گرفته و در مورد وجود و یا عدم آن اظهار نظر می‌شود. بعنوان مثال وقتی قرار است در مورد تأثیر یک کود بخصوص در افزایش بازده محصول گندم تحقیق شود ابتدا این کود را در قسمتی از زمین مورد آزمایش بکار برده و میزان محصول را در دو قسمت کود دار و بدون

---

1 - Estimation theory

2- Testing Hypotheses

3 - Analysis of Variance

کود ثبت می‌کنند . از این پس با استفاده از عملیات آماری دسته سوم که عمدتا" در دوروش تجزیه واریانس و رگرسیون[1] خلاصه می‌شوند راجع به تأثیر کود مزبور و یا عدم تأثیر آن قضاوت و اظهار نظر می‌نمایند .

مراحل انجام عملیات آماری

هریک از سه نوع عملیات آماری فوق که مورد استفاده باشند باید حداقل دو مرحله‌از سه‌مرحله زیر طی شود تا نتیجه مطلوب بدست آید . این مراحل عبارت از جمع‌آوری وثبت اطلاعات ، محاسبه پارامترها و انجام قضاوتهای مربوط به نمونه ، و بالاخره تعمیم نتایج حاصل از نمونه به جامعه و یا نتیجه‌گیری در مورد جامعه بر مبنای نتایج حاصل از نمونه می‌باشند .

۱ ــ جمع آوری و ثبت اطلاعات : عملیات‌آماری با هر هدفی‌که انجام گیرد بدون‌شک نیاز به مواد اطلاعاتی دارد . به‌عنوان مثال میانگین دستمزد روزانه کارگران کارخانه‌های صنعتی ارج وقتی قابل محاسبه‌است که میزان دستمزد هریک از کارگران بر محقق معلوم‌باشد که این کار درمرحله جمع‌آوری و ثبت‌اطلاعات انجام می‌گیرد . جمع‌آوری اطلاعات رابرحسب گروه مورد مطالعه می‌توان به‌دو نوع جمع‌آوری اطلاعات از جامعه وجمع‌آوری اطلاعات‌ازنمونه تقسیم نمود .

الف ــ جمع‌آوری اطلاعات از جامعه : گاهی اوقات جامعه مورد مطالعه کوچک است و مراجعه به‌خود جامعه زحمت چندانی ندارد و یا اینکه علیرغم بزرگ بودن حجم جامعه و با وجود این که‌جمع‌آوری اطلاعات از خود جامعه متضمن زحمت بسیار زیاد و دشواریهای فراوانی می‌باشد ، با این وجود ضرورت ایجاب می‌کند که برای جمع‌آوری اطلاعات به‌خود جامعه اصلی مراجعه شود . در چنین مواردی اطلاعات جمع‌آوری شده دقیق و قابل اعتماد هستند .

ب ــ جمع‌آوری اطلاعات از نمونه : بطور معمول نام آمار همواره مسئله تقریب و غیر یقینی بودن نتایج مطالعات آماری را در اذهان متبادر می‌سازد . این احساس درستی است که در اثر استفاده‌وسیع روشهای آماری از نمونه بجای جامعه‌اصلی‌به‌افراد مختلف‌دست داده است . زیرا وقتی که در یک مطالعه آماری بجای رجوع به‌جامعه اصلی از آن جامعه یک یاچند نمونه اخذ شده و مطالعات برروی این نمونه‌ها متمرکز می‌گردند و سپس نتایج حاصله‌ازنمونه به‌جامعه تعمیم داده می‌شوند همواره این اشکال در ذهن وجود خواهد داشت که این نتایج

---

1 - Regressinon

فقط در مورد خود آن نمونه‌ها مصداق قطعی دارند و اگر نمونه‌های مأخوذه آئینه‌های تمام‌نمائی از مسائل و خصوصیات جامعه اصلی مورد نظر نباشند آنگاه نتایج مربوط به نمونه در مورد جامعه اصلی انطباق صد درصد نخواهد داشت بلکه فقط می‌توان گفت که این نتایج با درجه‌ای از احتمال در مورد جامعه صادق خواهند بود. نیاز به‌دانستن نظریه احتمال [1] برای فردی که می‌خواهد به‌حل مسائل آماری بپردازد نیز از اینجا ناشی می‌شود که اکثریت قریب به‌اتفاق مسائل آماری به‌دلیل فوق یعنی به‌خاطر استفاده از نمونه درمرحله جمع‌آوری اطلاعات به‌مقدار زیادی با مسائل احتمال درهم آمیخته و وجود ترکیبی اما واحدی‌راایجاد نموده‌اند. به‌همین دلیل در اولین فصل این جزوه نظریه احتمال تقریبا" به‌تفصیل وازابعاد گوناگون مورد بحث قرار می‌گیرد.

ازآنجا که اصولا" نمونه‌ای برای مطالعه آماری مطلوب است که خصوصیات آن با خصوصیات جامعه اصلی حداکثر تشابه و همانندی را دارا باشد تا بدین ترتیب نتایج نمونه با اطمینان بیشتری در مورد جامعه بکار گرفته شود لازم است که در هنگام نمونه‌گیری از روشهائی که تئوری نمونه‌گیری پیشنهاد می‌کند استفاده شود تا مناسب‌ترین نمونه ممکن از جامعه اخذ گردد. به‌همین دلیل وجود مبحث نمونه‌گیری و روشهای‌آن دردرس‌آمار ضرورت دارد. اما ازآنجا که این کتاب به‌منظور حل مسایل آماری تدوین گردیده و در هنگام حل این مسائل معمولا" افراد، با اطلاعات جمع آوری شده و ثبت شده سروکار داشته و با مسئله نمونه‌گیری مواجه نیستند و نیز به‌لحاظ رعایت اختصار، در اینجا از ورود به‌بحث روشهای نمونه‌گیری خودداری‌می‌شود. واضح است‌که علاقمندان به‌مطالعه این تئوری می‌توانندبه‌کتب و جزواتی که در این زمینه نوشته شده است مراجعه کنند.

۲ ــ محاسبه پارامترها و انجام قضاوتهای مربوطه: پس از ثبت و جمع‌آوری اطلاعات از نمونه و یا جامعه اصلی و انجام طبقه‌بندیهایی که برای سهولت محاسبات معمولا" انجام می‌گیرند، زمان آن فرا می‌رسد که پارامترهایی از قبیل میانگین، واریانس، انحراف معیار و غیره را در مورد نمونه محاسبه کرد. این پارامترها بعضی اوقات مستقلا" مطلوب بوده و خودشان هدف عملیات آماری می‌باشند زیرا ارزیابی و قضاوت مورد نظر تنها با وجود همین پارامترها میسر است. اما گاهی اوقات نیز پارامترهای موردمحاسبه نقش ابزار و واسطه را دارند مثلا" برای استفاده از روشهایی مانند تجزیه واریانس و رگرسیون و غیره باید این پارامترها قبلا"محاسبه شده باشند. در صورتی که محاسبه پارامترها مطلوبیت اخیر راداشته

---

1 - Theory of Probability

و مسئله مورد نظر استفاده از روشهایی از قبیل تجزیه واریانس و غیره را نیز برای قضاوت در مورد نمونه لازم داشته باشد پس از محاسبه پارامترها باید روش مناسب را انتخاب نموده و با استفاده از آن قضاوت لازم را در مورد نمونه انجام داد. اگر اطلاعات جمع‌آوری شده در مرحله اول که در مرحله دوم مورد استفاده قرار گرفته‌اند از جامعه اصلی باشند و از نمونه‌ای باشند که نتایج مربوط به آن به‌طور قطع در مورد جامعه نیز صدق کنند کار مطالعه آماری در مرحله دوم پایان می‌یابد. به‌عنوان مثال وقتی که در یک قطعه زمین به‌عنوان نمونه، اثر یک نوع کود با درصدی از احتمال بر افزایش میزان محصول گندم از طریق تجزیه واریانس و یا رگرسیون مورد تأیید قرار می‌گیرد می‌توان این حکم را در مورد تمامی آن کودها و همه زمینه‌های مشابه به‌عنوان جامعه تعمیم داد. اما اگر نمونه مورد مطالعه به‌گونه‌ای باشد که نتوان با قاطعیت نتیجه آنرا در مورد جامعه نیز پذیرفت در آن صورت کار مطالعه آماری به مرحله سوم کشیده خواهد شد.

۳ ــ **نتیجه‌گیری در مورد جامعه بر مبنای نتایج نمونه:** در بیشتر مواردی که مطالعه آماری صورت می‌گیرد نتایج حاصل از نمونه به‌دلایل مختلف در مورد جامعه به‌طور قطعی و یقینی صادق نبوده و نمی‌توان صرفاً" برمبنای نتایج حاصل از نمونه در مورد وضعیت جامعه اظهارنظر کرد. در چنین مواردی با استفاده از پارامترهای مربوط به نمونه به‌عنوان مواد خام و افزودن بعضی اطلاعات دیگر از قبیل شکل توزیع، روشهایی مانند تجزیه واریانس، آزمون فرضیه‌ها و تئوری تخمین بکار گرفته شده و با استفاده از آنها نتایجی در مورد جامعه بطور تقریبی و احتمالی بدست می‌آید.

بدین ترتیب مبحث نمونه‌گیری که در اینجا از طرح آن خودداری شده است مربوط به مرحله اول عملیات آماری می‌باشد. مباحث چگونگی محاسبه پارامترها و مشخصه‌های آماری مختلف، رگرسیون و تا حدودی مبحث تجزیه واریانس از مباحث اصلی مربوط به مرحله دوم این گونه عملیات است و از جمله عملیات اصلی مربوط به مرحله سوم یعنی نتیجه‌گیری در مورد جامعه بر مبنای نتایج نمونه نیز می‌توان از تجزیه واریانس، نظریه تخمین و آزمون فرضیه‌ها نام برد و بالاخره دو مبحث احتمالات و قواعد مربوط به توزیعهای مختلف نیز اگر چه جزء مباحث اصلی آمار نیستند اما اولی به‌دلیل وجود مسئله نمونه در عملیات آماری جزء ضروریات و مقدمات لازم عملیات آماری محسوب می‌گردد و دومی یعنی توزیعهای مختلف نیز چون کمک بسیار موثری در مرحله سوم و تا حدودی در مرحله دوم به عملیات آماری می‌کنند جزء مطالب بسیار مهمی هستند که در درس آمار مورد بحث قرار می‌گیرند. این کمک به‌قدری مهم است که در صورت نبودن آن، نتیجه‌گیری در مورد جامعه علیرغم وجود نتایج مربوط به نمونه غیر ممکن می‌باشد.

تقسیم‌بندی موضوعات در این کتاب

با توجه به‌مطالب فوق طبیعی است‌که باید در هر نوشته مربوط به‌آمار، ابتدا دومبحث احتمالات و اشکال مختلف توزیع آماری مورد بحث قرار گیرند زیرا مقدمات عملیات آماری محسوب می‌گردند به‌این علت در اینجا ابتدا در مورد احتمالات بحث می‌شود. از آنجا که در هر شکلی از توزیع، مشخصه‌هائی از قبیل میانگین و واریانس وجود دارند و طبعاً" باید قبل از ورود به‌بحث اشکال مختلف توزیع این مشخصه‌ها بخوبی شناخته شده باشند پس از احتمالات راجع به‌مشخصه‌های مهم آماری توضیحات لازم داده شده و در فصل بعد اشکال مختلف توزیع مورد بحث قرار می‌گیرند. و پس از آن در جلد دوم کتاب وارد مطالب و اصلی آمار شده و پیرامون موضوعات تخمین، آزمون فرضیه رگرسیون، تجزیه واریانس‌و غیره صحبت خواهیم کرد.

در برخورد کلی‌آمار با مسائل‌اقتصادی،بعضی ازمشخصه‌های اقتصادی ـ آماری‌بوجود آمده‌اند که از طرفی مطالعه آنها برای محققان و پژوهشگران اقتصاد ضروری بوده و از طرف دیگر این مباحث، مباحث صد درصد آماری نبوده بلکه دارای دو بعد اقتصادی و آماری می‌باشند و صحبت پیرامون آنها در قالب مباحث آماری مطلق چندان مناسب نیست لذا در انتهای این‌کتاب فصلی تحت‌عنوان شاخصها آورده‌شده به‌توضیح پیرامون این‌مشخصه‌ها می‌پردازد. امید است که با تأیید خدای متعال مطالب ارائه شده از وضوح و کمال لازم برخوردار بوده و مورد استفاده کامل خواننده محترم قرار گیرند

فصل اول

تئوری احتمالات[1]

مقـــدمـه:

همانطورکه قبلا" گفته شد در بیشتر موارد وقتی که قرار است از طریق عملیات آماری نسبت به‌بعضی‌از خصوصیات و ویژگیهای یک جامعه نتیجه‌گیری و قضاوت شود به‌دلایل مختلفی ترجیح داده می‌شود که در مرحله جمع‌آوری اطلاعات بجای مراجعه به‌جامعه اصلی و کسب اطلاعات لازم از طریق مطالعه وضع همه افراد جامعه، نمونه و یا نمونه‌هائی رااز‌آن جامعه گرفته و این نمونه‌ها را مورد مطالعه قرار دهند. پس از اتمام این مطالعات، بـرروی اطلاعات جمع‌آوری شده از نمونه عملیاتی انجام گرفته و مشخصه‌ها و پارامترهای مورد نیاز مربوط به‌نمونه محاسبه‌می‌گردند. درمرحله بعد براساس نتایج مربوط به‌نمونه، با بهره‌گیری از بعضی از قوانین از قبیل قواعد مربوط به‌شکل توزیع و غیره با توسل به‌روشهایی مـانـنـد آزمون فرضیه‌ها و تخمین، در مورد جامعه اصلی نتیجه‌گیری بعمل آمده و قضاوت نـهـایـی صورت می‌گیـرد.

دلایل مختلفی برای رجحان نمونه بر جامعه می‌توانند وجود داشته باشند. دربعضی از موارد جامعه موردنظربه‌قدری بزرگ است که مراجعه به تک تک افراد آن مستلزم صـــرف هزینه‌های گزاف و وقت بسیـار می‌باشد که در بیشتر موارد فرصت و بودجه لازم در اختیار محقق قرار ندارد و یا آنکه ضرورتی ندارد که نتیجه مطالعه کاملا" دقیق بوده و کوچکتریـن خطایی در آن وجود نداشته باشد بلکه اگر بتوان با روشهای ساده‌تر و کم هزینه‌تری عمـل نموده و به‌طریقی اطمینان حاصل نمود که نتیجه حاصله به‌طور تقریبی و فقط تا حدودی بـا واقعیات جامعه موردنظر انطباق داردمقصود محقق از مطالعه آماری حاصل می‌گردد. ازآنجا که معمولا" مطالعه آماری از طریق نمونه‌گیری می‌تواند چنین نتایج تقریبی را بدست دهد.

---

1 - Theory of Pribability

مراجعه به‌جامعه اصلی برای کسب اطلاعات با توجه به‌مشکلات فراوان آن ضرورت خود را از دست خواهد داد.

به‌عنوان مثال فرض کنیم وزارت آموزش و پرورش می‌خواهد در مورد میزان درآمدهای جنبی معلمان و دبیران شاغل در سطح شهرتهران مطالعه نموده و بر اساس نتیجه مطالعه در میزان حقوق پرداختی خود به‌آنان تغییراتی ایجاد نماید. با توجه به‌جمعیت چندین ده هزار نفری جامعه مورد نظر مسلما" برای انجام چنین مطالعه‌ای، بسیار مشکل خواهد بود که تمام افراد جامعه اصلی یعنی همه معلمان و دبیران تهرانی مورد مطالعه قرارگیرند و به‌دلایل مختلفی از قبیل سرعت تحقیق، هزینه مطالعه و غیره استفاده از نمونه بهتر و مناسب‌تر خواهد بود.

در بعضی موارد همه افراد جامعه اصلی، مطلقا" قابل دسترسی نیستند و یا به دلایل دیگری مطالعه همه افراد جامعه نامعقول است. مثلا" وقتی قرار است در مورد کیفیت محصولات یک نوع بخصوص از پودر لباسشویی مثلا" پودر تاید در ۵ سال گذشته مطالعه شود عملا" فقط تعداد محدودی از محصولات تولید شده در طی ۵ سال گذشته در دسترس بوده و بیشتر محصولات تولیدی به‌مصرف رسیده و از دسترس آمارگر خارج شده‌اند. بهمین ترتیب در برخی موارد مطالعه همه افراد جامعه اصلی امری نامعقول بوده و در چنین مواردی الزاما" باید از روش نمونه‌گیری استفاده نمود. بارزترین نمونه این نوع جوامع، جامعه مربوط به‌بعضی از محصولات عکاسی از قبیل فیلم و کاغذ می‌باشد. واضح است که در هنگام مطالعه کیفیت یک فیلم عکاسی با مارک بخصوص نمی‌توان همه فیلمهای تولید شده را مورد مطالعه قرار داد زیرا در هنگام مطالعه، فیلمی که تحت مطالعه می‌باشد خراب شده و قابلیت انتفاع خود را از دست می‌دهد. دلایل دیگری نیز وجود دارند که باعث شوند در مطالعه آماری بجای مراجعه به‌جامعه اصلی در مرحله جمع‌آوری اطلاعات از نمونه استفاده شود که جهت اختصار از بیان آنها خودداری می‌گردد.

استفاده از نمونه در مطالعات آماری به‌هر دلیل که صورت گیرد باعث می‌گردد که نتیجه بدست آمده در مورد جامعه اصلی که بر مبنای نتایج یقینی مربوط به‌نمونه حاصل شده است همواره با تقریب و احتمال همراه باشد. به‌عنوان مثال در تئوری تخمین معمولا" گفته می‌شود حال که میانگین درآمد جنبی معلمان و دبیران شاغل در تهران که در نمونه مورد مطالعه وجود داشتند مقدار(X) را داراست با ۹۰٪ احتمال می‌توان گفت که میانگین درآمدهای جنبی همه افراد این جامعه بین دو رقم Y و Z قرار خواهند داشت. به‌همین ترتیب در مباحث آزمون فرضیه‌ها، تجزیه واریانس و غیره به‌دلیل استفاده از نمونه مجبور به قضاوت احتمالی در مورد جامعه خواهیم بود.

وجود مسئله تقریب و احتمال در عملیات مختلف آماری ایجاب می‌کند که مفهوم احتمال و شیوه‌های محاسبه آن در ضمن درس آمار تدریس شود زیرا اگرچه موضوع احتمال یک موضوع آماری نیست اما در روشهای آماری مختلف یکی از ارکان اساسی بحثها را تشکیل می‌دهد و معمولاً" بطور مستقل و یا ضمن هیچیک از دروس دانشگاهی نیز تدریس نمی‌شود و از همین روست که این فصل کتاب به‌بحث پیرامون احتمال و روشهای محاسبه آن اختصاص یافته‌است .

در این فصل ابتدا در مورد مفهوم احتمال و بعضی خصوصیات آن صحبت شده و سپس روش کلی محاسبه احتمال حوادث مختلف مورد بحث قرار می‌گیرد . از آنجا که معمولاً" حوادث و وقایع در آمار به دو دسته ساده و مرکب تقسیم می‌گردند و محاسبه احتمال حوادث ساده ، به‌سادگی قابل انجام بوده و در مقابل محاسبه احتمال حوادث مرکب در اغلب موارد دارای پیچیدگی‌های مختلفی می‌باشد ما نیز بهتر دانستیم این دو نوع حادثه را تفکیک نموده و پیرامون روش محاسبه احتمال هریک بصورت مستقل بحث کنیم . بدین ترتیب ادامه بحث در این فصل در عناوین کلی زیر صورت خواهد گرفت :

الف : مفهوم احتمال

ب : محاسبه احتمال حوادث مختلف

۱ ــ محاسبه احتمال حوادث ساده

۲ ــ محاسبه احتمال حوادث مرکب

قبل از ورود به‌بحث مفهوم احتمال یادآوری این نکته ضرورت دارد که چون استفاده از مثالهای مربوط به‌بعضی بازیها از قبیل پرتاب تاس و غیره به‌درک مفاهیم مربوط به‌احتمال کمک می‌کند علیرغم میل باطنی مجبور به‌استفاده از این مثالها شده و ازاین جهت از محضر خواننده عزیز پوزش می‌طلبیم .

الف ــ مفهوم احتمال

معمولاً" به‌درجه‌امکان وقوع یک‌حادثه ،احتمال وقوع آن حادثه گفته می‌شود . میزان احتمال وقوع یک حادثه به‌شرایط مختلف موجود در هنگام وقوع آن بستگی دارد . به‌عنوان مثال میزان احتمال آمدن شیر در پرتاب سکه‌ای که دو طرف آن از نظر خصوصیات فیزیکی یکسان هستند از میزان احتمال آن در هنگامی که یک سکه ناصاف که خصوصیات فیزیکی دو طرف آن باهم یکسان نیستند پرتاب می‌شود متفاوت‌است .همچنین احتمال‌تولید محصولات بی‌کیفیت در زمانی‌که بحران اقتصادی بر یک‌جامعه حکومت می‌کند با همین احتمال‌درشرایط عادی اقتصادی قطعاً"مساوی نخواهد بود .به‌همین ترتیب در مورد هر واقعه‌ای مسلماً"عوامل بسیاری وجود دارند که با درجات مختلفی از شدت و ضعف بر میزان احتمال وقوع‌آن‌حادثه

اثر می‌گذارند .

تعریفی که در پاراگراف قبلی از احتمال داده شد ، تعریفی آماری نبوده و ماهیتی فلسفی دارد . در آمار معمولاً" احتمال چنین تعریف می‌گردد .

" اگر در یک آزمایش چند حادثه مختلف امکان وقوع داشته باشند "
" که یکی از آنها مورد نظر و مطلوب آزمایش کننده باشد در صورتیکه "
" آن آزمایش بقدری زیاد تکرارشود که در این‌تکرارها وقایع شانسی یکدیگر را خنثی کنند "
" نسبت دفعات وقوع حادثه مطلوب به‌کل حوادث واقع شده را احتمال وقوع حادثه مطلوب نامند . "

بدین ترتیب اگر تعداد دفعاتی را که آزمایش تکرار می‌شود با n نشان داده و تعداد دفعاتی که حادثه مطلوب در n بار آزمایش اتفاق افتاده است با m نشان داده شود احتمال وقوع آن حادثه به‌صورت رابطه ریاضی ذیل قابل تعریف است .

(۱ — ۱)

$$P=\frac{m}{n} \quad \text{احتمال حادثه مورد نظر}$$

برای توضیح بیشتر فرض کنید سکه‌ای در اختیار شماست که درموردخصوصیات فیزیکی دوطرف آن اطلاعات چندانی ندارید و می‌خواهید بدانید احتمال رویت هریک از دو طرف سکه پس از یک پرتاب چقدر است . برای نیل به‌این هدف با پرتاب سکه شروع به‌آزمایش نموده و پس از ده بار پرتاب ملاحظه می‌کنید که ۶ مرتبه خط و بقیه شیر آمده است یعنی احتمال وقوع خط در این آزمایش ۰/۶ می‌باشد . آیا این احتمال ۰/۶ را به عنوان نتیجه نهائی قبول‌کرده و برای شما اطمینان حاصل می‌شود که در مورد این سکه خاص نتیجه‌حاصله (۰/۶)قطعیت دارد ؟ مسلماً" چنین نیست و شما با خود می‌گوییدکه‌احتمال اینکه‌این‌نتیجه شانسی بوده و در دفعات بعدی نتیجه‌ای بسیار متفاوت با این نتیجه عاید شود بسیار زیاد است . برای آنکه به‌نتیجه‌ای مطمئن‌تر دست یابید آزمایش را تکرار نموده و بطور طبیعی‌فکر می‌کنید که احتمال بدست آمده پس از تکرار بسیار زیاد آزمایش به‌واقعیت نزدیکتر است . بر اساس این تفکر بالاخره پس از انجام تعداد زیادی آزمایش به‌نقطه‌ای می‌رسیدکه احساس می‌کنید دیگر آزمایش کافی بوده و آزمایشهای انجام شده از نظر تعداد آنقدر زیاد هستند که در مورد حصول یک نتیجه تقریباً" واقع نما به‌انسان اطمینان خاطر دهند .به‌عبارت‌دیگر در نتیجه تکرار به‌دفعات زیاد آزمایش این احساس به‌افراد دست می‌دهد که اگر به‌طورمثال در ۵۰ بار تکرار پرتاب سکه ده مرتبه به‌صورت شانسی و تحت تأثیر عوامل خارجی از قبیل

وضعیت دست پرتاب‌کننده، جریان هوا و غیره شیرآمده‌است همین تعداد هم‌بصورت‌شانسی روی خط سکه وقوع‌یافته و این ۲۰ حادثه شانسی یکدیگر را خنثی نموده‌اند و در ۴۸۰پرتاب دیگر حوادث واقع شده نشانگر احتمال واقعی هریک از دو حادثه ممکن‌الوقوع (شیر یا خط) در پرتاب یک سکه است . و چون حوادث شانسی نیز چنان که گفته شد تا حدودی بایکدیگر خنثی شده‌اند با تقریب زیادی می‌توان گفت که در این تعداد انجام آزمایش این حوادث شانسی بر نتیجـــه آزمایش بی‌اثرند و لهذا نتیجه محاسبه شده که فرضاً" :

$$P_{خ}^{(1)} = \frac{۲۰۰}{۵۰۰} = ۰/۴ \qquad \text{احتمال آمدن خط در پرتاب سکه}$$

می‌باشد برای آزمایش کننده نتیجه‌ای تقریبا" قطعی تلقی خواهد شد .

ممکن است.این سئوال در ذهن خواننده نقش ببندد که آیا تنها راه تشخیص احتمال وقوع یک حادثه همین انجام آزمایش به‌دفعات بسیار زیاد است ؟ وآیا راه ساده‌تری بـــرای این تشخیص وجود ندارد ؟ در پاسخ به‌این سئوال باید گفت‌که اگرچه بهترین و مطمئن‌ترین راه نیل به‌هدف فوق تکرار زیاد آزمایش و محاسبه احتمال به‌طریق تجربی می‌باشداما از راه دیگری نیز می‌توان به‌این هدف نایل گردید و آن مطالعه عواملی است که وقوع یک حادثه‌را سبب می‌شوند بدین معنی که اگر عوامل موءثر بر وقوع دو یا چند حادثـه با یکدیگـر کاملا مشابه باشند آن‌گاه‌آن حوادث شانس کاملا"مساوی برای واقع شدن خواهند داشت . به‌عنوان مثال اگر خصوصیات فیزیکی دوطرف یک‌سکه مشابه هم باشند با اطمینان خاطر می‌توان‌گفت که احتمال وقوع هریک از دو روی سکه در نتیجه یک پرتاب مساوی روی دیگر آن خواهدبود. به‌همین ترتیب موردی را در نظر بگیرید که می‌خواهید از بین دو شیی، مثلا" دوکتاب که‌از نظر خصوصیات ظاهری کاملا" شبیه یکدیگرند یکی را که متعلق به‌دوستتان است و شما آن را نمی‌شناسید اما خود او از روی قراین و علامتهایی که در داخل کتاب است آنرا می‌شناســد انتخاب‌کنید . آیا دراینجا احتمال انتخاب هریک از دوکتاب فوق برای شما مساوی نیست ؟ مسلما" چرا ،علت آن است که عامل موءثر در انتخاب شما خصوصیات ظاهری کتاب است که این خصوصیات در مورد هر دو کتاب کاملا" یکسان است .

نتیجه‌ای که از بحث فوق گرفته می‌شود این است که اگر به‌صورت ظاهر، عوامل موءثر بر وقوع دو یا چند حادثه باهم مشابه باشند می‌توان وقوع آن حوادث را متساوی الاحتمـال فرض نموده و اگر این عوامل باهم مختلف باشند به‌میزان اختلاف، احتمال وقوع آن وقایع

---

۱ ــ مقصود از $P_{خ}$ احتمال وقوع خط و مقصود از $P_{ش}$ احتمال وقوع شیر در آزمایش پرتـاب یک سکه می‌باشد .

نیز باهم اختلاف خواهند داشت .

برای آنکه مطلب فوق بازهم از وضوح بیشتری برخوردار گردد ذیلا" دو مثال دیگر که یکی از آنها مربوط به‌پیشامدهای متساوی‌الاحتمال و دیگری مربوط به‌وقایع مختلف‌الاحتمال است ذکر می‌گردد . اولین مثال مربوط به‌یک تاس همتراز است واضح است که احتمال رؤیت هریک از ۶ جانب چنین تا سی با دیگر جوانب آن مساوی می‌باشد و به‌عنوان مثال دوم اگر فرض کنیم کیسه‌ای محتوی ۴ مهره قرمزرنگ و دو مهره به‌رنگ مشکی بوده و این مهره‌هاازنظر سایر خصوصیات کاملا" مشابـــه می‌باشند ، احتمال انتخاب تصادفی یک مهره قرمز از ایـــن کیسه ، با احتمال انتخاب یک مهره سیاه مساوی نبوده بلکه دوبرابر آن خواهد بود .

خصوصیات مهم احتمال

برای احتمال ویژگیهای مختلفی را می‌توان بر شمرد که دوتا از مهمترین آنها ذیلا"در قالب دو قضیه بیان می‌گردند .

قضیه۱ــ۱ : مقدار احتمال یک حادثه حداکثر مساوی یک و حداقل مساوی صفر بـــوده و مطمئنا" این مقداربزرگتر از یک وکوچکترازصفرنخواهدگردید . البته در بیشتر موارد مقدار احتمال بین صفر و یک بوده و مواردی که‌احتمال‌یک شیئی در نقاط حدی باشد معمـــولا" از دایره بحث احتمال خارج می‌باشد زیرا اگر احتمال وقوع یک حادثه برابر یک باشد یعنی‌آن حادثه یقینا" واقع شده و متقابلا" اگر احتمال وقوع حادثه‌ای مساوی صفر باشد یقینا" واقع نخواهد گردید . اثبات این قضیه که مقدار احتمال همواره بین صفر و یک است بسیار آسان است بدین ترتیب که مثلا" درآزمایش n بار پرتاب یک سکه ممکن است حداکثر تمامی‌موارد شیر باشد . بنابراین حداکثر مقدار احتمال شیر :

$$P_{\text{ش}} = \frac{n}{n} = ۱$$

خواهد بود . متقابلا" امکان دارد که هیچیک از n بار پرتاب شیر نباشد و در نتیجه‌حداقل مقدار احتمال شیر :

$$P_{\text{ش}} = \frac{۰}{n} = ۰$$

خواهد بود اما واقعی‌تر آنست که گفته شود از n بار پرتاب یک سکه m مرتبـــه شیر و n ــ m مرتبه خط می‌آید . $(n > m)$ و نتیجتا" مقدار احتمال آن شیر :

$$\circ < P_{ش} = \frac{m}{n} < ۱$$

خواهدبود . بدین ترتیب رابطه :

$$\circ < P < ۱ \qquad (۱-۲)$$

اثبات می‌گردد .

قضیه ۲ ــ ۱ : مجموع احتمالات مربوط به‌هریک از حوادث ممکن‌الوقوع در یک‌آزمایش مساوی یک می‌باشد . به‌عنوان مثال مجموع دو احتمال مربوط به‌وقوع شیر و خط در آزمایش پرتاب سکه مساوی یک است . برای اثبات این قضیه فرض کنید در نتیجه n بار پرتاب یک سکه m مرتبه شیر و m − n مرتبه خط آمده است n > m در نتیجه داریم :

$$P_{ش} = \frac{m}{n} \qquad و \qquad P_{خ} = \frac{n-m}{n}$$

که حاصل جمع آنها برابر :

$$P_{ش} + P_{خ} = \frac{m}{n} + \frac{n-m}{n} = \frac{n}{n} = ۱$$

خواهد شد و در نتیجه قضیه مورد نظر اثبات می‌گردد .

ب ــ محاسبه احتمال حوادث مختلف

همان‌طورکه در مقدمه این فصل گفته شد حوادثی‌که در آمار ازاحتمال وقوعشان صحبت می‌شود به‌دو نوع ساده و مرکب قابل تقسیم می‌باشند . ذیلا" هریک از این دو نوع حادثه تعریف شده و با ذکر مثالهایی توضیح داده می‌شوند . اما قبل از این تعاریف ، لازم است ابتدا مفهوم فضای نمونه که در آن تعاریف مورد استفاده قرار می‌گیرد بیان شده و با ذکر مثالهایی روشن گردد .

" اگر هریک‌از حوادث ممکن‌الوقوع در یک‌آزمایش را بصورت‌نقطه‌ای دریک مجموعه "
" نشان دهیم مجموعه‌ایجادشده را فضای نمونه و هریک از آن نقاط را یک‌نقطه نمونه نامند . "

مثال ۱ــ۱ . فضای نمونه مربوط به‌آزمایش پرتاب یک سکه را بصورت ریاضی و هندسی نشان دهید .

پاسخ :

S = { خـ , شـ }

مثال ۲ــ۱ . فضای نمونه مربوط به آزمایش پرتاب همزمان دو سکه را نشان دهید .
پاسخ : این فضا شامل ۴ نقطه است که شامل حوادث اولی و دومی هردو شیر ، اولی شیر و دومی خط ، اولی خط و دومی شیر و بالاخره هردو خط می‌باشد .

S = { (خـ , خـ) , (شـ , خـ) , (خـ , شـ) , (شـ , شـ) }

مثال ۳ــ۱ . فضای نمونه مربوط به آزمایش همزمان دو تاس را نشان دهید .
پاسخ : این فضا شامل ۳۶ نقطه بصورت زیر است :

(۶ و ۱) و (۵ و ۱) و (۴ و ۱) و (۳ و ۱) و (۲ و ۱) و (۱ و ۱)
(۶ و ۲) و (۵ و ۲) و (۴ و ۲) و (۳ و ۲) و (۲ و ۲) و (۱ و ۲)
(۶ و ۳) و (۵ و ۳) و (۴ و ۳) و (۳ و ۳) و (۲ و ۳) و (۱ و ۳)
(۶ و ۴) و (۵ و ۴) و (۴ و ۴) و (۳ و ۴) و (۲ و ۴) و (۱ و ۴)
(۶ و ۵) و (۵ و ۵) و (۴ و ۵) و (۳ و ۵) و (۲ و ۵) و (۱ و ۵)
(۶ و ۶) و (۵ و ۶) و (۴ و ۶) و (۳ و ۶) و (۲ و ۶) و (۱ و ۶)

شکل ۱ــ۱ . فضای نمونه مربوط به‌پرتاب همزمان دو تاس

حال پس از تعریف مفهوم فضای نمونه و روشن شدن این مفهوم از طریق چند مثال فوق به‌بحث اصلی بازگشته و حادثه ساده را بصورت زیر تعریف می‌کنیم .

" یک حادثه ساده ، حادثه‌ای است که به‌حوادث دیگری قابل تجزیه نباشد . "

" در یک فضای نمونه هریک از نقاط نمونه نشان دهنده یک حادثه ساده می‌باشند "

آمدن شیر در پرتاب یک سکه ، آمدن ۲ شیر در پرتاب همزمان دو سکه ، آمدن ۲ در پرتاب یک تاس ، آمدن ۶ و ۶ در پرتاب همزمان دو تاس ، انتخاب یک کتاب بخصوص از بین تعدادی کتاب مشابه ، انتخاب یک اتومبیل از بین دهها اتومبیل مشابه و غیره و غیره تماما " نمونه‌هایی از حوادث ساده می‌باشد .

نکته قابل توجه و بسیار مهمی که در مورد حوادث ساده وجود داشته و می‌تواند به‌عنوان پایه‌ای برای محاسبه احتمال حوادث بکار رود این است که در غالب موارد ، همه حوادث ساده‌ای که در یک آزمایش یا یک انتخاب امکان وقوع دارند از شانس مساوی برای

وقوع برخوردار هستند و اگر در مواردی هم این تساوی وجود ندارد سعی می‌کنند به طریقی آن را بوجود آورند . به‌عنوان مثال در آزمایش انتخاب یک کتاب از بین دو کتاب با ظواهر کاملا" مشابه احتمال انتخاب هریک از آنها مساوی احتمال انتخاب دیگری می‌باشد . اما در آزمایش پرتاب سکه هیچ دلیلی بر تساوی احتمال وقوع دوروی سکه با یکدیگر وجود ندارد . در چنین مواردی معمولا" فرض می‌کنند که سکه همتراز بوده‌و در نتیجه دو طرف آن از نظر وقوع متساوی الاحتمال می‌باشند و قس علی‌هذا .

" یک حادثه مرکب ، حادثه‌ای است که به چند حادثه ساده قابل تجزیه باشد "

" به‌عبارت دیگر حادثه مرکب حادثه‌ای است‌که شامل دو یا چند نقطه درفضای‌نمونه باشد . "

آمدن یک‌خط و یک‌شیر در پرتاب همزمان دو سکه، آمدن مجموع ۷ در پرتاب همزمان دو تاس، انتخاب یک‌مهره‌قرمز از کیسه‌ای که‌مثلا" ده‌مهره‌مشابه‌دارد که‌سه‌تای آنها قرمزاست، انتخاب یک دانشجوی سال سوم از کلاسی که ۶ نفر از سی نفر دانشجویان آن سال سوم می‌باشند و مثالهایی‌ازاین قبیل همه و همه نمونه‌هایی ازحوادث‌مرکب‌می‌باشند زیرا هریک از آنها به دو یا چند حادثه ساده قابل تجزیه‌اند . بدین ترتیب که حادثه مطلوب در پرتاب دو سکه به‌دو صورت قابل وقوع است ، یکی اینکه، سکه الف شیر و سکه ب خط بیاید و دیگری‌اینکه سکه الف خط و سکه ب شیر بیاید که هریک از این دو حالت ، حوادثی ساده هستند زیرا به‌حوادث ساده‌تری قابل تجزیه نیستند . بهمین ترتیب در پرتاب دو تاس الف و ب حادثه مطلوب ممکن است‌به‌یکی از ۶ صورت (۱ و ۶) ، (۲ و ۵) ، (۳ و ۴) ، (۴ و ۳) ، (۵ و ۲) و (۶ و ۱) وقوع یابد که در هر پرانتز ، عدد اول مربوط به‌تاس الف و عدد دوم به‌تاس ب مربوط می‌باشد . به‌هریک از این پرانتزها یک زوج مرتب گفته می‌شود و هریک از زوجهای مرتب فوق یک حادثه ساده هستند زیرا به‌حوادث ساده‌تری قابل تجزیه نیستند . و باز به‌همین ترتیب واقعه انتخاب مهره قرمز از کیسه‌حاوی مهره‌ها ممکن است به‌صورت انتخاب مهره قرمز الف یا انتخاب‌مهره قرمز ب و یا انتخاب مهره قرمز ج اتفاق بیفتد که هریک از این سه اتفاق یک حادثه ساده هستند . بالاخره‌حادثه انتخاب دانشجوی سال سوم از آن کلاس نیز به ۶ حادثه ساده‌قابل تقسیم است . انتخاب دانشجوی سال سوم شماره ۱ ، سال سوم شماره ۲ و ۰۰۰۰ و سال سوم شماره ۶ ، که مجموع این ۶ حادثه یک حادثه مرکب را تشکیل می‌دهند .

همان طوری که در مقدمه این فصل گفته شد چون طرز محاسبه احتمال یک حادثه‌ساده با احتمال یک حادثه مرکب تفاوت دارد و از طرفی روش محاسبه احتمال حادثه ساده پایه محاسبه‌احتمال حادثه مرکب‌است در اینجا ابتدا طرزمحاسبه احتمال حادثه ساده را توضیح داده و سپس در مورد شیوه محاسبه احتمال حادثه مرکب بحث می‌کنیم .

۱ ـ محاسبه احتمال حوادث ساده

یکی از مسائلی که همواره در حل مسائل مربوط به احتمال برای افراد مبتدی ایجاد مشکل می‌کند تشخیص حادثه ساده و حادثه مرکب می‌باشد. شکل متداولی از این مشکل این است که در بسیاری موارد اشکال ساده‌ای از حوادث مرکب با یک حادثه ساده اشتباه می‌شوند. برای اجتناب از چنین اشتباهاتی کافی است به تعریف حوادث ساده و مرکب توجه نموده و از آنها به عنوان معیار تشخیص هریک از این دو نوع حادثه استفاده نمود.

تذکر نکته مهمی در مورد احتمال حوادث ساده لازم است و آن این که معمولاً" افراد احتمال بسیاری از حوادث ساده و اشکال ساده‌تری از حوادث مرکب را به دلیل سهولت آنها، بصورت ذهنی حساب می‌کنند. به عنوان مثال اکثراً" بدون انجام هیچ محاسبه خاصی گفته می‌شود که احتمال آمدن هرروی یک سکه همتراز در یک پرتاب، $\frac{۱}{۲}$ و احتمال آمدن هریک از وجوه شش گانه یک تاس همتراز $\frac{۱}{۶}$ می‌باشد و به همین ترتیب احتمال انتخاب یک حلب روغن نباتی خاص از بین ده حلب مشابه $\frac{۱}{۱۰}$ می‌باشد و غیره و غیره.

اما اگر بخواهیم صرفنظر از مطلب فوق احتمال یک حادثه ساده را بر روی کاغذ محاسبه کنیم نیز کار دشواری در پیش نخواهیم داشت و می‌توان بسادگی احتمال مزبور را محاسبه نمود. برای این کار کافی است به خاطر آوریم که در جریان بحث پیرامون مفهوم حادثه ساده گفته شد که معمولاً" در یک آزمایش، کلیه حوادث ساده ممکن الوقوع از احتمال متساوی برخوردارند بنابراین اگر در یک آزمایش n حادثه امکان وقوع داشته باشد و به عبارت دیگر فضای نمونه این آزمایش دارای n نقطه به صورت زیر باشد:

$$S = \left\{A_۱, A_۲, \dots\dots\dots A_n\right\}$$

می‌توان مجموع احتمالات مربوط به این حوادث ساده را به دلیل تساوی احتمالاتشان بصورت زیر نوشت:

$$P_{A1}+P_{A2}+\dots\dots\dots+P_{A_i}+\dots\dots\dots+P_{A_n} = nP_{A_i}$$

ازطرفی طبق قضیه ۲ ـ ۱ می‌دانیم که مجموع احتمالات مربوط به کلیه حوادث در یک آزمایش برابر واحد است از تعمیم این قضیه به مورد فوق جمله زیر بدست می‌آید.

$$nP_{A_i} = 1$$

در نتیجه احتمال هریک از حوادث ساده ممکن الوقوع در این آزمایش بصورت رابطه (۳ـ۱)

که یک رابطه کلی است بدست می‌آید .

(۱-۳)

$$P_{A_i} = \frac{1}{n}$$

بعضی از حوادث ساده وقایعی دو یا چند مرحله‌ای بوده و بطور طبیعی محاسبه آنها پیچیده‌تر و مشکل‌تر از سایر موارد است . مثل پیشامد آمدن دو شیر و سه شیر و .... و n شیر که به ترتیب در پرتاب همزمان دو سکه، سه سکه و .... n سکه مورد آزمایش قرار می‌گیرند . محاسبه احتمال وقوع این گونه پیشامدها از طریق استفاده از رابطه ۳ ممکن است اما در بسیاری از موارد با مشکلات زیادی همراه است .برای اجتناب از این مشکلات روش ساده‌تری برای عمل درچنین مواردی بکار گرفته می‌شود . از آنجا که در این روش از مسئله‌ای بنام احتمال شرطی صحبت می‌شود ابتدا مختصرا" در مورد احتمال شرطی توضیحاتی داده و سپس محاسبه احتمال حوادث ساده چند مرحله‌ای مورد بحث قرار می‌گیرد .

احتمال شرطی

کاربرد اصلی مسئله احتمال شرطی در محاسبه احتمال حوادث ساده و یا مرکب چند مرحله‌ای می‌باشد اما درمواردی نیز احتمال شرطی مطلوبیت نفسی داشته و محاسبه آن به‌خودی خود مورد نظر است و در این موارد احتمال شرطی به‌عنوان یکی از مراحل محاسبه احتمال حوادث دیگری محسوب نمی‌گردد .در علم آمار این‌گونه احتمالات شرطی‌که مستقلا" مطلوبیت دارند عمدتا" در قالب رابطه بیس[1] مطرح می‌شوند . از آنجا که در این مرحله از بحث احتمال شرطی به‌عنوان یکی از مراحل محاسبه احتمال حوادث چند مرحله‌ای موردنیاز است لـــذا بصورت مختصر و فقط در حدودی که هدف مزبور تأمین گردد به آن پرداخته می‌شود و شکل پیچیده‌تر آن یعنی رابطه و یا قضیه بیس در خاتمه مبحث احتمالات مطرح می‌شود زیرا در این رابطه احتمالات نسبتا" پیچیده‌ای مورد استفاده قرار می‌گیرد که باید قبلا" توضیح داده شود و به‌همین دلیل قضیه بیس پس از توضیح روش‌های محاسبه همه انواع حوادث مورد بحث قرار می‌گیرد .

احتمال شرطی را بصورت زیر می‌توان تعریف نمود .

" احتمال وقوع حادثه B وقتی‌که بدانیم حادثه A اتفاق افتاده است رااحتمال‌شرطی" نامیده و آن را بصورت $P_{(B \mid A)}$ نشان می‌دهند . "

---

1 - Bayes Formula

جمله $P_{(B \mid A)}$ بصورت " احتمال‌وقوع‌حادثه B بشرط‌آنکه‌حادثه A قبلا"اتفاق‌افتاده باشد " خوانده می‌شود .

مثال ۴ـ۱ : در کیسه‌ای سه مهره مشابه وجود دارد که یکی از آنها برنگ سبز و دومی قرمز و بالاخره مهره سوم زردرنگ است . اگر از این کیسه یک مهره بیرون بیاوریم وآن مهره قرمز رنگ باشد احتمال این که مهره انتخابی دوم زرد رنگ باشد چقدر است .

پاسخ : قبل از اینکه مهره‌ای را از کیسه بیرون بیاوریم احتمال انتخاب هریک از سه مهره $\frac{۱}{۳}$ است اما پس از آنکه مهره قرمز بیرون می‌آید در کیسه فقط دو مهره باقی می‌مانـد و لـذا احتمال انتخاب مهره زرد رنگ $\frac{۱}{۲}$ خواهد شد .

$$P \text{ (زرد | قرمز) } = \frac{۱}{۲}$$

همچنانکه از مثال فوق بر می‌آید مسئله احتمال شرطی اصولا" وقتی قابل طرح است‌که حوادث مورد بحث به‌یکدیگر وابستگی داشته و از هم مستقل نباشند . هرگاه وقوع‌یک حادثه به‌نحوی‌احتمال وقوع حادثه دیگری را تحت تأثیر قرار دهد آن دو حادثه از هـم مستقـل نبوده وبرعکس‌هرگاه وقوع حادثه‌ای به‌هیچ وجه دراحتمال وقوع دیگری موءثر نباشد آن دو حادثه مستقل از یکدیگر خوانده می‌شوند .بدین ترتیب‌دو حادثه مورد بحث در مثال فـوق از هم مستقل نیستند وبه‌همین دلیل احتمال شرطی در مثال فوق قابل طرح می‌باشد .

مثال ۵ ـ ۱ : در صورتی که در آزمایش مثال (۴ـ۱) مهره قرمزی را که در مرتبه اول‌انتخاب می‌شود مجددا" به‌داخل کیسه برگردانیم احتمال انتخاب مهره زردرنگ در انتخـاب دوم چقدر است ؟

پاسخ : با جایگذاری مهره انتخاب شده‌اول ، حادثه اول و دوم از یکدیگر کاملا" مستقل شده و دیگر احتمال مورد سئوال این مثال احتمال شرطی نیست و احتمال انتخاب مهره زردرنگ در انتخاب دوم همانند احتمال انتخاب مهره قرمز رنگ در انتخاب اول $\frac{۱}{۳}$ است .

$$P \text{ زرد } = P \text{ قرمز } = \frac{۱}{۳}$$

محاسبه احتمال شرطی در شکل ساده‌ای که فعلا" موردنیاز است معمولا"بصورت ذهنـی انجام‌می‌گیرد روش محاسبه‌آن نیز بدین ترتیب‌است که فضای‌نمونه را پس از وقوع حادثه‌ای که به‌عنوان شرط در نظر گرفته شده است مشخص کرده و سپس احتمال حادثه مورد نظـر را همانند احتمال سایر حوادث ساده محاسبه می‌نمایند .

حال پس از ذکر این توضیحات پیرامون احتمال‌شرطی ، شرایط‌برای بحث پیرامـون احتمال حوادث ساده چند مرحله‌ای فراهم شده و لذا به‌این بحث می‌پردازیم .

محاسبه احتمال حوادث ساده چند مرحله‌ای

شکل ظاهری آزمایش مربوط به‌این‌گونه حوادث به‌دوصورت است‌که روش‌محاسبه‌احتمال حوادث در هردو شکل آزمایش یکسان می‌باشد . بدین ترتیب که برای انجام یک آزمایش n مرحله‌ای می‌توان با یک عنصر آزمایش را n بار تکرار نمود و یا اینکه در یک آزمایش n عنصررا باهم مورد آزمایش قرار داد که همان طور که قبلا" گفته شد ، در هر دو صورت نتیجه یکسان می‌باشد . به‌عنوان مثال برای محاسبه احتمال آمدن سه شیر در آزمایش پرتاب سکه ، می‌توانیم یک سکه را سه مرتبه متوالی پرتاب نموده و نتیجه را ثبت کنیم و این آزمایش را آنقدر تکرارنماییم تا اطمینان حاصل‌گردد که درنتیجه حاصله‌تصادف و شانس نقش چندانی ندارد . همین نتیجه را می‌توان از طریق آزمایش پرتاب همزمان سه سکه نیز بدست آورد . علت آنکه این حوادث چند مرحله‌ای نامیده شده‌اند آن است که احتمال آنها در چند مرحله متوالی محاسبه می‌گردد .

محاسبه احتمال این گونه حوادث بستگی به‌نوع حادثه داشته و به‌دوصورت کلی انجام‌پذیر است . اگر مراحل مختلف موجود در یک حادثه ساده از یکدیگر مستقل باشند محاسبه احتمال آن پیشامد ساده‌تر از وقتی است که این مراحل با یکدیگر وابستگی داشته و احتمال وقوع حوادث در هر مرحله از پیشامدهای مراحل قبلی اثر می‌پذیرند .

احتمال حوادث ساده چند مرحله‌ای با مراحلی مستقل از هم

برای محاسبه احتمال این گونه حوادث کافی است احتمال حادثه مطلوب درهر مرحله را محاسبه نموده و در یکدیگر ضرب کرد . اگر آزمایش شامل m مرحله مستقل بوده و در هر مرحله پیشامد ساده مطلوب باشد احتمال مربوطه از طریق رابطه زیر قابل محاسبه است .

(۱ – ۴)

$$P(A_1, A_2, \ldots\ldots A_i, \ldots A_m) = P_{A_1} \times P_{A_2} \times \cdots\cdots \times P_{A_i} \times \cdots \times P_{A_m}$$

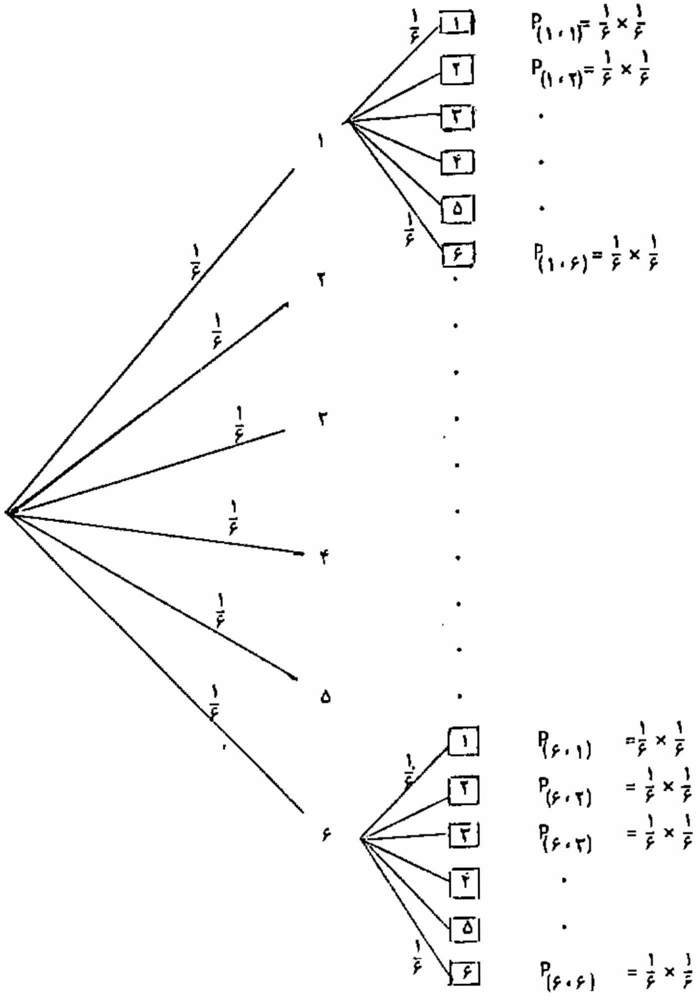
مثال ۶ – ۱ : مطلوب است محاسبه احتمال وقوع m شیر درآزمایش m مرتبه پرتاب یک سکه همتراز و یا درآزمایش پرتاب m سکه همتراز بطور همزمان

پاسخ : آمدن شیر در هریک از دفعات پرتاب و یا در مورد هر سکه از m سکه ، حادثه‌ای مستقل بوده و احتمال هر مرحله معادل $\frac{1}{2}$ است . با توجه به‌رابطه فوق داریم .

$$P_{(\text{شیر } m)} = \frac{1}{2} \times \frac{1}{2} \times \cdots\cdots\cdots \times \frac{1}{2} = \frac{1}{2^m}$$

مثال ۷ ـ ۱: مطلوبست محاسبه احتمال آنکه در آزمایش پرتاب همزمان دو تاس، تاس اول عدد ۶ و تاس دوم عدد ۳ را نشان دهد.

پاسخ: برای توضیح بیشتر در مورد شیوه محاسبه حوادث ساده چند مرحله‌ای، این مثال را ابتدا از طریق درخت احتمال نشان داده و سپس آن را حل می‌کنیم. در مرحله اول یعنی در پرتاب تاس اول وقوع ۶ حالت امکان‌پذیر است و لذا درخت ابتدا شش شاخه پیدا می‌کند حادثه مربوط به هریک از این شاخه‌ها که وقوع پیدا کند آن شاخه در مرحله بعد مجدد شش شاخه کوچک‌تر پیدا می‌کند زیرا در مرحله دوم نیز ۶ حالت ممکن است در اثر پرتاب تاس دوم اتفاق بیفتد و لذا درخت مزبور ۳۶ شاخه فرعی به‌شکل زیر پیدا می‌کند.



شکل ۲ـ۱: درخت احتمال مربوط به آزمایش پرتاب دو تاس

همانطور که در شکل مشاهده می‌شود احتمال مورد نظر $\frac{1}{36}$ می‌باشد . از طریق رابطه (۴-۱) نیز این مسئله قابل حل است . می‌دانیم که آمدن عدد ۶ در پرتاب تاس اول و آمدن عدد ۳ در پرتاب تاس دوم دو حادثه مستقل از یکدیگر بوده و احتمال هرکدام $\frac{1}{6}$ می‌باشد لذا احتمال حادثه مطلوب بقرار زیر است :

$$P_{(6 \text{ و } 3)} = P_6 \times P_3 = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

مثال ۸ - ۱ : مأمور نمونه‌گیری موءسسه استاندارد می‌خواهد از بین ۵ نوع مهره تولیدی یک کارگاه که در ۵ قوطی جداگانه قرار دارند ۵ مهره را به‌عنوان نمونه انتخاب کند (از هر قوطی یک مهره). مطلوب است احتمال آنکه به‌ترتیب مهره‌های دارای شماره‌های ۱، ۲، ۳، ۴، ۵ از قوطی اول تا پنجم انتخاب گردند . در صورتی که قوطی اول ده مهره، قوطی دوم ۹ مهره، قوطی سوم ۸ مهره، قوطی چهارم ۷ مهره و بالاخره قوطی پنجم ۶ مهره داشته باشد . و مهره‌ها در درون قوطی شماره‌گذاری شده باشند .

پاسخ : هریک از حوادث را در یک زوج مرتب نشان می‌دهیم . بدین ترتیب که مثلاً" جمله (۲-۲) به‌معنای انتخاب مهره شماره ۲ از قوطی دوم می‌باشد . برای حل این مسئله از رابطه (۴-۱) بصورت زیر استفاده می‌کنیم .

$$P\left\{(1-1),(1-2),(1-3),(1-4),(1-5)\right\} = P_{(1-1)} \times P_{(2-2)} \times P_{(3-3)} \times P_{(4-4)} \times P_{(5-5)}$$

برای محاسبه احتمال مزبور کافی است احتمال موجود در سمت راست معادله فوق را جدا جدا حساب کرده و در هم ضرب کنیم : داریم .

$$P_{(1-1)} = \frac{1}{10} \text{ و } P_{(1-2)} = \frac{1}{9} \text{ و } P_{(1-3)} = \frac{1}{8} \text{ و } P_{(1-4)} = \frac{1}{7} \text{ و } P_{(1-5)} = \frac{1}{6}$$

در نتیجه احتمال مورد نظر مقدار زیر را خواهد داشت .

$$P\left\{(1-1),(2-2),(3-3),(4-4),(5-5)\right\} = \frac{1}{10} \times \frac{1}{9} \times \frac{1}{8} \times \frac{1}{7} \times \frac{1}{6} = \frac{1}{30240}$$

احتمال حوادث ساده چند مرحله‌ای با مراحل غیر مستقل

احتمال این گونه حوادث نیز همانند مورد قبل معادل حاصلضرب احتمالات مربوط به‌هر مرحله می‌باشد . تنها مسئله‌ای که این گروه از حوادث را از پیشامدهای چند مرحله‌ای دارای مراحل مستقل متمایز می‌سازد آن است که به‌جز احتمال مربوط به‌مرحله اول ، احتمالات

مربوط به‌هریک ازسایر مراحل احتمالاتی‌مشروط و شرطی می‌باشند .بدین‌ترتیب اگر یک‌حادثه ساده شامل m مرحله غیر مستقل از A_1 تا A_m بوده باشد احتمال آن به‌صورت زیر محاسبه می‌گردد .

(۱ـ۵)

$$P_A=P(A_1,A_2,\ldots A_n)=P_{A_1}\times P(A_2|A_1)\times P(A_3|A_1,A_2)\times\cdots\times P(A_m|A_1,\ldots,A_{m-1})$$

مثال ۹ ــ ۱ . از کیسه‌ای که چهار مهره مشابه با چهار رنگ متفاوت قرمز، سبز، آبی و سیاه دارد می‌خواهیم سه مهره را بدون بازگردانی (یعنی مهره‌های انتخاب شده را به کیسه بر نمی‌گردانیم)، انتخاب کنیم .احتمال آنکه مهره‌های انتخابی به‌ترتیب قرمز، سبز و آبی رنگ باشند چقدر است؟

پاسخ: حادثه مطلوب یک حادثه سه مرحله‌ای بوده‌و آن را A می‌نامیم .چون سه‌مرحله از یکدیگر مستقل نیستند می‌توان از رابطه (۵ـ۱) استفاده کرد .

$$A=(\text{آبی وسبزو قرمز})\,،\ A_1=\text{قرمز}\,،\ A_2=\text{سبز}\,،\ A_3=\text{آبی}$$

از رابطه ۵ ـ ۱ داریم :

$$P_A=P_{A_1}\times P(A_2|A_1)\ \times P(A_3|A_1,A_2)$$

احتمال مربوط به‌هریک از مراحل را جداگانه حساب می‌کنیم . همانطورکه قبلا" گفته شد معمولا" این عمل بصورت ذهنی انجام می‌گیرد . در مرحله اول از بین ۴ مهره مشابه یکی انتخاب می‌شود . بنابراین احتمال آنکه این مهره قرمز باشد $P_{A_1}=\frac{1}{4}$ می‌باشد . در مرحله دوم فقط سه مهره برای انتخاب باقیمانده است و لذا $P_{(A_2|A_1)}=\frac{1}{3}$ است و بالاخره در مرحله سوم از بین دو مهره باقیمانده یکی انتخاب خواهد شد و لذا $P(A_3|A_1,A_2)=\frac{1}{2}$ خواهد بود . که حاصلضرب این سه احتمال، احتمال مطلوب یعنی P_A را تشکیل می‌دهد .

$$P_A=\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}=\frac{1}{24}$$

مثال ۱۰ ــ ۱: در یک مراسم قرعه‌کشی ۵۰ نفر و از جمله آقایان محمدی و علوی شرکت دارند نحوه قرعه کشی بدین ترتیب است که فردی با چشمان بسته از بین ۵۰ برگ کاغذ که برروی هرکدام اسامی یکنفر از ۵۰ نفر مزبور نوشته شده است ۵ برگ‌راانتخاب می‌کند ، احتمال آنکه

در انتخابهای اول و دوم بهترتیب آقایان محمدی و علوی انتخاب شوند چقدر است ؟
پاســـخ :

$$A_۲ = \text{علوی} \quad ، \quad A_۱ = \text{محمدی} \quad ، \quad A = (\text{علوی ومحمدی})$$

$$P_A = P_{A_۱} \times P_{(A_۲|A_۱)}$$

اما می‌دانیم که $P_{A_۱} = \frac{۱}{۵۰}$ و $P_{(A_۲|A_۱)} = \frac{۱}{۴۹}$ می‌باشد ، لذا احتمال حادثه A مقدارزیرا خواهد داشت :

$$P_A = \frac{۱}{۵۰} \times \frac{۱}{۴۹} = \frac{۱}{۲۴۵۰}$$

۲ ــ محاسبه احتمال حوادث مرکب

پیچیدگی اصلی د رمحاسبه احتمالات پیشامدهای مختلف عمدتا" مربوط به احتمال وقوع حوادث مرکب است . اگرآن‌طورکه بر مبنای تعریف احتمال در رابطه (۱ــ۱) گفته شد طرز محاسبه احتمال هر حادثه را $p = \frac{m}{n}$ بدانیم در بسیاری از حوادث مرکب محاسبه صورت و مخرج این کسر یعنی (m و n) بقدری پیچیده و مشکل است که عملا" استفاده از رابطه فوق را غیر ممکن می‌سازد . در حالی که در حوادث ساده صورت کسر همیشه $m = ۱$ بوده و مقدار مخرج n نیز معمولا" براحتی قابل تشخیص است اگرچه در این قبیل وقایع نیز محاسبات از نظرسهولت یکسان نبوده و بعضی‌از آنها از برخی‌دیگر پیچیده‌ترمی‌باشند .

به‌هرحال رابطه (۱ــ۱) اگرچه رابطه‌ای عام و فراگیر بوده و احتمال هرگونه حادثه‌ای با هردرجه از پیچیدگی ازنظر منطقی با استفاده از این رابطه قابل محاسبه‌است‌اما همان‌طور که قبلا" گفته شد استفاده ازاین‌رابطه به‌لحاظ مشکلات فراوانی که دارد در بسیاری از موارد غیر ممکن است و لذا برای سهولت حل مسائل مربوط به‌پیشامدهای مرکب ، این پیشامدها به‌چند دسته تقسیم ، و برای هرکدام راه حل ساده و مناسبی پیش بینی شده است. براساس این تقسیم‌بندی حوادث مرکب به‌چند دسته کلی به‌صورت زیر تقسیم می‌گردند .

۲ ــ الف ــ اجتماع چند حادثه ساده

۲ ــ ب ــ اجتماع چند حادثه مرکب

۲ ــ ج ــ اشتراک چند حادثه مرکب

هریک ازاین سه نوع خود به‌انواع مشخص‌تری تقسیم می‌شوند که ذیلا" به تفصیل مورد بحث قرار می‌گیرند .

۲ ـ الف ـ اجتماع چند حادثه ساده

این گروه از حوادث مرکب ، حوادثی هستند که فقط و فقط به‌چند حادثه ساده قابل تجزیه می‌باشند به‌عنوان مثال انتخاب یک مهره قرمز از کیسه‌ای که ۲ مهره قرمز و سه مهره مشکی رنگ دارد از این نوع حوادث بوده و فقط قابل تجزیه به‌دو حادثه ساده است یکی از این پیشامدهای ساده انتخاب مهره قرمز رنگ اول و حادثه دیگر انتخاب مهره قرمزرنگ‌دوم می‌باشد . برای محاسبه احتمال این گونه حوادث از قضیه زیر استفاده می‌شود .

قضیه ۳ ـ ۱ : احتمال یک حادثه مرکب که به m حادثه ساده قابل تجزیه است برابر مجموع احتمالات مربوط به هریک از آن حوادث ساده می‌باشد .

با استفاده از این قضیه می‌توان رابطه عام و کلی (۱ـ۱) را بدست آورد . اگر فضای نمونه مربوط به‌یک آزمایش n نقطه نمونه‌داشته باشد و حادثه موردنظر A نیز حادثه‌مرکبی باشد که به m حادثه ساده (A_1 و A_2 ..... A_m) تجزیه شوداحتمال این حادثه یعنی P_A به‌صورت زیر قابل محاسبه است .

$$P_A = P_{A_1}+P_{A_2}+\cdots\cdots+P_{A_m}$$

از طرفی می‌دانیم که حوادث A_1، A_2، ..... ، A_m وقایعی ساده بوده و احتمال هریک از آنها $\frac{1}{n}$ می‌باشد لذا خواهیم داشت .

$$P_A = \left(\frac{1}{n}\right)_1 + \left(\frac{1}{n}\right)_2 + \cdots\cdots + \left(\frac{1}{n}\right)_m = m\left(\frac{1}{n}\right) = \frac{m}{n}$$

بدین ترتیب رابطه کلی محاسبه احتمالات یعنی رابطه (۱ـ۱) اثبات می‌گردد . این رابطه چنان که قبلا" گفته شد می‌تواند برای محاسبه احتمال هرگونه حادثه‌ای مورد استفاده قرارگیرد اما به‌دلیل‌دشواریها و پیچیدگیهای‌موجود در بسیاری از موارد ، عملا" فقط درمورد محاسبه احتمال حوادث موضوع این بند به‌کار گرفته می‌شود .

مثال ۱۱ ـ ۱ . در کیسه‌ای دو مهره قرمز و سه مهره مشکی رنگ قرار دارد . مطلوب است احتمال آن که در یک انتخاب تصادفی ، از این کیسه یک مهره قرمز رنگ بیرون آید .

پاسخ : حادثه مطلوب همان‌طوری‌که قبلا" گفته شد فقط به‌دو حادثه ساده انتخاب مهره قرمز رنگ اول یا دوم قابل تجزیه بوده و فضای نمونه آن نیز دارای ۵ نقطه نمونه می‌باشد . بنابراین احتمال آن مقدار زیر را خواهد داشت .

$$P_A = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

چنان که قبلا" گفته شد تشخیص این‌که حادثه مطلوب (حادثه‌مربوط به‌صورت کسر) و یا فضای نمونه (مخرج کسر) به‌چند حادثه ساده قابل تجزیه‌اند ، همیشه امری ساده‌نیست. در بسیاری از موارد مقدار m و n باید با استفاده از روشهای خاص ریاضی از قبیل ترکیب [1] و ترتیب [2] به‌دست آیند که ذیلا" موارد مشخص و استاندارد این روشها بدون اثبات و بطور مختصر توضیح داده شده و برای هریک از آنها یک مثال آورده می‌شود .

کیسه‌ای‌را در نظر بگیریدکه حاوی ۳ مهره بوده و برروی هریک از این مهره‌ها به‌ترتیب اعداد ۱ ، ۲ ، ۳ نوشته شده است . می‌خواهیم بدانیم درآزمایش انتخاب دو مهره از این کیسه چند حالت می‌تواند اتفاق بیفتد ؟ این سئوال به‌دو صورت می‌تواند پاسخ گفته شود . ممکن است برای ما به‌عنوان مثال ، حادثه انتخاب مهره ۱ در انتخاب‌اول و مهره۲ در انتخاب دوم با حادثه انتخاب مهره ۲ در انتخاب اول و انتخاب مهره ۱ در انتخاب دوم بی‌تفاوت باشد یعنی ترتیب انتخاب شماره‌ها اهمیتی نداشته و فقط‌وجود اصل هریک از سه شماره مهم باشد . این حالت را اصطلاحا" ترکیب نامیده و در این صورت مسئله فوق سه حالت{۲و۱}و {۳و۱}و {۳و۲} را می‌تواند داشته باشد . اما اگر ترتیب انتخاب شماره‌ها نیز مورد نظر باشد پاسخ مسئله ۶ می‌باشد زیرا هریک ازحوادث سه‌گانه فوق به دو حالت تقسیم می‌شوند . این حالت را ترتیب یا تبدیل نامیده و ۶ زوج مرتب زیر پاسخ مسئله فوق در حالت تبدیل می‌باشد .

(۲ و۱) و (۱ و۲) و (۳ و۱) و (۱ و۳) و (۳ و۲) و (۲ و۳)

در حالت کلی مطلب فوق را می‌توان در قالب قضایای ذیل عنوان نمود .

قضیه ۴-۱ . تعداد ترتیبهای m شی از n شیء متمایز برابر است با

$$P_{n,m} = \frac{n!}{(n-m)!} \qquad (۱-۶)$$

مثال ۱۲ - ۱ : با ارقام (۱،۲،۳،۴،۵) چند عدد سه رقمی می‌توان نوشت بطوری که هر عدد فقط یکبار مورد استفاده قرار گیرد .

پاسخ : m=۳ و n=۵ می‌باشد لذا جواب مسئله مقدار زیر را خواهد داشت :

$$P_{۵,۳} = \frac{۵!}{(۵-۳)!} = ۶۰$$

مثال ۱۳ - ۱ : با ارقام مسئله قبل چند عدد ۵ رقمی می‌توان نوشت در حالی که ازهر عدد فقط یکبار استفاده شده باشد .

1 - Combination

2 - Permutation

پاسخ : $n = m = ۵$ می‌باشد لذا داریم :

$$P_{۵،۵} = \frac{۵!}{(۵-۵)!} = \frac{۵!}{۰!}$$

می‌دانیم که $۰! = ۱$ می‌باشد لذا جواب مسئله به‌صورت زیر در می‌آید :

$$P_{۵،۵} = ۵! = ۱۲۰$$

قضیه ۵ ــ ۱ : تعداد ترکیبهای m شیئ از n شیئ متمایز برابر است با :

$$\binom{n}{m} = C_n^m = \frac{n!}{m!\,(n-m)!} \qquad (۱-۷)$$

مثال ۱۴ ــ ۱ : به‌چند طریق می‌توان از بین ۵ نفر یک مجموعه سه‌تائی تشکیل داد .

پاسخ : چون درمجموعه‌ها ترتیب قرارگرفتن عناصر موردنظرنیست‌می‌توان ازرابطه (۷ــ۱) استفاده نمود .

$$\binom{۵}{۳} = C_۵^۳ = \frac{۵!}{۳!\,(۵-۳)!} = ۱۰$$

در قضیه ۵ــ۱ . اگر درست دقت شود در باره تعداد حالاتی که می‌توان n عنصر متمایز را در دو گروه m تایی و $n - m$ تایی تقسیم نمود صحبت می‌شود زیرا وقتی که یک مجموعه m عنصری از n عنصر جدا می‌شود عناصر باقیمانده یک مجموعه $n-m$ عنصری‌را تشکیل می‌دهند . می‌توان این قضیه را از شکل فوق که مربوط به‌دو گروه است به‌قضیه زیر که تقسیم n عنصر در K گروه است نیز تعمیم داد .

قضیه ۶ ــ ۱ : تعداد طرقی که می‌توان یک مجموعه n عنصری را به K گروه جداگانه تقسیم نمود به‌طوری که $n_۱$ عنصر گروه اول و $n_۲$ عنصر در گروه دوم و . . . . n_K عنصر در گروه K ام قرار گیرند برابر است با :

(۱ ــ۸)

$$\binom{n}{n_۱\, n_۲ \ldots\ldots n_k} = \frac{n!}{n_۱!\; n_۲!, \ldots n_k!}$$

مثال ۱۵ ــ ۱ : به‌چند طریق می‌توان ۷ نفر را در یک اطاق سه‌تختی و دو اطاق دو تختی‌جا داد .

پاسخ :

$$\binom{۷}{۳و۲و۲} = \frac{۷!}{۳!\,۲!\,۲!} = ۲۱۰$$

مواردی که تا کنون مورد بحث قرار گرفتند همه دارای یک وجه مشترک هستند و آن

این است که تعداد n عنصر اولیه همگی از هم متمایز بوده و به‌عبارت دیگر همگی‌غیر مشابه هستند اما اکنون قضیه‌ای را مطرح می‌کنیم که در آن n عنصر موجود در K گروه جا دارند که هر گروه حاوی تعدادی عناصر مشابه هم می‌باشد. بدین ترتیب از n عنصر موجود n_1 عنصر با یکدیگر مشابه و با سایر عناصر متفاوت هستند و به‌همین ترتیب n_2 عنصر باهم مشابه و سایرین متفاوت و ...... و بالاخره n_K عنصر باهم مشابه و با بقیه متفاوت می‌باشند.

قضیه ۷ ــ ۱: تعداد ترتیب‌ها و یا تبدیلهای n تائی متمایز n شیئ که n_1 تای آن یک نوع و n_2 تای آن از نوع دوم و ...... و بالاخره n_K تای آن از نوع K ام است از رابطه (۸ــ۱) قابل محاسبه می‌باشد.

$$\frac{n!}{n_1!\,n_2!\,\ldots\ldots\,n_k!}$$

مثال ۱۶ ــ ۱: به‌چند طریق متفاوت می‌توان دو درخت کاج و یک درخت بلوط را کنار یکدیگر کاشت.

پاسخ: روشن است که دو مجموعه ذیل با یکدیگر تفاوتی ندارند و بنابراین هردو یک مجموعه تلقی می‌شوند.

{کاج اول، بلوط، کاج دوم} = {کاج دوم، بلوط، کاج اول}

با توجه به‌مطلب فوق کل حالات ممکنه را می‌توان به‌صورت سه مجموعه زیر نشان داد.

{کاج و کاج و بلوط} و {کاج و بلوط و کاج} و {بلوط و کاج و کاج}

همین نتیجه را می‌توان از رابطه (۸ــ۱) نیز بدست آورد.

$$\frac{۳!}{۲!\times ۱!} = \frac{۶}{۲} = ۳$$

مثال ۱۷ ــ ۱: به‌چند طریق متمایز می‌توان ۳ لامپ قرمز، ۴ لامپ زرد و ۲ لامپ آبی را در یک ردیف نصب کرد.

پاسخ: از رابطه (۸ــ۱) استفاده می‌کنیم.

$$\frac{۹!}{۳!\times ۴!\times ۲!} = ۱۲۶۰$$

در مثالهای فوق صرفاً در مورد تعداد ترکیب‌ها و ترتیبها در حوادث مختلف صحبت می‌شد و از احتمال حرفی در میان نبود اما در دو مثال ذیل نقش قضایای فوق در حل مسائل

مربوط به احتمالات مربوط به حوادث مرکب که از مجموع چند حادثه ساده ایجاد شده است به‌خوبی روشن می‌شود .

مثال ۱۸ ــ ۱ : از یک کیسه حاوی ۴ مهره سفید و ۲ مهره سیاه ، سه مهره را می‌خواهیم انتخاب کنیم . احتمال آنکه ۲ مهره سفید و یک مهره سیاه انتخاب شوند چقدر است ؟

پاسخ : حادثه مورد نظر یک حادثه مرکب است که از مجموع چند حادثه سه‌مرحله‌ای پدید آمده است زیرا این حادثه می‌تواند چنان که بعداً مشخص می‌شود به ۱۲ حادثه ساده سه مرحله‌ای تقسیم شده و احتمال آن از جمع احتمالات این ۱۲ حادثه بدست آید که جهت سهولت از قضایای ترکیب برای حل آن استفاده می‌شود .

$$P_A = \frac{m}{n}$$

در این رابطه n مساوی است با تعداد ترکیب‌های ۳ تایی ممکن‌الحصول که ۲ مهره سفید و یک مهره سیاه داشته باشند . با کمی دقت می‌توان دریافت که فقط دو حالت برای انتخاب مهره سیاه وجود دارد و با هریک از این دو حالت $\binom{۴}{۲}$ حالت برای انتخاب ۲ مهره سفید از ۴ مهره سفید می‌تواند متناظر شود . بنابراین صورت کسر مقدار زیر خواهد بود :

$$m = \binom{۲}{۱}\binom{۴}{۲} = ۲ \times ۶ = ۱۲$$

n یعنی مخرج کسر نیز با کل ترکیبات سه‌تائی ممکن‌الحصول بدون هیچگونه قید و شرطی برابر است (فضای نمونه) و بنابراین مقدار آن از رابطه زیر بدست می‌آید .

$$n = \binom{۶}{۳} = \frac{۶!}{۳!\ (۶ - ۳)!} = ۲۰$$

از تقسیم دو مقدار فوق برهم مقدار P_A بدست می‌آید .

$$P_A = \frac{۱۲}{۲۰} = ۰/۶ = ۶۰\ \%$$

مثال ۱۹ ــ ۱ : مؤسسه‌ای ۵۰ نفر پرسنل دارد که ۱۰ تن از آنان مدیر و ۴۰ نفر باقیمانده کارمند جزء می‌باشند . مسئولین مؤسسه می‌خواهند سه جایزه را برحسب قرعه به سه نفر از پرسنل مؤسسه اعطا نمایند اما مایلند که حداکثر یکی از جوایز به‌گروه مدیران تعلق گیرد . احتمال تأمین این خواسته را در شرایط طبیعی و بدون استفاده از روشهایی از قبیل سهمیه بندی و غیره تعیین کنید .

پاسخ : احتمال مطلوب حاصل جمع دو احتمال است یکی اینکه هیچیک از جوایز به‌مدیران

تعلق نگیرد $P_{(\cdot)}$ و دیگری اینکه یکی از جوایز به‌آنها تعلق بگیرد $P_{(1)}$ که هریک‌از اینها خود احتمالات مربوط به‌حوادث مرکبی هستند که از مجموع چندین حادثه ساده سه‌مرحله‌ای‌پدید آمده‌اند و شایسته است خواننده محترم جهت تمرین تعدادی از این حوادث ساده رابرروی کاغذ مشخص نماید . بدین ترتیب حادثه مورد نظر را به‌صورت زیر می‌توان نوشت .

$$P_A = P_{(\cdot)} + P_{(1)}$$

عبارت $P_{(\cdot)}$ به‌معنای آن است که هر سه نفر از بین ۴۰ نفر کارمند جزء انتخاب شوند یعنی هرسه فرد انتخابی از گروه ۴۰ نفری انتخاب شوند . تعداد کل ترکیبات ممکن‌الوقوع‌مطلوب از رابطه ترکیب $\binom{۴۰}{۳}$ مشخص می‌گردند .

$$P_{(\cdot)} = \frac{m}{n} = \frac{\binom{۴۰}{۳}}{\binom{۵۰}{۳}}$$

به‌همین ترتیب $P_{(1)}$ به‌صورت زیر قابل محاسبه است .

$$P_{(1)} = \frac{m}{n} = \frac{\binom{۱۰}{۱}\binom{۴۰}{۲}}{\binom{۵۰}{۳}}$$

حاصل جمع این دو احتمال ، احتمال مورد نظر را معین می‌کند .

$$P_A = P_{(\cdot)} + P_{(1)} = \frac{\binom{۴۰}{۳}}{\binom{۵۰}{۳}} + \frac{\binom{۱۰}{۱}\binom{۴۰}{۲}}{\binom{۵۰}{۳}} = \frac{\binom{۴۰}{۳} + \binom{۱۰}{۱}\binom{۴۰}{۲}}{\binom{۵۰}{۳}}$$

$$P_A = \frac{\left(\frac{۴۰\times۳۹\times۳۸}{۳!}\right) + \left(\frac{۱۰}{۱}\right)\left(\frac{۴۰\times۳۹}{۲}\right)}{\frac{۵۰\times۴۹\times۴۸}{۳!}} = \frac{۲۲۱}{۲۴۵} = ۰/۹ = ۹۰\ \%$$

حال پس ازمشخص شدن شیوه حل مسائل‌مربوط به‌احتمال حوادث مرکبی که از اجتماع چند حادثه ساده پدید آمده‌اند نوبت به‌بحث پیرامون حوادث مرکبی می‌رسد که از اجتماع چند حادثه مرکب تشکیل یافته‌اند و مثال ۱۹ــ۱ یکی از اشکال ساده این قبیل حوادث بود

۲ ـ ب ـ اجتماع چند حادثه مرکب

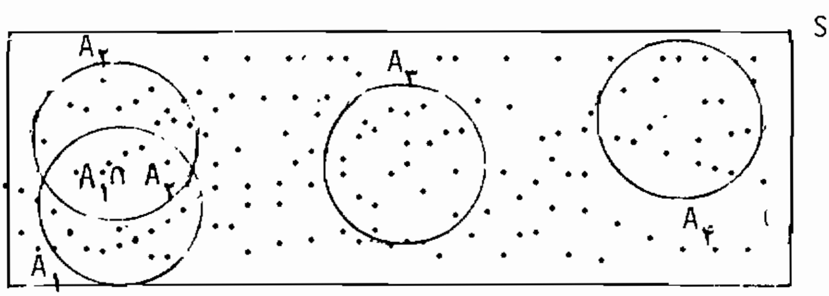
این روش مرسوم ومتداول است‌که پیشامدهای مرکب‌موضوع این‌بند را به‌زبان مجموعه‌ها بیـان می‌کنندو این کـار از آن جهت که به‌فهم بهتر مطلب کمک می‌کند روشی پسندیده و نیکوست‌و به‌همین دلیل در اینجا نیز مراعات گردیده است .

محاسبه احتمال حوادث موضوع بند ب وقتی میسر است که حوادث سازگار و ناسازگار از یکدیگر تمیز داده شوند . به‌همین جهت ابتدا راجع به‌این دونوع حادثه صحبت نمـوده و سپس بحث اصلی را دنبال می‌کنیم .

حوادث سازگار و ناسازگار

در شکل ۳ـ۱ مجموعه S را که شامل n عنصر می‌باشد در نظر بگیرید . به‌همین‌ترتیب چهار مجموعه A_1 ، A_2 ، A_3 و A_4 را که هریک به‌ترتیب دارای m_1 ، m_2 ، m_3 و m_4 عنصربوده و هریک زیر مجموعه‌ای از مجموعه S می‌باشند ملاحظه کنید . اگر هریک از ایـن چهـار زیر مجموعه ، نشان دهنده یک حادثه خاص باشند و پیشامد A_1 و A_2 را نسبت به‌یکدیگر سازگار و یا قابل جمع نامیده و دو پیشامد A_3 و A_4 را ناسازگار و یا مانعةالجمع نسبت بیکدیگـر بنامیـم ، واضح است که حوادث A_3 و A_4 نسبت‌به‌حوادث A_1 و A_2 نیز ناسازگار می‌باشند .

" دو پیشامد را ناسازگار گویند وقتی که هیچ یک از حوادث ساده مربوط به‌آنها "
" با یکدیگر مشترک نباشند . بعبارت دیگر $A \cap B = \phi$ "
" و بعکس دو پیشامد را سازگار گویند وقتیکه حداقل یکی از حوادث ساده مربوط "
" بـه آنها در هردو مشترک باشد و بعبارت دیگر $A \cap B \neq \phi$ "



شکل ۳ـ۱ حوادث سازگار و ناسازگار

درآزمایش پرتاب یک تاس آمدن عدد زوج و آمدن عدد فرد دوحادثه ناسازگار و مانعه‌الجمع هستند در حالی که آمدن عدد زوج و آمدن عدد کوچکتر از ۵ دو حادثه سازگار می‌باشند .

واضح است که فقط حوادث مرکب می‌توانند با یکدیگر سازگار باشند و کلیـه حوادث ساده با یکدیگر ناسازگارند زیرا حادثه ساده دارای اجزائی نیست که بتواند با دیگر حوادث مشترک باشد .

حال‌که مفهوم سازگاری و ناسازگاری حوادث توضیح داده شد می‌توان درمورد احتمال حادثه مرکبی که از اجتماع چند حادثه مرکب دیگر پدید آمده است صحبت نمود . به زبان مجموعه‌ها اگرحادثه مثل A از مجموع دو حادثه دیگر مثل $A_۱$ و $A_۲$ به‌وجودآمده‌باشدمی‌توان آنرا به‌صورت زیر نشان داد .

$$A=A_۱ \cup A_۲$$

حال با توجه به‌شکل ۳ـ۱ اگر بخواهیم تعداد نقاط موجود در مجموعه A را بشمریم می‌توانیم بصورت زیر عمل کنیم .

$$A=A_۱+A_۲-(A_۱ \cap A_۲)$$

علت‌اینکه‌عبارت $(A_۱ \cap A_۲)$ را کم‌می‌کنیم این‌است که تعدادنقاط موجود دراین‌مجموعه دومرتبه شمارش شده‌اند .اگر تعداد نقاط موجود در مجموعه $(A_۱ \cap A_۲)$ را با $m_۵$ وتعداد نقاط موجود در مجموعه A را با m نشان دهیم با توجه به‌اینکه قبلا" گفتیم $m_۱$ و $m_۲$ نشان ـ دهنده تعداد نقاط موجود در مجموعه‌های $A_۱$ و $A_۲$ می‌باشند واضح است که رابطه زیـــر را می‌توان در مورد تعداد نقاط موجود در A نوشت .

$$m = m_۱ + m_۲ - m_۵$$

اگر فضای نمونه مربوط به‌حوادث A و $A_۱$ و $A_۲$ مجموعه S باشد که تعداد نقاط نمونه آن‌برابر n است واضح است که احتمال مربوط به هریک از حوادث فوق به‌صورت عبارت زیـر قابـل نمایش است .

$$P_A= \frac{m}{n} \quad , \quad P_{A_۱}= \frac{m_۱}{n} \quad , \quad P_{A_۲}= \frac{m_۲}{n} \quad , \quad P_{(A_۱ \cap A_۲)}= \frac{m_۵}{n}$$

حال طرفین رابطه قبلی را بر n تقسیم می‌کنیم .

$$\frac{m}{n}= \frac{m_۱}{n} + \frac{m_۲}{n} - \frac{m_۵}{n}$$

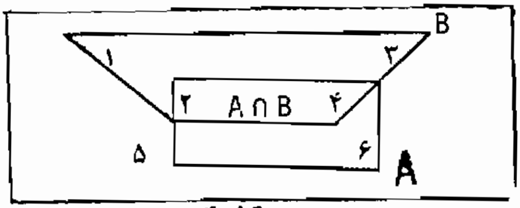
با توجه به‌احتمالات مربوط به هریک از حوادث ، معادله فوق را بصورت زیر نیز می‌تـــوان نوشت .

(۱ــ۹)

$$P_{(A_1 \cup A_2)} = P_A = P_{A_1} + P_{A_2} - P_{(A_1 \cap A_2)}$$

مثال ۲۰ــ۱ : در آزمایش پرتاب یک تاس A نشان دهنده پیشامد وقوع عدد زوج و B نشان دهنده حادثه‌وقوع عددکوچکتر یا مساوی ۵ می‌باشد .مطلوبست احتمال آنکه در این‌آزمایش عدد زوج و یا کوچکتر از ۵ بیاید .

پاسـخ : برای وضوح بیشتر ابتدا دو حادثه A و B را بر روی نمودار مشخص می‌کنیم .



شکل ۴-۱

همان طور که در شکل مشخص است دو حادثه A و B با یکدیگر سازگار بوده و بنابراین برای حل مسئله می‌توان از رابطه ۹ــ۱ استفاده کرد .

$$P_{(A \cup B)} = P_A + P_B - P_{(A \cap B)}$$

با توجه به‌اینکه فضای نمونه ۶=N نقطه نمونه دارد احتمالات طرف راست رابطه فوق را از روی شکل می‌توان بدست آورد .

$$P_A = \frac{۳}{۶} = \frac{۱}{۲} \quad و \quad P_B = \frac{۴}{۶} = \frac{۲}{۳} \quad و \quad P_{(A \cap B)} = \frac{۲}{۶} = \frac{۱}{۳}$$

و در نتیجه جواب مسئله مقدار زیر خواهد بود .

$$P_{(A \cup B)} = \frac{۳}{۶} + \frac{۴}{۶} - \frac{۲}{۶} = \frac{۵}{۶}$$

رابطه ۹ــ۱ رابطه‌ای کلی است و می‌تواند برای کلیه حوادثی که از دو حادثه مرکب تشکیل شده‌اند مورد استفاده قرار گیرد اعم از آنکه این حوادث مرکب با یکدیگر سازگار و یا ناسازگار باشند . واضح است که وقتی دو حادثه A_1 و A_2 ناسازگار باشند با یکدیگر وجه مشترکی نداشته و طبعاً عبارت $A_1 \cap A_2 = \phi$ شده و در نتیجه احتمال آن مساوی صفر خواهد شد :

$$P_{(A_1 \cap A_2)} = 0$$

و بدین ترتیب درحوادث ناسازگار جملات مربوط به‌اشتراکات مجموعه‌ها مساوی صفر شده و از رابطه ۹ــ۱ حذف می‌شوند و در نتیجه رابطه ۹ــ۱ برای حوادث ناسازگار به‌صورت رابطه ۱۰ــ۱ در می‌آید . اگر A_1 و A_2 دو حادثه ناسازگار باشند داریم

(۱ ــ ۱۰)

$$P_{(A_1 \cup A_2)} = P_{A_1} + P_{A_2}$$

اگر حادثه‌ای از جمع m پیشامدمرکب ناسازگار ایجاد شده‌باشد چون هیچیک از حوادث‌باهم وجه مشترکی ندارند لذا جملات مشترک وجود نداشته و رابطه ۱۰ــ۱ را می‌توان به‌صورت رابطه ۱۱ــ۱ برای m حادثه ناسازگار نیز نوشت .

(۱ ــ ۱۱)

$$P_{(A_1 \cup A_2 \cup \cdots\cdots \cup A_m} = P_{A_1} + P_{A_2} + \cdots\cdots + P_{A_m}$$

مثال ۲۱ ــ ۱ : کیسه‌ای حاوی ۵ مهره قرمز ، ۶ مهره سفید ، ۲ مهره زرد ، ۳ مهره سبز و ۴ مهره آبی رنگ می‌باشد .احتمال آنکه در یک انتخاب‌مهره قرمز یا سفید یازرد و یا سبزبیرون بیاید چقدر است ؟

پاسخ : اگر حادثه انتخاب مهره قرمز را با A ، مهره سفید را با B ، زرد را با C و سبز را با D نشان دهیم واضح است که این حوادث با یکدیگر ناسازگار بوده و لذا از رابطه ۱۱ ــ ۱ می‌توان استفاده نمود .

$$P_{(A \cup B \cup C \cup D)} = P_A + P_B + P_C + P_D$$

جملات سمت راست معادله را جداگانه حساب نموده و سپس در معادله قرار می دهیم .

$$P_A = \frac{5}{20} \text{ و } P_B = \frac{6}{20} \text{ و } P_C = \frac{2}{20} \text{ و } P_D = \frac{3}{20}$$

بنابراین

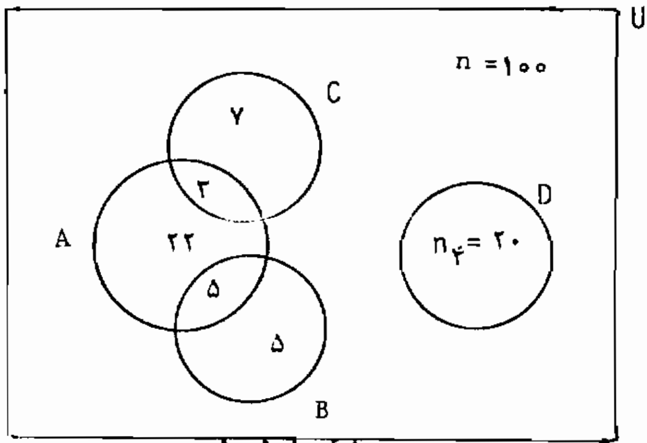
$$P_{(A \cup B \cup C \cup D)} = \frac{۵}{۲۰} + \frac{۶}{۲۰} + \frac{۲}{۲۰} + \frac{۳}{۲۰} = \frac{۱۶}{۲۰} = ۰/۸ = ۸۰\ \%$$

برای محاسبه احتمال حادثه مرکبی که از اجتماع n حادثه مرکب پدید آمده‌است واین حوادث همگی و یا بعضی از آنها بایکدیگر سازگار هستند رابطه کلی و فراگیری وجود ندارد. زیرا نحوه سازگاری حوادث در تمامی موارد یکسان نیستند ممکن است در یک مورد ، حادثه A با B سازگار و هردوی این پیشامدها با واقعه C ناسازگار باشند و یا اینکه A با B و C سازگار اما B و C با یکدیگر ناسازگار باشند و یا بعکس B با A و C هردو سازگار و A و C با یکدیگر ناسازگار باشند و یا هرسه حادثه باهم سازگار و . . . . نتیجه‌چنین‌تفاوتهائی‌آن است که در روابطی که برای محاسبه هریک از این حوادث در نظر گرفته می‌شوند جملات مربوط به‌اشتراک حوادث‌که برای حذف دوباره‌کاری‌ها در نظر گرفته می‌شوند با یکدیگرتفاوت داشته و نمی‌توان یک رابطه واحد و کلی را برای تمامی موارد ارائه نمود .

البته معنی جملات فوق این نیست که برای تشخیص و بدست آوردن این‌روابط هیچ ضابطه کلی وجود ندارد . برای تشخیص رابطه کلی مربوط به‌محاسبه یک حادثه مرکب از نوع موردبحث می‌توان براساس قواعد مربوط به‌مجموعه‌ها ،از طریق جمع کردن نقاط موجود درچند مجموعه سازگاری که حادثه مورد نظر را تشکیل می‌دهند و کسر کردن نقاطی‌که‌دوبار و یا بیشتر محاسبه شده‌اند به‌این مهم دست یافت .

مثال ۲۲ــ۱ : بر طبق ضوابط موجود در یک اداره که یکصدنفر پرسنل دارد در هرسال ۳۰ نفر از پرسنل می‌توانند از درآمدهای حاصل از اضافه‌کار A ، ۱۰ نفر از درآمدهای مربوط به‌مأموریتهای داخلی خارج از مرکز B ۱۰ نفر از درآمدهای مربوط به‌مأموریتهای خارج از کشور C و ۲۰ نفر نیز از مزایای پاداشهای غیرنقدی D استفاده نمایند . بر طبق‌همین مقررات هیچیک از افراد گروه D نمی‌توانند عضو سه گروه دیگر باشند . همچنین هیچیک از افراد دوگروه B و C نیز نمی‌توانند درگروه دیگر عضویت داشته باشند بالاخره معمولاً" سه نفر از افراد گروه A در گروه C و ۵ نفر نیز در گروه B عضویت دارند . مطلوبست محاسبه احتمال آنکه یکنفر از پرسنل این‌اداره بتواند در سال حداقل از یکی از مزایای درآمدی‌فوق استفاده نماید .

پاسخ : مطابق نمودار زیر که به‌دیاگرام ون معروف است حوادث (A و B) و (C و A) با یکدیگر سازگاربوده و حوادث (B با C)و نیز حوادث (A و B و C با D) با یکدیگرناسازگار هستند و حادثه موردنظر پیشامد ($A \cup B \cup C \cup D$) می‌باشد .



شکل ۵-۱

مطابق شکل می‌توان برای محاسبه احتمال مورد نظر مجموع نقاط موجود در چهار مجموعه را از طریق رابطه زیر محاسبه نمود .

$$(A \cup B \cup C \cup D) = A + B + C + D - (A \cap B) - (A \cap C)$$

حال برای محاسبه احتمال حادثه مطلوب رابطه فوق را می‌توان به‌شکل زیر تغییر داد .

$$P_{(A \cup B \cup C \cup D)} = P_A + P_B + P_C + P_D - P_{(A \cap B)} - P_{(A \cap C)}$$

عبارات طرف راست رابطه فوق با توجه به‌شکل به راحتی قابل محاسبه‌اند .

$$P_A = \frac{۳۰}{۱۰۰} = ۰/۳ \quad \text{و} \quad P_C = \frac{۱۰}{۱۰۰} = ۰/۱ \quad \text{و} \quad P_D = \frac{۲۰}{۱۰۰} = ۰/۲ \quad \text{و} \quad P_{(A \cap B)} = ۰/۰۵$$

$$P_B = \frac{۱۰}{۱۰۰} = ۰/۱ \quad \text{و} \quad P_{(A \cap C)} = ۰/۰۳$$

از جایگذاری این ارقام در رابطه اصلی احتمال موردنظر محاسبه می‌گردد .

$$P_{(A \cup B \cup C \cup D)} = ۰/۱ + ۰/۳ + ۰/۱ - ۰/۲ - ۰/۰۵ - ۰/۰۳ = ۰/۶۲ = ۶۲\%$$

نکته‌ای که به‌عنوان پیش درآمد بحث بندج قابل ذکر است این‌است که احتمال حوادثی از قبیل $A \cap B$ و نظایر آن در همه موارد به‌راحتی قابل محاسبه نیست و باید برای محاسبه

آن از شیوه‌های بخصوصی استفاده نمود . که این شیوه‌ها در بند ج که موضوع آن اشتراک چند حادثه مرکب می‌باشد توضیح داده خواهد شد .

۲ ـ ج ـ اشتراک چند حادثه مرکب

حادثه‌ای که از اشتراک چند حادثه مرکب پدید می‌آید در واقع شکل پیچیده‌تری از حوادث ساده چند مرحله‌ای می‌باشد . از آنجا که این شکل ساده یعنی حوادث ساده چند مرحله‌ای قبلاً" به‌تفصیل مورد بحث قرار گرفته‌اند در اینجا از بحث تفصیلی پیرامون شکل پیچیده‌تر آنها یعنی اشتراک چند حادثه مرکب خودداری کرده و به توضیح مختصری پیرامون این قبیل پیشامدها اکتفا می‌شود .

تفاوت اصلی یک پیشامد مرکب ایجاد شده از اشتراک چند حادثه مرکب با یک واقعه چند مرحله‌ای ساده آن‌است که در پیشامد مرکب موردگفتگو ، هریک از مراحل‌خود حادثه‌ای مرکب می‌باشند در حالی که مراحل مربوط به یک حادثه ساده تماما" حوادثی ساده هستند و به‌همین دلیل محاسبه احتمال هریک از مراحل یک پیشامد مرکب معمولا" پیچیده‌تر از یک حادثه ساده مشابه می‌باشد .

تفاوت فوق‌الذکر هیچ تفاوتی را در راه حل این دو گروه از حوادث ایجاد نکرده و احتمال هر دونوع این وقایع با شیوه‌های همانندی که در بحث حوادث ساده چند مرحله‌ای مشروحا" مورد بحث قرار گرفتند ، محاسبه می‌گردد . در اینجا جهت یادآوری ، این روشها در قالب عناوین مناسب با بحث فعلی مجددا" بیان می‌گردند .

همانند یک‌حادثه ساده چند مرحله‌ای روش محاسبه احتمال اشتراک چند حادثه‌مرکب نیز بسته به‌استقلال و یا عدم استقلال هریک از آن حوادث مرکب به‌دونوع تقسیم می‌گردد. که در اینجا به‌بحث پیرامون آن دو می‌پردازیم .

۲ـج ـ۱ـ اشتراک چند حادثه مرکب مستقل

مفهوم واقعی حادثه مرکبی که از اشتراک چند حادثه مرکب دیگر پدید آمده است و به‌زبان مجموعه‌ها به‌صورت ($A \cap B \cap C \ldots\ldots$) نشان داده می‌شود چیزی جزیک‌حادثه مرکب چند مرحله‌ای نیست . بدون شک چنین حادثه‌ای فقط می‌تواند از ترکیب چند حادثه مرکب سازگار با یکدیگر ایجاد گردد چون حوادث ناسازگار هیچ وجه مشترکی ندارند تا از اشتراک آنها یک حادثه جدید پدیدار گردد .

آمدن دو عدد زوج در پرتاب دو تاس ، انتخاب چند مهره خاص از دو کیسه که‌هرکدام حاوی تعدادی مهره هستند و نظایر اینها نمونه‌های ساده‌ای از اشتراک چند حادثه مرکب

مستقل هستند که معمولا" در آمار به‌عنوان مثال مورد استفاده قرار می‌گیرند . رابطه اصلی برای محاسبه احتمال این قبیل حوادث ، همان رابطه (۴-۱) می‌باشد .

(۱ - ۴)

$$P_{(A\cap B\cap C\cdots\cdots)}=P_A\times P_B\times P_C\times\cdots\cdots$$

مثال ۲۳ - ۱ : مطلوبست محاسبه احتمال آمدن سه عدد زوج در پرتاب همزمان سه تاس همتراز ؟
پاسخ : احتمال آمدن عدد زوج را در سه تاس به‌ترتیب با P_A ، P_B و P_C نشان می‌دهیم چون احتمال آمدن عدد زوج در پرتاب هر تاس $\frac{۱}{۲}=\frac{۳}{۶}$ است لذا داریم .

$$P_A = P_B = P_C = \frac{۱}{۲}$$

با استفاده از رابطه ۴-۱ پاسخ مسئله به‌صورت زیر خواهد بود .

$$P_{(A\cap B\cap C)} = \frac{۱}{۲}\times\frac{۱}{۲}\times\frac{۱}{۲}=\frac{۱}{۸}$$

علت این امر نیز واضح است . تنها نیمی از حالاتی که در پرتاب تاس اول اتفاق می‌افتند مورد قبول پرتاب کننده هستند (فقط اعداد زوج) و به‌همراه هریک از این حوادث مورد قبول نیمی از حالات ممکن‌الحصول در پرتاب تاس دوم قابل قبول می‌باشند (بازهم فقط اعداد زوج) که در نتیجه در پرتاب دوتاس فقط $\frac{۱}{۴}=\frac{۱}{۲}\times\frac{۱}{۲}$ حالات قابل قبول است و $\frac{۳}{۴}$ حالات ممکن‌الحصول دیگر از نظر پرتاب کننده مردود می‌باشند . این سه حالت غیر قابل قبول حالات زیر می‌باشند .

(تاس دوم فرد و تاس اول فرد) ، (تاس دوم فرد و تاس اول زوج) و (تاس دوم زوج و تاس اول فرد) وضعیت در پرتاب سه‌تاس و یا بیشتر نیز عینا" همانند مورد دو تاس است که جهت جلوگیری از اطاله کلام از توضیح آنها خودداری می‌شود .
مثال ۲۴ - ۱ : دو گروه ده نفری هرکدام شامل سه فرد بیکار و ۷ فرد شاغل وجود دارد . احتمال آنکه در یک قرعه‌کشی دو نفر بیکار از گروه اول و دونفر شاغل از گروه دوم انتخاب شوند چقدر است .
پاسخ : اگر انتخاب بیکار از گروه اول را با A و انتخاب شاغل از گروه دوم را با B نشان دهیم حوادث A و B دو حادثه مستقل می‌باشند که حادثه مورد نظر یعنی$(A\cap B)$ را ایجاد می‌کنند . لذا برای محاسبه احتمال مورد نظر از رابطه (۴-۱) استفاده می‌کنیم .

$$P_{(A \cap B)} = P_A \times P_B$$

حال باید P_A و P_B را مستقلا" محاسبه کنیم . این دو حادثه هردو مرکب بوده و هرکدام از آنها از مجموع چند حادثه ساده ایجاد شده‌اند .

$$P_A = \frac{\binom{۳}{۲}}{\binom{۱۰}{۲}} = \frac{\frac{۳!}{۲!۱!}}{\frac{۱۰!}{۲!۸!}} = \frac{۳}{۴۵}$$

$$P_B = \frac{\binom{۷}{۲}}{\binom{۱۰}{۲}} = \frac{\frac{۷!}{۲!۵!}}{\frac{۱۰!}{۲!۸!}} = \frac{۲۱}{۴۵}$$

$$P_{(A \cap B)} = \frac{۳}{۴۵} \times \frac{۲۱}{۴۵} = ۰/۰۳۱$$

اشتراک چندحادثه مرکب غیر مستقل

شیوه‌محاسبه‌این قبیل حوادث‌مثل مورد قبل است با این تفاوت‌که چون‌مراحل‌مختلف این حوادث از یکدیگر مستقل نیستند احتمال وقوع حادثه‌ای که در مرحله دوم و یا n ام قرار دارد از حوادث‌قبلی متأثر بوده و لذا باید درمراحل بعدی‌از احتمال شرطی هرحادثه استفاده‌شود . بدین‌ترتیب رابطه کلی‌که برای محاسبه احتمال این قبیل وقایع مورداستفاده قرار می‌گیرد همان رابطه (۵ ــ ۱) می‌باشد که جهت یادآوری یکبار دیگر در قالبی جدید بیان می‌شود .

(۱ ــ ۵)

$$P_{(A \cap B \cap C \cdots\cdots)} = P_A \times P_{(B \mid A)} \times P_{(C \mid A \cap B)} \times \cdots\cdots$$

آمدن عدد زوج و کوچکتر از ۴ در پرتاب یک تاس، آمدن عدد فرد کوچکتر از ۴ در پرتاب یک تاس و غیره نمونه‌های ساده‌ای از این قبیل پیشامدها می‌باشند .

مثال ۲۵ ــ ۱ : مطلوبست احتمال آنکه در پرتاب یک تاس همتراز عدد فردی بیاید که از۴ کوچکتر باشد .

پاسـخ: اگر واقعه آمدن عدد فرد را با A و واقعه آمدن عدد کوچکتر از ۴ را با B نشـان دهیم مسلما" این دو حادثه از یکدیگر مستقل نیستند چون در شرایط طبیعی $P_B = \frac{۳}{۶} = \frac{۱}{۲}$ میباشد اما بهمحض اینکه بدانیم عددی که آمده فرد است این احتمال به $P_{(B \mid A)} = \frac{۲}{۳}$ افزایش پیدا میکند و لذا برای کل آن از رابطه (۵ ـ ۱) استفاده میشود .

$$= P_{(A \cap B)} = P_A \times P_{(B \mid A)} = \frac{۳}{۶} \times \frac{۲}{۳} = \frac{۱}{۳}$$

مثال ۲۶ ـ ۱: وضعیت شغلی و درآمدی یک جمعیت ۹۰۰ نفری مطابق جدول زیر است .

وضع شغل / وضع درآمد	مستمری بگیر	پیشه ور
کم درآمد	۴۶۰	۴۰
پردرآمد	۱۴۰	۲۶۰

مطلوب است احتمال آنکه در قرعهکشی برای سه قطعه زمین ،سه مستمری بگیر کم درآمد برنده شوند .

پاسـخ: حادثه انتخاب سه مستمری بگیر را با A و انتخاب سه کم درآمد را با B نشـان میدهیم واضح است که این دو حادثه ازیکدیگر مستقل نبوده و باید برای حل مسئله از رابطه (۵ ـ ۱) استفاده شود .

$$P_{(A \cap B)} = P_A \times P_{(B \mid A)}$$

حال دو عبارت طرف راست معادله را جداگانه حساب میکنیم .

$$P_A = \frac{\binom{۶۰۰}{۳}}{\binom{۹۰۰}{۳}} = \frac{\frac{۶۰۰!}{۳!\ ۵۹۷!}}{\frac{۹۰۰!}{۳!\ ۸۹۷!}} = \frac{۶۰۰ \times ۵۹۹ \times ۵۹۸}{۹۰۰ \times ۸۹۹ \times ۸۹۸} = ۰/۲۹۶$$

مفهوم $P_{(B \mid A)}$ آن است که اگر بدانیم سه مستمری بگیر انتخاب شده‌اند احتمال اینکه این سه نفر ازافراد کم درآمد باشند چقدر است . این‌احتمال بصورت زیر محاسبه‌می‌گردد .

$$P_{(B \mid A)} = \frac{\binom{۴۶۰}{۳}}{\binom{۶۰۰}{۳}} = \frac{۴۶۰ \times ۴۵۹ \times ۴۵۸}{۶۰۰ \times ۵۹۹ \times ۵۹۸} = ۰/۴۵$$

با قراردادن‌این ارقام در رابطه اصلی احتمال مورد نظر بدست می‌آید .

$$P_{(A \cap B)} = P_A \times P_{(B \mid A)} = ۰/۲۹۶ \times ۰/۴۵ = ۰/۱۳ = ۱۳\ \%$$

با حل این دو مثال بحث اشتراک چندحادثه مرکب به‌پایان می‌رسد و مطابق فهرست مطالب پیش بینی شده در مقدمه این فصل با اتمام این بحث باید مبحث احتمالات نیز به‌پایان خود رسیده باشد ، اما همان طور که در بحث احتمال شرطی گفته شد ، نوعی از احتمال‌شرطی تحت‌عنوان رابطه بیس وجود دارد که چون در آن از احتمالات‌مربوط به‌انواع حوادث ساده و مرکب استفاده می‌شود ضرورتا" باید در انتهای فصل مربوط به‌احتمالات مورد بحث قرار گیرد . به‌همین جهت در اینجا پیرامون این رابطه بحث نموده و با حل دو مثال در باره آن ، فصل احتمالات را به پایان می‌بریم .

همان طورکه مشاهده شد تا کنون احتمال شرطی به‌عنوان ابزاری جهت حل حوادث ساده و مرکب چند مرحله‌ای با مراحل غیر مستقل مورد استفاده قرار می‌گرفت و به‌خودی خود امری مطلوب‌نبود . به‌عبارت دیگر احتمالات شرطی که تا کنون مطرح شده‌اند مطلوبیت غیری داشته‌اند ، اما نمونه‌ای از احتمالات شرطی وجود دارد که مطلوبیت آن معمولا" نفسی بوده و البته گهگاهی هم به‌عنوان ابزارحل سایر مسائل مورد استفاده قرار می‌گیرد . از آنجا که راه حل استانداردی برای این احتمال شرطی خاص توسط فردی به‌نام بیس[1] مطرح شده است ، احتمال شرطی فوق‌الذکر معمولا" در کتب آمار تحت عنوان رابطه بیس[2] یاقانون‌بیس مطرح می‌گردد .

---

1 - Bayes

2 - Bayes Formula

ج ــ رابطه بیس

بر طبق رابطه (۵ ــ ۱) برای محاسبه احتمال یک حادثه دو مرحله‌ای با مراحل غیر مستقل رابطه کلی زیر وجود دارد:

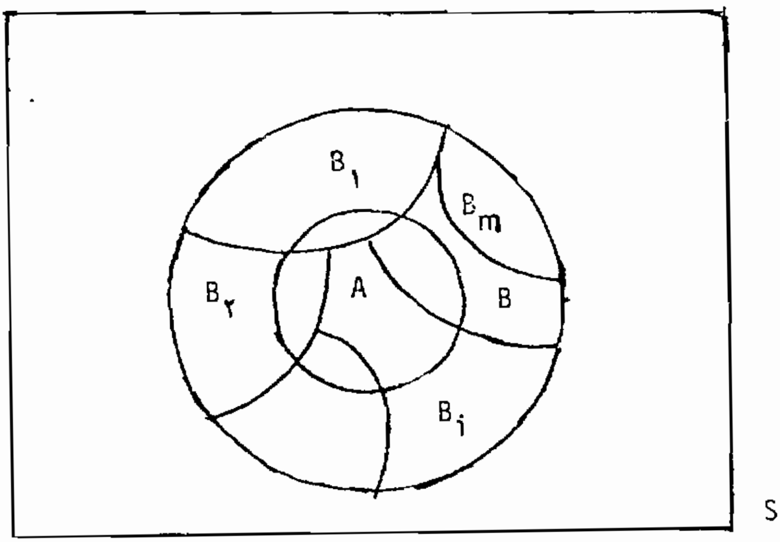
$$P_{(A\cap B)}=P_A\times P_{(B\mid A)}$$

رابطه فوق را بصورت زیر نیز می‌توان نوشت:

$$P_{(B\mid A)}=\frac{P_{(A\cap B)}}{P_A}$$

حال اگر مطابق شکل (۶ ــ ۱) حوادث A و B دو حادثه در فضای نمونه S باشند و حادثه B مجموع حادثه (B_1، B_2، ...... B_m) باشد به‌طوری که حادثه A را نیز به K قسمت تقسیم نماید در آن صورت احتمال مشروط یکی از حوادث n گانه B از طریق رابطه زیر بدست می‌آید.

$$P_{(B_i\mid A)}=\frac{P_{(B_i\cap A)}}{P_A}$$



شکل ۶-۱

واضح است که در وضعیت فوق حادثه A را به‌صورت زیر نیز می‌توان محاسبه نمود .

$$P_A = P(B_1 \cap A) + P(B_2 \cap A) + \cdots\cdots + P(B_m \cap A)$$

در نتیجه رابطه $P(B_i \mid A)$ بصورت زیر نیز قابل نمایش است :

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(B_1 \cap A) + P(B_2 \cap A) + \cdots + P(B_m \cap A)}$$

که این رابطه از نظر ریاضی بصورت زیر نیز قابل نمایش بوده و به‌آن رابطه بیس گویند .

(۱ – ۱۲)

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{\sum_{i=1}^{n} P(B_i \cap A)}$$

مثال ۲۷ – ۱ : دو جعبه‌کاملا" مشابه داریم . در جعبه اول دو مهره قرمز و درجعبه دوم یک مهره قرمز و یک مهره سفید وجود دارد . یکی از این دو جعبه را به‌تصادف انتخاب می‌کنیم و یک مهره را از آن بیرون‌می‌کشیم و ملاحظه می‌کنیم‌که این مهره قرمز است . چقدر احتمال دارد که مهره دیگری که در این جعبه قرار دارد نیز قرمز باشد ؟

پاسخ : در واقع احتمال مطلوب مثال فوق ، احتمال انتخاب جعبه اول است‌که‌بصورت زیر محاسبه می‌گردد . واقعه انتخاب مهره قرمز = C و واقعه انتخاب جعبه دوم = B و واقعه انتخاب جعبه اول = A و $P(A \mid C) = ?$

$$P(A \mid C) = \frac{P(A \cap C)}{P(A \cap C) + P(B \cap C)}$$

جملات صورت و مخرج کسر را جداگانه محاسبه نموده و در رابطه فوق قرار می‌دهیم .

$$P(A \cap C) = P(A) \times P(B \mid A) = \frac{1}{2} \times 1 = \frac{1}{2}$$
$$P(B \cap C) = P(C) \times P(B \mid C) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

و در نتیجه جواب مسئله مقدار زیر است .

$$P_{(A\mid C)}=\frac{\frac{۱}{۲}}{\frac{۱}{۲}+\frac{۱}{۴}}=\frac{\frac{۱}{۲}}{\frac{۳}{۴}}=\frac{۲}{۳}=۶۶/۶\ \%$$

مثال ۲۸ ــ ۱ : درجامعه‌ای ۰/۰۱ افراد آن مبتلا به‌بیماری $B_۱$ و ۰/۰۳ افراد آن به‌بیماری $B_۲$ مبتلا بوده و ۰/۹۶ بقیه که با $B_۳$ نشان داده می‌شوند سالم هستند . در نتیجه یک آزمایش پزشکی خاص معمولاً" بیماری ۰/۹۷ مبتلایان به‌بیماری $B_۱$ و ۱۰% مبتلایان به‌بیماری $B_۲$ تشخیص داده می‌شود و در مقابل ۵% از افراد سالم نیز به‌غلط بیمار معرفی می‌گردند . اگر از جامعه فوق یک نفر بطور تصادفی برای آزمایش مذکور انتخاب شده و مریض تشخیص داده شود ، احتمال آنکه این فرد از مبتلایان به بیماری $B_۱$ باشد چقدر است ؟

پاسخ : این حادثه که فرد انتخابی در نتیجه آزمایش مریض تشخیص داده شود را با A نشان می دهیم در نتیجه می‌توان احتمال مطلوب را بصورت $P_{(B_۱\mid A)}=?$ نشان داد، از قضیه بیس یعنی رابطه (۱۲ــ۱) داریم :

$$P_{(B_۱\mid A)}=\frac{P_{(B_۱\cap A)}}{\sum_{i=۱}^{۳}P_{(B_i\cap A)}}$$

جمله مخرج به‌سه جمله قابل بسط است‌که ابتدا این بسط را انجام داده‌و سپس مقادیرصورت و مخرج را حساب می‌کنیم :

$$\sum_{i=۱}^{۳}P_{(B_i\cap A)}=P_{(B_۱\cap A)}+P_{(B_۲\cap A)}+P_{(B_۳\cap A)}$$

$$P_{(B_۱\cap A)}=P_{(B_۱)}\times P_{(A\mid B_۱)}=(۰/۰۱)(۰/۹۷)$$

$$P_{(B_۲\cap A)}=P_{(B_۲)}\times P_{(A\mid B_۲)}=(۰/۰۳)(۰/۱۰)$$

$$P_{(B_۳\cap A)}=P_{(B_۳)}\times P_{(A\mid B_۳)}=(۰/۹۶)(۰/۰۵)$$

از جایگذاری مقادیر بدست آمده در رابطه اصلی جواب مسئله به‌دست می‌آید .

$$P_{(B_۱\cap A)}=\frac{(۰/۰۱)(۰/۹۷)}{(۰/۰۱)(۰/۹۷)+(۰/۰۳)(۰/۱۰)+(۰/۹۶)(۰/۰۵)}=۰/۱۶=۱۶\ \%$$

مسائل فصل اول

۱ ــ احتمال را تعریف نموده و نقش آنرا در عملیات آماری بیان کنید .

۲ ــ ثابت کنید که مقدار احتمال همواره $1 \geq P \geq 0$ میباشد .

۳ ــ ثابت‌کنید که مجموع‌احتمالات مربوط به‌همه پیشامدهای ممکن‌الوقوع در یک‌حادثه همواره مساوی یک خواهد بود .

۴ ــ فضای نمونه مربوط به‌پرتاب همزمان ۴ سکه را نشان دهید .

۵ ــ حادثه مرکب‌آمدن سه خط و یک‌شیر در پرتاب همزمان چهارسکه را به‌تمام حوادث ساده ممکن تجزیه کنید .

۶ ــ حوادث مرکب ذیل را به‌تمام حوادث ساده ممکن تجزیه کنید .

الف : انتخاب یک توپ قرمز از کیسه‌ای که ۵ توپ قرمز و دو توپ سفیددارد .

ب : انتخاب دو توپ قرمز از کیسه فوق

ج : وقوع دوعدد که حاصل‌جمع‌آنها مساوی ۸ باشد در پرتاب‌دوتاس همتراز

۷ ــ می‌خواهیم از بین ده نفر شامل آقایان عباسی و اکبری دونفر را بطور تصادفی و به‌قید قرعه انتخاب کنیم احتمال آنکه ابتدا آقای عباسی و سپس‌آقای اکبری انتخاب شوند چقدر است ؟ درصورتی‌که بدانیم هر برگه حاوی اسم پس ازانتخاب مجددا" داخل برگه‌دان برگردانده می‌شود .

۸ ــ مسئله قبل را برای حالتی حل کنید که برگه اول پس از انتخاب به‌داخل‌برگه‌دان برگردانده نشود .

۹ ــ مطلوبست احتمال آنکه در سه مرتبه پرتاب یک تاس همتراز به‌ترتیب اعداد ۲ و ۳ و ۴ بیایند .

۱۰ ــ احتمال هریک از حوادث ذیل را تعیین کنید .

الف : آمدن عدد فرد در یکبار پرتاب یک تاس همتراز

ب : آمدن حداقل یک شیر در دوبار انداختن یک سکه همتراز

ج : از انداختن یک جفت تاس همتراز عدد ۷ حاصل شود .

د : اگر از ۱۰۰ بار انداختن یک سکه ۵۶ شیر آمده باشد احتمال آمدن خط در پرتاب مجدد سکه .

هـ : آمدن سه‌بار عدد ۶ در ۵ بار پرتاب یک تاس همتراز

۱۱ ــ در یک موءسسه که ۵۰ نفر پرسنل دارد ، می‌خواهند ۵ دستگاه اتومبیل را به‌قید

قرعه به ۵ نفر از پرسنل آن موءسسه واگذار نمایند . تعداد نقاط نمونه n موجود در فضای نمونه مربوط به‌این حادثه را حساب کنید .

۱۲ — اگر درقرعه‌کشی مسئله ۱۱ ارزش اتومبیل‌ها بایکدیگر تفاوت‌داشته و اتومبیل‌های مزبور را به‌ترتیب‌گرانی‌قیمت به‌نفر اول تا پنجم واگذار نمایند .فضای نمونه‌مربوط به‌آزمایش چند نقطه نمونه خواهد داشت ؟

۱۳ — درآزمایشی می‌خواهند یک عدد زوج را از بین یک مجموعه اعداد ۳ رقمی انتخاب کنند ، تعداد حوادث ساده موجود دراین حادثه مرکب (انتخاب عدد زوج) را تعیین نمایید درصورتی که بدانیم‌اعداد فوق از ۵ عدد اصلی ۱ و۲و۵ و۶ و ۹ تشکیل شده و هریک از این ۵ عدد فقط یکبار مورد استفاده قرار گرفته است .

۱۴ — تعداد نقاط فضای نمونه و احتمال مربوط به‌حادثه مسئله ۱۳ را تعیین کنید .

۱۵ — ۵ مهره با ۵ رنگ مختلف را به چند طریق می‌توان مرتب کرد ؟

۱۶ — به چند طریق می‌توان ده شیی رابه‌دوگروه شامل ۴ شیی و ۶شیئی تقسیم‌نمود ؟

۱۷ — به چند طریق می‌توان یک کمیته ۵ نفری از بین ۹ نفر انتخاب کرد ؟

۱۸ — اگر تمام اعداد چهار رقمی که ممکن است با ده عد (۹ . . . . . و ۱و ۰) ایجاد شوند را در سه حالت ذیل روی کاغذهایی مشابه نوشته و در کیسه‌ای بیندازیم و سپس یک عدد را از این کیسه بطور تصادفی انتخاب کنیم ، احتمال انتخاب عدد ۴۲۹۰ را در این‌سه حالت پیدا کنید .

الف : تکرار هر عدد مجاز باشد .

ب : تکرار مجاز نباشد .

ج : رقم آخر هر عدد صفر شود و تکرار نیز مجاز نباشد .

۱۹ — در یک قرعه کشی ۷ نفر شامل ۳ دانشجوی سال چهارم و ۴ دانشجوی سال‌سوم شرکت‌دارند و قرار است یک‌کمیته سه نفری به‌قید قرعه‌انتخاب شود. احتمال آنکه درکمیته انتخابی ۲ دانشجوی سال سوم و ۱ دانشجوی سال چهارم شرکت داشته باشند چقدراست ؟

۲۰ — ۴ کتاب مختلف ریاضی ، ۶ کتاب مختلف فیزیک و دو کتاب مختلف شیمی باید به‌وسیله یک کودک خردسال در قفسه‌ای مرتب شوند . احتمالات حالات ذیل‌راتعیین کنید .

الف : کتابهای هر موضوع خاص کنار هم قرار گیرند .

ب : اقلا" کتابهای ریاضی کنارهم قرار گیرند .

۲۱ — از بین ۵ ریاضیدان و ۷ فیزیکدان ، یک کمیته علمی ۵ نفری باید تشکیل شود . احتمال آنکه کمیته انتخابی شامل ۲ ریاضی دان و ۳ فیزیک دان باشد در سه حالت ذیل چقدر است ؟

الف : همه افراد فوق می‌توانند عضو کمیته شوند .

ب : فیزیک‌دان خاصی باید عضو کمیته شود .

ج : دو ریاضی‌دان معین نمی‌توانند عضو کمیته شوند .

۲۲ ــ در مسئله ۲۰ احتمال آنکه کتابهای ریاضی کنارهم و یا همه کتابهای شیمی در کنار یکدیگر قرار گیرند چقدر است ؟

۲۳ ــ در مسئله ۲۱ احتمال آنکه همه افراد انتخابی فیزیک دان و یا همه آنها ریاضی دان باشند چقدر است ؟

۲۴ ــ برای استفاده از یکی از دروس اقتصادی دانشگاهی ۶۰ نفر در کلاس درس حاضر می‌شوند . وضعیت این ۶۰ نفر از نظر شکل تحصیل و میزان تحصیلات مقدماتی مطابق جدول زیر است .

وضعیت تحصیلی / تحصیلات مقدماتی	قابل قبول	اندک
دانشجوی رسمی	۳۰	۱۰
مستمع آزاد	۱۵	۵

یک نفر از این جمعیت به‌تصادف انتخاب می‌شود . احتمال آنکه فرد مزبور دانشجو بوده و یا تحصیلات مقدماتی قابل قبولی داشته باشد چقدر است ؟

۲۵ ــ از جعبه‌ای شامل ۶ توپ قرمز ، ۴ توپ سفید و ۵ توپ آبی یک توپ را بطور تصادفی انتخاب می‌کنیم احتمال انتخاب الف : قرمز ، ب : سفید ، ج : آبی ، د : غیر قرمز ، و هـ : قرمز یا سفید چقدر است ؟

۲۶ ــ از جعبه مسئله قبل سه توپ را به‌طور متوالی خارج می‌کنیم در دو حالت ذیل احتمال آنکه توپهای انتخابی به‌ترتیب قرمز ، سفید و آبی باشند چقدر است ؟

الف : با جایگذاری

ب : بدون جایگذاری

۲۷ ــ در یک‌کیسه ۴ توپ سفید و ۲ توپ سیاه و در کیسه دیگری ۳ توپ سفید و ۵ توپ سیاه وجود دارد . اگر از هرکیسه یک توپ انتخاب کنیم احتمال حالات ذیل را تعیین‌کنید .

الف : هردو سفید باشند ب : هردو سیاه باشند ج : یکی سفید و یکی سیاه باشد .

۲۸ ــ دو تیم فوتبال A و B دوازده مرتبه باهم بازی می‌کنند در نتیجه ۶ مرتبه تیم A و ۴ مرتبه تیم B برنده شده و در دو نوبت دیگر بازی با نتیجه مساوی به‌پایان می‌رسد. A و B تصمیم به‌انجام سه‌بازی‌جدید با یکدیگر می‌گیرند پیدا کنید احتمالات ذیل را:
الف: برد A در هر سه بازی، ب: تساوی در دو بازی، ج: A و B یک‌درمیان‌ببرند
د: B اقلا" یکبار ببرد.

۲۹ ــ احتمال آنکه در مسئله ۲۴ نفرات انتخابی دانشجوی رسمی بوده و تحصیلات مقدماتی قابل قبولی نیز داشته باشند چقدر است؟

۳۰ ــ در جعبه‌ای ۸ توپ قرمز، ۳ توپ سفید و ۹ توپ آبی قرار دارد. اگر به طور تصادفی ۳ توپ از جعبه برداشته شود احتمالات مربوط به‌حالات ذیل را در دو حالت با جایگذاری و بدون جایگذاری محاسبه کنید.

الف: هر سه توپ قرمز باشند   ب: هر سه سفید باشند.
ج: ۲ قرمز و یک سفید باشد   د: حداقل یکی سفید باشد
هـ: یکی از هر رنگ باشد   و: به‌ترتیب قرمز و سفید و آبی باشند.

۳۱ ــ یک دانشکده ۲۰۰ نفر دانشجوی سال اول، ۱۵۰ نفر دانشجوی سال دوم، ۱۰۰ نفر دانشجوی سال سوم و بالاخره ۵۰ نفر دانشجوی سال چهارم دارد. قرار است از بین دانشجویان مزبور ۹ نفر از طریق یک مراسم قرعه کشی انتخاب شوند. احتمال آنکه سه نفر اول دانشجوی سال اول و سه نفر دوم دانشجوی سال دوم بوده و بالاخره دو نفر از سه نفر سوم دانشجوی سال سوم و یک نفر باقیمانده سال چهارم باشد چقدر است؟

۳۲ ــ مهره‌های رنگی در سه جعبه مشابه به‌صورت زیر توزیع شده‌اند

شماره جعبه ←	۱	۲	۳
سرخ	۲	۴	۳
سفید	۳	۱	۴
آبی	۵	۳	۳

جعبه‌ای به‌تصادف انتخاب شده و یک مهره از آن بطور تصادفی انتخاب شده است اگر این مهره سرخ رنگ باشد احتمال اینکه جعبه دوم انتخاب شده باشد چقدر است؟

۳۳ ــ فرض کنید که از نیروی کار جامعه‌ای ۴۰٪ فارغ‌التحصیل دبستان C_1، ۵۰٪ فارغ‌التحصیل دبیرستان C_2 و ۱۰٪ فارغ‌التحصیل دانشگاه C_3 باشند. در گروه C_1 ۱۰٪ و در گروه C_2 ۵٪ و در گروه C_3 ۲٪ بیکار هستند اگر از این جامعه فردی بطور تصادفی انتخاب شده و بیکار E باشد احتمال اینکه این شخص فارغ‌التحصیل دانشگاه باشد چقدر است؟

حل مسائل فرد فصل اول

پاسخ سئوالات اول و سوم بطور کامل در مباحث فصل اول وجود دارد و لذا از تکرار آنها خودداری می‌شود .

۵ ــ حادثه مربوط به سکه‌ها به‌ترتیب با اعداد ۱ ، ۲ ، ۳ ، ۴ نشان داده شده و کلیه حوادث ساده موجود در حادثه مطلوب ، به‌صورت مجموعه زیر قابل نمایش است .

A = { (خ۴ ، خ۳ ، ش۲ ، خ۱) و (خ۴ ، خ۳ ، خ۲ ، ش۱)
(ش۴ ، خ۳ ، خ۲ ، خ۱) و (خ۴ ، ش۳ ، خ۲ ، خ۱) }

۷ ــ چون برگه‌های انتخابی مجدداً" به‌داخل برگه‌دان بر می‌گردد حادثه مطلوب یک حادثه ساده چند مرحله‌ای با مراحل مستقل است و برای حل آن از رابطه (۴ــ۱) می‌توان استفاده نمود .

$$P_{(A_1, A_2)} = (P_{A_1})(P_{A_2})$$

$P_{A_2} = ۰/۱$ = احتمال انتخاب آقای اکبری و $۰/۱ = P_{A_1}$ = احتمال انتخاب آقای عباسی

$$P\,(A_1 \text{ و } A_2) = (۰/۱)(۰/۱) = ۰/۰۱$$

۹ ــ کلیات مسئله شبیه مسئله ۷ است و لذا بازهم از رابطه (۴ــ۱) استفاده می‌کنیم .

احتمال وقوع ۲ در پرتاب اول = $P_{A_1} = \frac{1}{6}$

احتمال وقوع ۳ در پرتاب دوم = $P_{A_2} = \frac{1}{6}$

احتمال وقوع ۴ در پرتاب سوم = $P_{A_3} = \frac{1}{6}$

$$P_A=P_{(A_۱,A_۲,A_۳)}=(P_{A_۱})(P_{A_۲})(P_{A_۳})=\left(\frac{۱}{۶}\right)^۳=\frac{۱}{۲۱۶}$$

۱۱ ــ چون ترتیب انتخاب افراد لحاظ نمی‌شود و نفر اول انتخابی با نفر پنجم ازنظر دریافت جایزه وضعیت مشابهی دارند در هریک از نقاط نمونه و به‌عبارت بهتر در هریک از ترکیبات انتخابی باید حداقل یک نفر متفاوت با سایر ترکیبات وجود داشته باشد . صورت کلی فضای نمونه بصورت زیر است .

$$S=\left\{\begin{array}{l} (۱ \text{ و } ۲ \text{ و } ۳ \text{ و } ۴ \text{ و } ۵)\text{ و }(۱ \text{ و } ۲ \text{ و } ۴ \text{ و } ۵ \text{ و } ۶)\text{ و }\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ (۵۰ \text{ و } ۴۹ \text{ و } ۴۸ \text{ و } ۴۷ \text{ و } ۴۶)\cdots\cdots \end{array}\right\}$$

تعداد نقاط نمونه این فضا از رابطه (۷ــ۱) قابل محاسبه است .

$$n=\binom{۵۰}{۵}=\frac{۵۰!}{۵!(۵۰-۵)!}=۲۱۱۸۷۶۰$$

۱۳ ــ فقط دو عدد می‌تواند به‌عنوان اولین رقم سمت راست نوشته شود و برای هریک از این دو عدد چهار عدد به‌عنوان سومین رقم سمت راست نوشته شود بنابر این تعداد اعداد زوج ممکن‌الحصول عبارت از مقدار زیر خواهد بود .

$$(۳)(۴)(۲)=۲۴$$

۱۵ ــ از رابطه ۶ــ۱ مسئله حل می‌شود .

$${}_{۵}P_{۵}=\frac{۵!}{(۵-۵)!}=\frac{۱۲۰}{۱}=۱۲۰$$

۱۷ ــ چون در کمیته ترتیب انتخاب افراد تأثیری در پاسخ مسئله نـدارد از رابطـه (۷ــ۱) استفاده می‌کنیم

$$\binom{۹}{۵}=\frac{۹!}{۵!(۹-۵)!}=۱۲۶$$

۱۹ ــ حادثه مطلوب ، یک حادثه مرکب است که از مجموع چند حادثه ساده تشکیـل شده است و لذا احتمال آن از رابطه (۱ــ۱) قابل محاسبه است .

$$P_A = \frac{m}{n}$$

تعداد کل ترکیبات حاوی دو دانشجوی سال سوم و یک دانشجوی سال چهارم = m وانتخاب دودانشجوی سال سوم‌ویک‌دانشجوی سال‌چهارم = A و تعدادکل‌ترکیبات‌سه‌نفری ممکن = n

$$m = \binom{۴}{۲}\binom{۳}{۱} = (۶)(۳) = ۱۸$$
$$n = \binom{۷}{۳} = ۳۵$$
$$P_A = \frac{۱۸}{۳۵} = ۵۱\ \%$$

۲۱ ــ بازهم حادثه مطلوب ،یک حادثه‌مرکب‌است که ازمجموع m حادثه ساده تشکیل شده است . لذا از رابطه (۱ــ۱) برای حل آن استفاده می‌کنیم .

$$P_A = \frac{m}{n}$$

الف : تعداد ترکیبات ممکن شامل ۲ ریاضی دان و ۳ فیزیک دان = m و انتخاب ۲ ریاضی‌دان و ۳ فیزیک دان = A و تعداد ترکیبات ۵ نفری ممکن = n

$$m = \binom{۵}{۲}\binom{۷}{۳} = (۱۰)(۳۵) = ۳۵۰$$
$$n = \binom{۱۲}{۵} = ۷۹۲$$
$$P_A = \frac{۳۵۰}{۷۹۲} = ۴۴/۱\%$$

ب : با این فرض در واقع انتخاب از بین فیزیک‌دانها به ۲ نفر از ۶ نفر و در کل انتخاب به ۴ نفر از ۱۱ نفر محدود می‌شود .

$$m = \binom{۵}{۲}\binom{۶}{۲} = (۱۰)(۱۵) = ۱۵۰$$

$$n = \binom{۱۱}{۴} = ۳۳۰$$

$$P_A = \frac{۱۵۰}{۳۳۰} = ۴۵/۱\ \%$$

ج : با این فرض تعداد ریاضی‌دانهای قابل انتخاب از ۵ نفر به سه نفر و تعداد کل افراد قابل انتخاب از ۱۲ نفر به ۱۰ نفر کاهش پیدا می‌کند .

$$m = \binom{۳}{۲}\binom{۷}{۳} = (۳)(۳۵) = ۱۰۵ \quad \text{و} \quad n = \binom{۱۰}{۵} = ۲۵۲$$

$$P_A = \frac{۱۰۵}{۲۵۲} = ۴۲\ \%$$

۲۳ ــ حادثه مورد نظر، اجتماع دو حادثه ناسازگار مرکب است و مقدار احتمال آن از رابطه (۱۰ــ۱) محاسبه می‌گردد.

$$P_{(A_۱ \cup A_۲)} = P_{A_۱} + P_{A_۲}$$

$$\text{احتمال انتخاب ۵ ریاضی‌دان از جمع مزبور} = P_{A_۱} = \frac{m_۱}{n} = \frac{\text{تعداد ترکیبات ۵ نفری از ریاضی‌دانان یا } m_۱}{\text{تعداد ترکیبات ۵ نفری ممکن یا } n} = \frac{\binom{۵}{۵}}{\binom{۱۲}{۵}} = \left(\frac{۱}{۷۹۲}\right)$$

$$\text{احتمال انتخاب ۵ فیزیکدان از این جمع} = P_{A_۲} = \frac{m_۲}{n} = \frac{\text{تعداد ترکیبات ۵ نفری از فیزیک‌دانان یا } m_۲}{\text{تعداد ترکیبات ۵ نفری ممکن یا } n} = \frac{\binom{۷}{۵}}{\binom{۱۲}{۵}} = \left(\frac{۲۱}{۷۹۲}\right)$$

$$P_A = P_{(A_۱ \cup A_۲)} = \left(\frac{۱}{۷۹۲}\right) + \left(\frac{۲۱}{۷۹۲}\right) = ۳\ \%$$

۲۵ ــ الف : حادثه مورد نظر حادثه مرکبی است که از مجموع m حادثه ساده تشکیل شده و احتمال آن از رابطه (۱ــ۱) محاسبه می‌شود.

$$P_{ق} = \frac{m_{ق}}{n} = \frac{۶}{۱۵} = ۴۰\ \%$$

ب : احتمال این حادثه نیز مثل بند الف قابل محاسبه است.

$$P_{ـ} = \frac{m_{ـ}}{n} = \frac{۴}{۱۵} = ۲۷\ \%$$

ج : احتمال مورد نظر مثل دو بند قبلی قابل محاسبه است.

$$P_{آ} = \frac{m_{آ}}{n} = \frac{۵}{۱۵} = ۳۳\ \%$$

د : این حادثه در واقع حادثه سفید یا آبی می‌باشد و چون دو حادثه سفید و یا آبی از هم مستقل هستند از رابطه (۱۰ــ۱) استفاده می‌کنیم.

$$P(\text{غیرقرمز}) = P(\text{سفید} \cup \text{آبی}) = P_{ـ} + P_{آ} = \%۲۷ + \%۳۳ = ۶۰\ \%$$

هـ : این بند نیز مثل بند " د " قابل حل است .

$$P(\text{قرمز} \cup \text{سفید}) = P_{\text{ق}} + P_{\text{س}} = (۴۰+۲۷)\% = ۶۷\%$$

۲۷ ــ مسئله مربوط به یک حادثه مرکب است که از اشتراک دو حادثــه مرکب مستقل ایجاد شده است و لذا احتمال آن از رابطه (۴ــ۱) قابل محاسبه است .

الف :

$$P(\text{سفید}۱ \cap \text{سفید}۲) = \left(P_{\text{سفید}۱}\right)\left(P_{\text{سفید}۲}\right)$$

$$P_{\text{سفید}۱} = \frac{۴}{۶} \quad \text{و} \quad P_{\text{سفید}۲} = \frac{۳}{۸}$$

$$P_{\text{هردوسفید}} = P(\text{سفید}۱ \cap \text{سفید}۲) = \left(\frac{۴}{۶}\right)\left(\frac{۳}{۸}\right) = \frac{۱}{۴} = ۲۵\%$$

ب :

$$P(\text{سیاه}۱ \cap \text{سیاه}۲) = \left(P_{\text{سیاه}۱}\right)\left(P_{\text{سیاه}۲}\right)$$

$$P_{\text{سیاه}۱} = \frac{۲}{۶} \quad \text{و} \quad P_{\text{سیاه}۲} = \frac{۵}{۸}$$

$$P_B = P(\text{سیاه}۱ \cap \text{سیاه}۲) = \left(\frac{۲}{۶}\right)\left(\frac{۵}{۸}\right) = \frac{۱۰}{۴۸} = ۲۱\%$$

ج : حادثه موردنظر به‌دوحادثه از نوع موجود در بندهای‌الف وب قابل تجزیه‌است .

$$P(\text{سفید} \cap \text{سیاه}) = P(\text{سفید}۱ \cap \text{سیاه}۲) + P(\text{سیاه}۱ \cap \text{سفید}۲)$$

$$P(\text{سفید}۱ \cap \text{سیاه}۲) = \left(\frac{۴}{۶}\right)\left(\frac{۵}{۸}\right) = \frac{۲۰}{۴۸}$$

$$P(\text{سیاه}۱ \cap \text{سفید}۲) = \left(\frac{۲}{۶}\right)\left(\frac{۳}{۸}\right) = \frac{۶}{۴۸}$$

از جایگذاری ارقام‌بدست آمده در رابطه‌کلی فوق مقدار احتمال موردنظر بــدست می‌آید .

$$P(\text{سفید} \cap \text{سیاه}) = \frac{۲۰}{۴۸} + \frac{۶}{۴۸} = \frac{۲۶}{۴۸} = ۵۴\%$$

۲۹ ــ حادثه مورد نظر یک حادثه مرکب تشکیل شده از اشتراک دو حادثه مرکب غیــر مستقل از یکدیگر می‌باشد و لذا احتمال آن از رابطه (۵ ــ ۱) محاسبه می‌شود .

$$P_{(A \cap B)} = P_A \cdot P_{(B \mid A)}$$

A = انتخاب یک دانشجوی رسمی

B = انتخاب یک دانشجو با تحصیلات مقدماتی قابل قبول

$$P_A = \frac{۴۰}{۶۰} \quad \text{و} \quad P_{(B \mid A)} = \frac{۳۰}{۴۰}$$

$$P_{(A \cap B)} = \left(\frac{۴۰}{۶۰}\right)\left(\frac{۳۰}{۴۰}\right) = \frac{۳۰}{۶۰} = ۵۰\ \%$$

۳۱ـ حادثه مورد نظر یک حادثه مرکب سه مرحله‌ای با مراحلی غیر مستقل است و احتمال آن از رابطه (۵ ـ ۱) محاسبه می‌شود .

$$P_A = P_{(A_1 \cap A_2 \cap A_3)} = P_{A_1} \cdot P_{(A_2 \mid A_1)} \cdot P_{(A_3 \mid A_2 \cap A_1)}$$

A_1 = انتخاب سه نفر دانشجوی سال چهارم

A_2 = انتخاب سه دانشجوی سال دوم

A_3 = انتخاب دو دانشجوی سال سوم و یک دانشجوی سال چهارم

$$P_{A_1} = \frac{\binom{۲۰۰}{۳}}{\binom{۵۰۰}{۳}} = \frac{(۲۰۰)(۱۹۹)(۱۹۸)}{(۵۰۰)(۴۹۹)(۴۹۸)} = ۲\ \%$$

$$P_{(A_2 \mid A_1)} = \frac{\binom{۱۵۰}{۳}}{\binom{۴۹۷}{۳}} = \frac{(۱۵۰)(۱۴۹)(۱۴۸)}{(۴۹۷)(۴۹۶)(۴۹۵)} = ۳\ \%$$

$$P_{(A_3 \mid A_1 \cap A_2)} = \frac{\binom{۱۰۰}{۲}\binom{۵۰}{۱}}{\binom{۴۹۴}{۳}} = \frac{(۴۴۵)(۵۰)}{۱۹۹۷۰۴۴۴} = ۱\ \%$$

$$P_A = P_{(A_1 \cap A_2 \cap A_3)} = (۰/۰۲)(۰/۰۳)(۰/۰۱) = ۰/۰۰۰۶\ \%$$

۳۲ـ احتمال مطلوب یک احتمال شرطی است که از رابطه بیس حل می‌شود .

$$P_{(C_{۳} \mid E)} = \frac{P_{(C_{۳} \cap E)}}{\sum_{i=۱}^{۳} P_{(C_i \cap E)}}$$

جملات موجود در صورت و مخرج کسر را جداگانه حساب می‌کنیم .

$$P_{(C_{۳} \cap E)} = P_{C_{۳}} \cdot P_{(E \mid C_{۳})} = (\text{۱۰ ٪})(\text{۲ ٪}) = \text{۰/۰۰۲}$$

$$P_{(C_{۱} \cap E)} = P_{C_{۱}} \cdot P_{(E \mid C_{۱})} = (\text{۴۰ ٪})(\text{۱۰ ٪}) = \text{۰/۰۴}$$

$$P_{(C_{۲} \cap E)} = P_{C_{۲}} \cdot P_{(E \mid C_{۲})} = (\text{۵ ٪})(\text{۵ ٪}) = \text{۰/۰۲۵}$$

$$P_{(C_{۳} \mid E)} = \frac{\text{۰/۰۰۲}}{\text{۰/۰۰۲} + \text{۰/۰۴} + \text{۰/۰۲۵}} = \text{۳ ٪}$$

پاسخ مسائل زوج فصل اول

۶ـ الف ــ ۵ حادثه   ب : ده حادثه   ج : ۵ حادثه

۸ــ ۱/۱ %

۱۰ـ الف : ۵۰ % ، ب : ۷۵% ج : ۱۷ % د : ۴۴ % هـ : ۳۲ %

۱۲ــ ۲۵۴۲۵۱۲۰۰

۱۴ــ N = ۶۰ و P_A = ۴۰ %

۱۶= ۲۱۰

۱۸ـ الف : ۰/۰۱ % ، ب : ۰/۰۲ % ، ج : ۰/۲ %

۲۰ـ الف : ۰/۰۴ % ، ب : ۱/۸ %

۲۲ــ ۱۸ %

۲۴ــ ۶۳ %

۲۶ـ الف : ۳/۵ % ، ب : ۴ %

۲۸ـ الف : ۱۵/۴ % ، ب : ۸ % ، ج : ۱۳ % ، د : ۶۷ %

۳۰ـ با جایگذاری

الف : ۶/۴ % ، ب : ۰/۰۳ % ، ج : ۷/۲ % ، د : ۳۹ %

هـ ، ۱۶ % ، و : ۳ %

بدون جایگذاری

الف : ۵ % ، ب : ۰/۰۹ % ، ج : ۷/۳ % ، د : ۴۰ %

هـ : ۱۹ % ، و : ۳/۲ %

۳۲ــ ۵۰ %

فصل دوم

مشخصه‌های تمرکز و پراکندگی

مقدمه:

چنانکه در مقدمه کتاب گفته شد هدف عملیات آماری معمولا "رسیدن به مشخصه‌ها و پارامترهای خاصی از جامعه است که بر آساس آنها بتوان قضاوت نهائی را در مورد آن جامعه به عمل آورد. میانگین واریانس و جذر آن یعنی انحراف معیار، مد و غیره نمونه‌هائی از مشخصه‌های فوق‌الذکر می‌باشند.

همچنین در مباحث قبلی گفته شد که محقق برای دست یافتن به مشخصه‌های مزبور در مورد جامعه مورد مطالعه‌اش در بیشتر اوقات به دلایل مختلفی مجبور است از اطلاعاتی استفاده کند که از یک یا چند نمونه به دست آمده‌اند. اجبار فوق‌الذکر به‌دو صورت بر کار محقق مورد بحث‌اثر می‌گذارد. اولا"تحقیق براساس نمونه‌سبب می‌شودکه اونتواند مشخصه‌های جامعه را مستقیم و بلاواسطه بر مبنای اطلاعات حاصله از خود جامعه محاسبه کند. این اثر به‌نوبه خود اثر دوم را پدید می‌آورد و آن اثر این است‌که نتایج حاصله در مورد جامعه با تقریب و احتمال همراه بوده و محقق هیچگاه نمی‌تواند یقین کند که این نتایج بر جامعه مورد مطالعه‌اش صدق قطعی دارند. علت این امر نیز واضح است. چنانکه گفته شد برای اخذ نتیجه در مورد وضع یک جامعه اطلاعات اولیه‌ای در دسترس او هستند که از یک یا چند نمونه بدست آمده و فقط در مورد همان نمونه و یا نمونه‌ها صدق قطعی دارند. از طرف‌دیگر این مسئله سبب می‌شود که محقق همان نتایج غیر مسلم و تقریبی ممکن‌الحصول در مورد جامعه را با واسطه و در دو مرحله مختلف تحصیل نماید. زیرا در این روش تحقیق، بین محقق و نتایج‌مربوط به‌جامعه‌چیزی به‌نام اطلاعات‌پرورش یافته‌مربوط به نمونه‌و به‌عبارت بهتر مشخصه‌های مربوط به‌نمونه حائل می‌شود که این اطلاعات مربوط به‌نمونه در واقع در

ساحل رودخانه‌ای قرار دارند که ساحل دیگر آن را نتایج تقریبی مربوط به‌جامعه تشکیل می‌دهند . بدین ترتیب محقق مورد بحث در بیشتر اوقات مجبور است ابتدا با پرورش اطلاعات خام حاصل از نمونه ازطریق محاسبه مشخصه‌های موردلزوم ، خود را به‌ساحل نمونه این‌رودخانه‌رسانده و سپس با نصب پلی برروی این رودخانه خودرا به‌ساحل دیگرآن‌برساند و اطمینان حاصل کند که بالاخره به‌نتایجی هرچند تقریبی در مورد جامعه‌دست‌یافته‌است .

بدین‌ترتیب وقتی که در عملیات آماری از اطلاعات خام حاصل از نمونه به عنوان اطلاعات اولیه استفاده می‌گردد ، اطلاعات پرورش یافته مربوط به نمونه واسطه رسیدن به‌نتایج تقریبی ممکن‌الحصول در مورد جامعه می‌باشند و لذا اولین کار یک محقق آماری آن است که اطلاعات اولیه در اختیار خود را طوری پرورش دهد که بتواند بر مبنای آنها و با توسل به‌روشهایی‌که موضوع بحث فصول‌بعدی هستند در موردجامعه استنتاج‌نماید . به‌همین دلیل دانستن چگونگی پرورش اطلاعات حاصله از نمونه و به‌عبارت بهتر محاسبه مشخصه‌ها و پارامترهای مربوط به‌نمونه که موضوع بحث این فصل است شرط لازم و مقدمه ضروری هر نوع عملیات آماری می‌باشد .

چنانکه بعداً" خواهدآمد در بیشتر اوقات برای دست‌یابی به‌هریک ازمشخصه‌های‌جامعه اصلی ، دانستن مقدار و یا مقادیر همان مشخصه در مورد نمونه ر یا نمونه‌ها ضرورت داشته و مبنای کار محسوب می‌گردد . به‌عنوان مثال اگر میانگین درآمدی معادل ۷۰۰۰۰ ریال برای کلیه افراد شاغل ساکن شهر تهران بیانگر یک وضعیت درآمدی رضایتبخش باشد و یک محقق بخواهد ازطریق نمونه‌گیری در مورد رسیدن و یا نرسیدن میانگین درآمد جامعه شاغلین‌فوق به‌حد نصاب مزبور تحقیق کند ، ابتدا باید میانگین درآمد افراد موجود در نمونه‌را به‌دست آورده و سپس از طریق تئوری تخمین و یا آزمون فرضیه‌ها که هریک از آنها یک پل رابط بین نمونه و جامعه می‌باشند میانگین جامعه را به‌صورت تقریبی به‌دست آورد .

ممکن است با مشاهده مثال فوق این سئوال به‌ذهن متبادر گردد که بر چه اساسی مشخصه خاصی مثل میانگین در مثال فوق به‌عنوان ملاک قضاوت در مورد کیفیت درآمدافراد جامعه مورد مطالعه انتخاب شده و کمیت ۷۰۰۰۰ ریال چگونه تعیین گردیده است ؟ شکل کلی‌تر این سئوال آن است که ملاک قضاوت و به‌عبارت بهتر مشخصه‌های آماری که بر مبنای آنها در موردچگونگی وضعیت جامعه مورد مطالعه‌از یک جهت‌خاص مثل درآمد و غیره‌قضاوت می‌شود بر اساس چه ضابطه‌ای انتخاب می‌گردند ؟ پاسخ این سئوال در دو اصل هدف‌تحقیق و خواص مشخصه‌های آماری خلاصه می‌شود . این دو اصل در قالب دو مثال که یکی از آنها همان مثال قبلی می‌باشد بیشتر توضیح داده می‌شوند . در مثال قبل هدف مطالعه آماری می‌تواند این مسئله باشد که آیا کل درآمدی که بین جامعه شاغلین ساکن شهر تهران توزیع

می‌گردد ، جوابگوی نیازهای متعارف آنان می‌باشد یا خیر؟ با توجه به‌این هدف و نیز با در نظر گرفتن خواص مشخصه‌های آماری مختلف به‌نظر می‌رسد که میانگین بهترین ملاک قضاوت باشد زیرا به‌راحتی می‌توان استدلال نمود که اگر بطور متوسط به‌هریک از افراد این جامعه مقدارخاصی به‌عنوان درآمد پرداخت گردد جامعه مزبور از نظر کل درآمد در وضعیت‌قابل قبولی قرار دارد و می‌توان گفت که کل درآمدی که بین افراد جامعه توزیع می‌گردد جوابگوی نیازمندیهای آنان بوده و اگر در این جامعه افراد بسیار کم درآمدی دیده می‌شوند ریشه مسئله را در علل و عوامل دیگری باید جستجو نمود . مقدار مطلوب هر مشخصه که در مثال فوق مبلغ ۷۰۰۰۰ ریال برای میانگین در نظر گرفته شده با توجه به‌ضوابط غیرآماری ازقبیل تجربیات شخصی ، نظریات مختلف اقتصادی و غیره تعیین می‌گردد و از حیطه بحث آماری خارج است . به‌عنوان مثال فرض‌کنید که هدف یک مطالعه آماری تحقیق پیرامون بلند قد بودن افراد یک جامعه خاص باشد . تجربه نشان می‌دهد که اگر متوسط‌قد افراد یک جامعه ۱۸۰ سانتیمتر و یا بیشتر باشد می‌توان آن را جامعه افراد بلند قد دانست .

نتیجه‌ای که ازمطالب فوق گرفته می‌شوداین‌است‌که اگر فردی‌بخواهد در مورد‌یک مسئله خاص مربوط به‌یک جامعه از طریق مطالعه آماری تحقیق کند اولا" باید با خواص مشخصه‌های آماری مختلف آشنا باشد تا بتواند ملاک قضاوت مناسبی را از بین این مشخصه‌ها انتخاب نماید و ثانیا"طرزمحاسبه این مشخصه‌ها را بداند زیراهمانطورکه‌گفته شد بدست‌آوردن‌مقدار تقریبی هریک از مشخصه‌های مربوط به‌جامعه منوط به‌محاسبه همان مشخصه و یا مشخصه‌های دیگری در مورد نمونه می‌باشد . به‌همین دلیل این فصل به‌توضیح خواص مهم و طرز محاسبه این مشخصه‌ها اختصاص یافته است .

در بیشتر مطالعات آماری وضعیت به‌گونه‌ای متفاوت با دو مثال قبل است بدین معنی که حصول نتیجه در مورد فقط یک پارامتر و یک مشخصه جامعه از قبیل میانگین آن تأمین‌کننده هدف تحقیق نیست و ممکن است نیل به‌هدف محتاج محاسبه دو پارامتر و یا بیشتر و یا حتی آمیزه و ترکیبی از دو یا چند پارامتر و مشخصه باشد .مثالی‌را که میانگین‌درآمدی معادل ۷۰۰۰۰ ریال ملاک خوبی درآمد افراد جامعه بود مجددا" بخاطر بیاورید . فرض‌کنید هدف محقق آماری تغییر کرده و هدف جدید تحقیق این مسئله باشد که " آیا هریک از افراد این جامعه درآمدی نسبتا" کافی و نه چندان کم و فقیرانه به‌دست می‌آورند ؟" به‌نظر شما با هدف جدیدی‌که برای مطالعه آماری در نظر گرفته شده‌است آیا می‌توان گفت که اگر به‌هریک از افراد جامعه بطور متوسط درآمدی معادل ۷۰۰۰۰ ریال برسد هدف فوق تأمین شده‌است ؟ برای گفتن پاسخ عجله نکنید . فرض‌کنید تعداد افراد جامعه مورد مطالعه ۱۰۰۰۰۰ نفربوده و محقق پس از انجام مطالعات لازم به‌این نتیجه برسد که به‌احتمال زیاد میانگین درآمد

افراد این جامعه معادل مقدارفوق می‌باشد یعنی کل افراد این جامعه معادل ۷۰۰۰۰۰۰۰۰۰ ریال درآمد ماهیانه دارند . با حصول چنین نتیجه‌ای نمی‌توان هدف محقق را تأمین شده تلقی نمود زیرا اگرچه درآمد متوسط ۷۰۰۰۰ ریال برای هر نفر درآمد قابل قبولی است اما دلیلی در دست نیست که درآمد کل ۷ میلیارد ریالی مزبور طوری بیـن افراد جامعه‌توزیع شودکه به‌هریک ازافراد چیزی درحدود ۷۰۰۰۰ ریال فوق‌الذکریعنی مثلا" ۵۰۰۰۰، ۶۰۰۰۰، ۷۰۰۰۰، ۸۰۰۰۰، یا ۹۰۰۰۰ برسد بلکه ممکن است ۱۰۰۰۰ نفر از افراد این جامعه هریـک رقمی معادل ۵۰۰۰۰۰ ریال درآمد ماهیانه داشته‌باشند و باقیمانده درآمد بین ۹۰۰۰۰نفر دیگر توزیع شده و به‌هرنفرتقریبا" معادل ۲۲۰۰۰ ریال برسد .در حالت اخیرعلیرغم وجود میانگین درآمدی معادل ۷۰۰۰۰ ریال بهیچوجه نمی‌توان‌گفت هدف محقق تأمین شده‌است . زیرا برای تأمین این هدف اولا" باید میانگین درآمد افراد جامعه معادل ۷۰۰۰۰ ریال بوده و ثانیا" ثابت شود که درآمد هیچیک از افراد جامعه با مقدار متوسط فوق فاصله بسیار زیـاد و غیر قابل قبول ندارد . در حالی که در این مثال فقط بند اولا" بطور تقریبی تأیید شـده است .

بدین ترتیب در اکثر مطالعات آماری دست‌یابی به‌نتایجی در مورد فقط یک مشخصـه جامعه برای نیل به‌مقصود کافی نیست بلکه این هدف وقتی محقق می‌شود که مقادیر تقریبی حداقل دو مشخصه جامعه در مورد صفت یا خاصیت و یا متغیری که مطالعه می‌گـردد تعییـن گردد . در مثال فوق درآمد هریک از افراد جامعه متغیر یا صفت مورد مطالعه بودمعمولا"یکی از این دو مشخصه ، مشخصه‌ای است که وضعیت تمرکز مقادیر متغیر مورد مطالعـه را مشخـص می‌نماید وقتی‌که این مقادیر برحسب اندازه مرتب شده‌باشند . به‌عنوان مثال این‌مشخصه‌ها تعیین می‌کنند که متوسط مقادیر متغیر مورد مطالعه چه مقداری می‌باشد یا اینکه کدامیـک‌از مقادیر متغیر مورد مطالعه بیش از سایر مقادیر تکرار شده است . فرضا" در مثال فوق ممکـن است به‌وسیله این مشخصه نتیجه‌گیری شود که گروه افرادی که درآمد ماهیانه معادل ۵۰۰۰ تومان دارند تعدادشان از تک تک گروههای دیگر درآمدی بیشتر است . این مشخصه‌هـا که به‌دلیل‌خاصیت اصلی‌که دارند . درآمار به‌مشخصه‌های‌تمرکز معروف هستندویکی‌از موضوعات اصلی مورد بحث در این فصل می‌باشند . دومین مشخصه‌ای که معمولا" در مطالعات آمـاری برای نیل به‌هدف موردنیاز است باید وضعیت پراکندگی مقادیر متغیر مورد مطالعه را حـول مرکز این مقادیرمشخص نماید . فرضا" در مثال فوق این مشخصه تعیین می‌کند کهاگرمیانگین درآمد افراد ۷۰۰۰۰ ریال است آیا درآمد همه افراد مبلغی نزدیک به‌همین ۷۰۰۰۰ است یا اینکه سایر ارقام درآمدی فاصله بسیاری با این مقدار متوسط دارند . گروهی از مشخصه‌های آماری با انگیزه بررسی وضعیت پراکندگی مقادیر متغیر مورد مطالعه وضع شده‌اند و به‌همین

دلیل به‌آنها مشخصه‌های پراکندگی گفته می‌شود . بحث پیرامون این مشخصه‌ها یکی دیگراز موضوعات اصلی مورد بحث در این فصل می‌باشد .

از آنجا که‌معمولا" اطلاعات جمع‌آوری شده ازنمونه و یا جامعه‌به‌همان صورت ابتدائی که اخذ شده‌اند قابل استفاده برای عملیات آماری مختلف نبوده و وضعیت بسیار آشفته‌ای دارند درمرحله‌دوم عملیات‌آماری یعنی‌مرحله‌پرورش اطلاعات‌جمع‌آوری‌شده وبامرحله‌محاسبه پارامترهای‌مورد نیاز ازنمونه، به‌نظرمی‌رسد سازمان‌دادن این‌آشفته بازار اطلاعات و به‌عبارت بهتر طبقه‌بندی اطلاعات جمع‌آوری شده بر هر کار دیگری تقدم دارد . به‌دلیل وجود همین تقدم و مراعات آن ، در این فصل ابتدا چگونگی طبقه‌بندی اطلاعات مورد بحث قرارگرفته و سپس پیرامون مشخصه‌های تمرکز و پراکندگی‌که موضوعات اصلی بحث در این‌فصل هستند صحبت خواهد شد .

الف : طبقه‌بندی اطلاعات :

عملیات آماری ماهیتا" با کمیات و اعداد سرکار دارند و برای استفاده از این‌عملیات در هر نوع تحقیقی الزاما" باید اطلاعات جمع‌آوری شده بصورت کمی باشند و اگر چنین نیستند ابتدا باید این اطلاعات کیفی را به‌صورت کمی تبدیل کرده و سپس در عملیات آماری مورد استفاده قرار گیرند . به‌عنوان مثال اگر رنگ پوست افراد یک جامعه مورد مطالعه بوده‌واطلاعات حاصله ۵حالت سفید، سفید مایل به‌سبزه، سبزه، سبزه‌مایل به‌سیاه‌وبالاخره‌سیاه است را شامل می‌شوند ،این اطلاعات وقتی قابل‌استفاده در عملیات آماری هستند که هریک‌ازاین حالات به‌صورت یک کمیت تعریف گردند . در مثال فوق می‌توان حالات پنج‌گانه مزبور را با اعداد ۱ الی ۵ و یا هر عدد مناسب دیگری‌نشان داده و سپس با طبقه‌بندی‌آنها مشخصه‌های آماری مورد نظر را محاسبه کرده و سایر عملیات آماری را انجام داد .

با توجه به‌این خصوصیت عملیات آماری یعنی ماهیت کمی آنها ، مناسب‌ترین نوع طبقه‌بندی اطلاعات خام حاصل از نمونه ، طبقه‌بندی آنها بر اساس مقادیر مختلفی است که برای متغیر تصادفی مورد نظر از نمونه به‌دست‌آمده است . چگونگی انجام این عمل بدین‌ترتیب است که همانند جدول (۱-۲) در ستون اول مقادیر مختلفی که برای متغیر تصادفی از نمونه به‌دست‌آمده‌اند به‌ترتیب‌ازکوچک به‌بزرگ‌مرتب‌می‌گردند. سپس درمقابل هریک از این مقادیر تعداد اعضاء نمونه که اندازه آنها مساوی آن مقدار خاص می‌باشد را می‌نویسند . به‌هریک ازاعدادی که دراین ستون و درمقابل یک مقدارخاص از متغیر تصادفی نوشته‌می‌شود فراوانی مطلق[1] آن مقدار خاص متغیر تصادفی گفته می‌شود . معمولا" در جداول آماری برای

---

1 - Absolute Frequency

نمایش این ستون از حرف F که اولین حرف کلمه فراوانی یا تکرار به‌زبان انگلیسی می‌باشد استفاده می‌کنند . در نتیجه ثبت اطلاعات به‌شیوه فوق‌الذکر یک جدول دوستونی به دست می‌آید که به‌هریک از سطورآن یک‌طبقه گفته‌می‌شود .به‌عنوان مثال جدول (۱–۲) یک جدول ۱۲ طبقه‌ای می‌باشد البته معمولا" بنا به‌ضرورت به‌این جدول ستونهای دیگری نیز افزوده می‌گردد که بعدا" در این مورد بیشتر صحبت خواهد شد .

مهمترین مسئله‌ای که معمولا" یک آمارگر در هنگام طبقه‌بندی اطلاعات پیش رو دارد این است که اطلاعات خام در اختیار خود را در چند طبقه جای دهد که از یک طرف حجم محاسبات موردنیاز به‌خاطر کثرت طبقات خیلی زیاد نشود و از طرف دیگر تعداد طبقات آنقدرکم نباشندکه نتیجه عملیات را از دقت مطلوب محقق ساقط نمایند . غالب صاحبنظران بر این عقیده‌اند که اگر این تعداد بین ۲۰–۱۰ طبقه باشد می‌توان گفت دو ضابطه مزبور رعایت شده است یعنی اولا" با این تعداد طبقات محاسبه مشخصه‌های آماری به‌راحتی‌صورت گرفته و از حدود حوصله آمارگر بیرون نیست و ثانیا" نتیجه حاصله از عملیات نیز از دقت قابل قبولی برخوردارخواهد بود . در عمل جداول ۱۰ طبقه (۱۰سطری) بیش از سایرجداول مورد استفاده قرارمی‌گیرند . اینک در قالب چند مثال چگونگی تلخیص اطلاعات در ۲۰–۱۰ و یا هر تعداد دلخواه دیگری از طبقات توضیح داده می‌شوند .

مثال ۱ – ۲ : به‌منظور بررسی وضع جمعیت روستاهای یک منطقه کشور ۴۰ روستا به عنوان نمونه انتخاب شده و نتیجه شمارش جمعیت در این نمونه ۴۰ تایی به‌صورت زیر بوده‌است .

۲۹۵، ۲۷۰، ۲۹۵، ۲۸۰، ۴۵۴، ۶۰۵، ۵۶۶، ۲۵۵، ۳۷۰، ۴۰۱، ۲۹۰، ۲۶۷، ۳۹۸، ۴۲۸

۵۶۶، ۳۸۵، ۲۷۷، ۵۴۴، ۴۸۵، ۳۶۰، ۳۸۵، ۵۶۸، ۳۸۵، ۳۲۶، ۲۶۳، ۲۷۴، ۵۹۵، ۴۶۷،

۵۲۰، ۵۵۰، ۳۶۶، ۲۹۷، ۲۹۸، ۲۶۸، ۴۸۳، ۳۸۴، ۲۸۲، ۴۶۷، ۴۸۱، ۵۹۰،

این اطلاعات را به‌منظور انجام عملیات آماری بعدی طبقه‌بندی کنید .

پاسخ : اگر بخواهیم اطلاعات خود را در یک جدول ۱۰ طبقه‌ای نشان دهیم مسلما" در مواردی مثل مورد فوق باید هر طبقه تعدادی از مقادیر متغیر تصادفی را که در اینجا تعداد جمعیت هر روستا می‌باشد در بر گیرد . به‌عنوان مثال می‌توان فاصله دو عدد (۲۸۰–۲۵۰) را به‌عنوان اولین‌طبقه انتخاب‌نموده و تعداد روستاهائی‌که جمعیت آنها در این فاصله‌قرار دارد را به‌عنوان فراوانی مطلق این طبقه در ستون دوم در مقابل آن نوشت . به‌تفاضل این دو عدد عرض طبقه اول گفته می‌شود .

C_1 = عرض طبقه اول = ۲۸۰ - ۲۵۰ = ۳۰

روش اصولی برای پیدا کردن عرض طبقات‌مختلف درچنین مواردی آن است‌که تفاضل

کمترین مقدار متغیر تصادفی و بیشترین مقدار آن را بدست آورده و بر تعداد طبقات مورد نظر تقسیم نمائیم . در مورد فوق حداقل مقدار مشاهده شده برای متغیر تصادفی ۲۵۵ و حداکثر مقدار آن ۵۹۵ بوده و تعداد طبقات نیز معادل ۱۰ اختیار گردیده ، نتیجتا" در این مورد داریم :

$$۶۰۵ - ۲۵۵ = ۳۵۰ \div ۱۰ = ۳۵$$

بدین‌ترتیب در مثال فوق باید اولین طبقه از ۲۵۵ شروع شده و به ۲۹۰ ختم گردد و به‌همین ترتیب طبقه دوم از ۲۹۰ شروع شده و به ۳۲۵ ختم گردد و بالاخره آخرین طبقه با عدد ۶۱۵ ختم شود . اما معمولا" رسم بر این است که حدود طبقات و عرض آنها را با اعداد ساده‌تر و مشخص تری نمایش داده و سعی شود حتی‌المقدور از این جهت جدول واضح‌تر و گویاتری تنظیم گردد . برای این کار در مثال فوق کافیست عرض هر طبقه را از ۳۵ به ۳۰ کاهش داده و حد پائین طبقه اول را بجای عدد ۲۵۵ ، ۲۵۰ در نظر گرفت . بدین ترتیب اطلاعات فوق به‌صورت جدول ۱ـ۲ طبقه بندی می‌گردند .

مقدار متغیر تصادفی (x)	فراوانی مطلق (f_i)
۲۵۰-۲۸۰	۶
۲۸۰-۳۱۰	۶
۳۱۰-۳۴۰	۱
۳۴۰-۳۷۰	۲
۳۷۰-۴۰۰	۷
۴۰۰-۴۳۰	۳
۴۳۰-۴۶۰	۱
۴۶۰-۴۹۰	۵
۴۹۰-۵۲۰	۰
۵۲۰-۵۵۰	۲
۵۵۰-۵۸۰	۴
۵۸۰-۶۱۰	۳
	n = ۴۰

جدول ۱-۲

چنین جداولی را به‌دلیل اینکه چگونگی تقسیم فراوانی مقادیر متغیر تصادفی را در طبقات مختلف نشان می‌دهند جدول توزیع فراوانی می‌نامند . نکته‌ای که در مورد جدول فوق محتاج توضیح است این است که مقصود از عبارت (۲۵۰ــ۲۸۰) و نظایر آن که درسطور مختلف جدول توزیع فراوانی نوشته می‌شود معمولاً" طبق فرض (۲۸۰> x ≥۲۵۰) می‌باشد . یعنی هر طبقه شامل کلیه اعدادی که در عرض آن قرار دارد می‌شود اما حد بالای هر طبقه جزء آن نبوده بلکه به‌عنوان حد پائین طبقه بعد جزء طبقه بعدی محسوب می‌گردد .به‌همین دلیل عدد ۲۸۰ که در اطلاعات فوق قرار دارد جزء طبقه دوم منظور گردید . در تعدادی از کتب آمار بجای نمایش حدود طبقات به‌صورت جدول (۱ــ۲) از روش دیگری استفاده می‌گردد . بدین‌ترتیب که طبقه اول را با اعداد (۲۵۰ــ۲۷۹) و طبقه دوم را با اعداد (۲۸۰ ــ ۳۰۹) نمایش می‌دهند . این روش در مورد متغیرهای گسسته که معمولاً" فقط می‌توانند اعداد صحیح را به‌عنوان مقداراختیارکنند قابل اعمال است زیرا اگر فرض کنیم که مثلاً" در جدول (۱ــ۲) طبقه اول تمام اعداد بین دوعدد (۲۵۰ــ۲۸۰) به‌استثنای خود عدد ۲۸۰ را در بر می‌گیرد درآن صورت می‌توان فاصله (۲۵۰ــ۲۷۹) را نیز برای بیان همین منظور به‌کار برد زیرا در فاصله بین دو عدد ۲۷۹ و ۲۸۰ متغیر تصادفی گسسته مقداری را نمی‌تواند اختیار کند . اما اگر متغیر تصادفی موردمطالعه متغیر پیوسته باشد درآن صورت عبارت (۲۵۰ــ۲۷۹) شامل تمام مقادیری که عبارت (۲۵۰ــ۲۸۰) آنها را در بر می‌گیرد نخواهد شد زیرا در فاصله دو عدد (۲۷۹ــ۲۸۰) یک متغیر تصادفی پیوسته مثل قد و وزن افراد و اشیاء می‌تواندبینهایت مقدار اختیار کند . به‌عبارت دیگر بین دو عبارت مورد بحث رابطه عموم و خصوص مطلق برقرار است و عبارت (۲۵۰ــ۲۸۰) دربرگیرنده تمام مقادیر عبارت دوم نیز هست‌درحالی که عکس‌این جریان صادق نیست و به همین‌دلیل در این جزوه عبارت‌فراگیرترمورداستفاده قرار گرفته است .

معمولاً" جهت سهولت محاسبات چنین عمل می‌کنند که میانگین دو حد بالا و پائین هرطبقه را که طبیعتاً" مقدارآن درست وسط مقادیر دو حد بالا و پائین یک طبقه است ، به‌عنوان علامت طبقه در نظر می‌گیرند به‌عنوان مثال علامت طبقه اول جدول (۱ــ۲) عدد ۲۶۵ می‌باشد و این عمل براین فرض استوار است که مقادیری که در هرطبقه قرار دارند به‌گونه‌ای در فاصله آن طبقه توزیع شده‌اند که میانگین آنها با میانگین دوحد طبقه برابر است . به‌عنوان مثال در جدول (۱ــ۲) معنای فرض مزبور عبارت زیر است :

حاصل‌جمع مقادیر عناصر مربوط به‌طبقه اول = (۲۶۵) (۶)

که معنای عبارت فوق بصورت زیر نیز قابل بیان است .

$$\text{علامت طبقه} = \frac{\text{حاصل‌جمع مقادیر دوحد بالا و پائین}}{۲} = \frac{\text{حاصل جمع مقادیر عناصر موجود در طبقه اول}}{f_i}$$

بدین ترتیب جدول (۱ـ۲) در شکل ابتدائی به‌جدول سه‌ستونی زیر که جهت اختصار فقط سه سطر اول جدول قبلی در آن منعکس شده است تغییر شکل می‌دهد .

F_i	علامت طبقه (X)	مقدار متغیر تصادفی (X)
۶	۲۶۵	۲۵۰ - ۲۸۰
۶	۲۹۵	۲۸۰ - ۳۱۰
۱	۳۲۵	۳۱۰ - ۳۴۰

جـــدول (۲-۲)

معمولاً" در جداول توزیع فراوانی ستونهای دیگری را نیز بسته به‌ضرورت به‌چند ستون فوق اضافه می‌کنند . این ستونهای اضافی بیشتر اوقات مربوط به‌فراوانی مطلق تجمعی یا تراکمی ، فراوانی نسبی و فراوانی نسبی تراکمی هرطبقه می‌باشد . تعاریف مربوط به این سه نوع فراوانی بصورت زیر است .

" فراوانی تراکمی مطلق هرطبقه عبارت از حاصل جمع فراوانیهای مطلق آن طبقه و طبقات ماقبل آن می‌باشد . "

" نسبت فراوانی مطلق هرطبقه به‌کل فراوانیها را فراوانی نسبی آن طبقه گویند "

" نسبت فراوانی تراکمی هرطبقه به‌کل فراوانیها را فراوانی‌تراکمی نسبی آن طبقه گویند "

مثال ۲ ـ ۲ : اطلاعات حاصل از یک نمونه ۲۰ تائی از مصنوعات یک کارخانه که به‌منظور تحقیق در مورد وزن این مصنوعات اخذ شده است برحسب کیلوگرم به‌قرار زیر است .

۲۳/۷۵، ۲۱/۵، ۲۶، ۲۸/۴۲، ۲۱/۱، ۲۴/۲۵، ۲۱/۵۵، ۲۸/۵، ۲۲/۲، ۲۳/۱، ۲۸،
۲۰/۰۵، ۲۶/۱۳، ۲۹/۹۵، ۲۰/۴۸، ۲۷/۶، ۲۲/۲۲، ۲۷/۱۱، ۲۰/۵۵، ۲۵

مطلوبست جدول توزیع فراوانی مطلق ،نسبی، فراوانی تجمعی مطلق و فراوانی تجمعی‌نسبی برای اطلاعات .

پاسخ: همانطور که قبلا" گفته شد حد بالای هر طبقه در سیستم مورد استفاده این جزوه ، جزء آن طبقه محسوب نمی‌گردد . مثل مثال قبل ابتدا ده طبقه را برای این طبقه‌بندی ، در نظر می‌گیریم و عرض هر طبقه را به‌صورت زیر حساب می‌کنیم .

$$\frac{۲۹/۹۵ - ۲۰/۰۵}{۱۰} = \frac{۹/۹}{۱۰} = ۰/۹۹ \qquad \text{عرض هرطبقه}$$

واضح است‌که اگر به‌جای عدد ۲۰/۰۵ از عدد ۲۰ به‌عنوان حد پائین طبقه اول استفاده‌نموده و عرض هرطبقه را نیز به‌جای ۰/۹۹ فعلی‌معادل عدد ۱ درنظربگیریم طبقه‌بندی مناسب‌تری انجام خواهد شد .

X	X'	F_i	فراوانی نسبی F_i'	فراوانی تجمعی مطلق Σf_i	فراوانی تجمعی نسبی $\frac{\Sigma f_i}{n}$
۲۰-۲۱	۲۰/۵	۳	۰/۱۵	۳	۰/۱۵
۲۱-۲۲	۲۱/۵	۳	۰/۱۵	۶	۰/۳
۲۲-۲۳	۲۲/۵	۲	۰/۱	۸	۰/۴
۲۳-۲۴	۲۳/۵	۲	۰/۱	۱۰	۰/۵
۲۴-۲۵	۲۴/۵	۱	۰/۰۵	۱۱	۰/۵۵
۲۵-۲۶	۲۵/۵	۱	۰/۰۵	۱۲	۰/۶
۲۶-۲۷	۲۶/۵	۲	۰/۱	۱۴	۰/۷
۲۷-۲۸	۲۷/۵	۲	۰/۱	۱۶	۰/۸
۲۸-۲۹	۲۸/۵	۳	۰/۱۵	۱۹	۰/۹۵
۲۹-۳۰	۲۹/۵	۱	۰/۰۵	۲۰	۱
		۲۰	۱		

جدول ۳-۲

یکی از موارد عمده استفاده از فراوانی نسبی و تجمعی نسبی در محاسبه احتمالات است بدین ترتیب‌که فراوانی نسبی مربوط به‌هرطبقه احتمال انتخاب هریک از عناصرموجود درآن طبقه به‌طورتصادفی می‌باشد و به‌همین ترتیب فراوانی تجمعی نسبی مربوط به‌هرطبقه احتمال‌انتخاب تصادفی هریک از عناصر موجود درآن طبقه و طبقات قبلی رانشان‌می‌دهد .

مثال ۳ـ۲ : اگر از ۲۰ کالای موجود درنمونه‌مربوط به‌مثال ۲ـ۲ یکی‌را به‌طورتصادفی‌انتخاب کنیم :

الف : احتمال آنکه مقدار آن ($۲۵ > x \geq ۲۴$) باشد چقدر است؟
ب : احتمال آنکه مقدار آن ($۲۵ > x$) باشد چقدر است؟

پاسخ: الف : این مقدار برابر فراوانی نسبی طبقه پنجم جدول (۳-۲) می‌باشد .

$$(۲۴ \leq x < ۲۵) = ۰/۰۵ = ۵\ \%$$

ب : این مقدار برابر فراوانی نسبی تجمعی طبقه پنجم جدول (۳-۲) می‌باشد .

$$(x < ۲۵) = ۰/۵۵ = ۵۵\ \%$$

ب : نمایش نموداری اطلاعات طبقه‌بندی شده :

اطلاعات طبقه‌بندی شده در جداول توزیع فراوانی را می‌توان به‌صورت نمودار نیز نشان داد . متداولترین اشکال نمایش جدول توزیع فراوانی نمودارهای میله‌ای و هیستوگرام می‌باشند .

ب - ۱ - نمودار میله‌ای :

این نوع نمودارمعمولا" برای نمایش جداول‌توزیع فراوانی به‌کارمی‌رود که‌متغیرتصادفی آنها گسسته بوده و مقادیر مشاهده شده برای این متغیر نیز زیاد متنوع نباشند . به‌عنوان مثال اگر متغیر تصادفی ما تعداد حالت‌های مشاهده شده درصد بار پرتاب یک تاس همتراز باشداولا" این متغیر ،متغیری‌گسسته است و ثانیا"این‌متغیر فقط می‌تواند ۶ مقدار را اختیار کند و مقادیرآن خیلی‌متنوع‌نیستند و درنتیجه نمایش توزیع‌فراوانی چنین آزمایشی‌ازطریق نمودار میله‌ای مناسب‌تر است . البته معنای جمله مزبور این نیست که سایر جداول توزیع فراوانی را نمی‌توان از طریق نمودار میله‌ای نشان داد بلکه مقصود آن است که نمودارهای هیستوگرام به‌دلیل آنکه‌اطلاعات بیشتری را نسبت به‌نمودارهای‌میله‌ای منعکس می‌کنند برای نمایش جداول پیچیده‌تر مناسبت بیشتری دارند .

برای رسم نمودار میله‌ای کافی است مقادیر متغیر تصادفی را برروی محور افقی و فراوانی مطلق یا نسبی و یا هر فراوانی دلخواه دیگری را برروی محور عمودی نشان داده و سپس تصویرفراوانی هرمقدار را به‌صورت‌یک خط برروی نقطه‌ای که در محور افقی ،آن‌مقدار خاص را نشان می‌دهد عمود کنیم .

مثال ۴ - ۲ : جدول (۴-۲) نتیجه مطالعه برروی تعداد فرزندان یک نمونه ۵۰ خانواری را

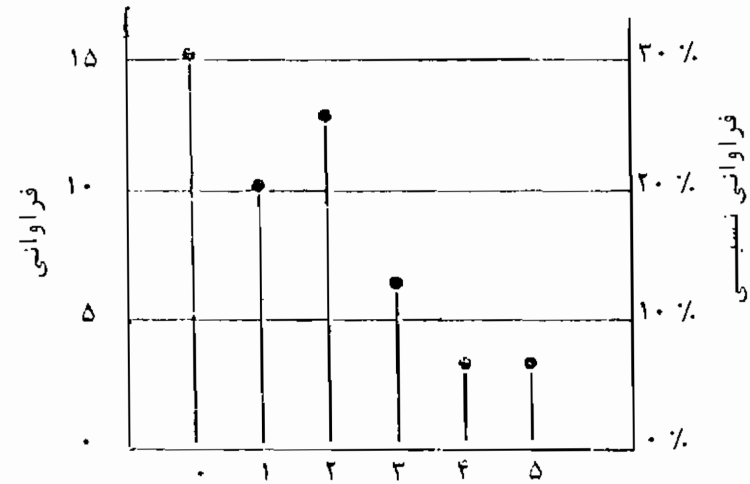
نشان می‌دهد. نمودار میله‌ای مربوط به این جدول را برحسب فراوانی مطلق و نسبی رسم کنید.

(۱) تعداد فرزندان	(۲) خط نشان	(۳) فراوانی (f)	(۴) فراوانی نسبی (f/n)
۰	///// ///// /////	۱۵	۰/۳۰
۱	///// /////	۱۰	۰/۲۰
۲	///// ///// ///	۱۳	۰/۲۶
۳	///// /	۶	۰/۱۲
۴	///	۳	۰/۰۶
۵	///	۳	۰/۰۶

$\Sigma\left(\frac{f}{n}\right)=۱/۰۰ \quad \Sigma f=۵۰=n$

جدول ۴-۲

پاسخ: در محور عمودی سمت چپ فراوانی مطلق و در محور عمودی سمت راست فراوانی نسبی را نشان می‌دهیم.



شکل ۱-۲

ب: ۲ ــ نمودار هیستوگرام

مورد استفاده این نمودار جداول توزیع فراوانی پیچیده مربوط به متغیرهای گسسته و تمام جداول مربوط به متغیرهای پیوسته می‌باشد. به‌عبارت دیگر تمام جداول توزیع فراوانی

که‌دارای طبقات غیرنقطه‌ای و عریض هستند توسط نمودارهای‌هیستوگرام‌نشان داده‌می‌شوند. خاصیت اصلی این نمودارها که در واقع مزیت آنها بر نمودارهای میله‌ای محسوب می‌گردد این است که در این نمودارها می‌توان حدود هرطبقه و نیز وسط هرطبقه را نمایش داد.

برای ترسیم این نمودارها باید ابتدا حدود هرطبقه را برروی محور افقی یک دستگاه مختصات مشخص نموده و فراوانی موردنظر (مطلق، نسبی، یا تراکمی) را روی محور عمودی نشان داد. سپس تصویر فراوانی مربوط به هرطبقه را به‌صورت یک مستطیل برروی فاصله‌ای که در محور افقی، نمایش دهنده آن طبقه است عمود کرد.

مثال ۵ ـ ۲: جدول (۵ ـ ۲) وضعیت قد ۲۰۰ مرد موجود در یک نمونه را نشان می‌دهد. مطلوبست نمودار هیستوگرام مربوط به‌این جدول براساس فراوانیهای مطلق و نسبی.

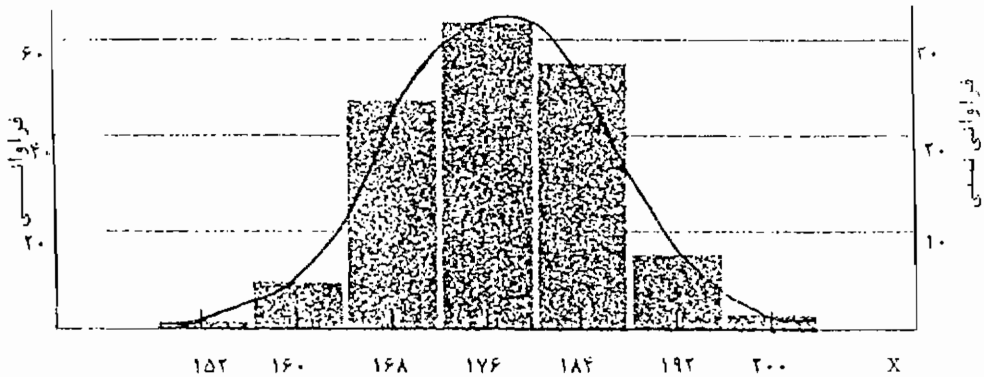
شماره طبقه	حدودطبقه	علامت طبقه X'	خط نشان	فراوانی f	فراوانی نسبی $\frac{f}{n}$
۱	۱۴۸-۱۵۶	۱۵۲	||	۲	۰/۰۱
۲	۱۵۶-۱۶۴	۱۶۰	𝍸 𝍸	۱۰	۰/۰۵
.	.	۱۶۸	.	۴۸	۰/۲۴
.	.	۱۷۶	.	۶۴	۰/۳۲
.	.	۱۸۴	.	۵۶	۰/۲۸
.	.	۱۹۲	.	۱۶	۰/۰۸
۷	۱۹۶-۲۰۴	۲۰۰	|||	۴	۰/۰۲

$\Sigma F_i =۲۰۰ =n$ $\qquad$ $\frac{\Sigma F_i}{n} =۱$

جـــــــدول ۵-۲

پاسخ: همانند مثال قبل فراوانی مطلق را در محور عمودی سمت چپ و فراوانی نسبی را روی محور سمت راست نشان می‌دهیم. ضمناً" ازآنجا که در بین آمارگران معمول‌است که با اتصال وسط اضلاع بالائی هریک ازمستطیلهای یک نمودار هیستوگرام به‌یکدیگر نمودارتوزیع فراوانی را به‌صورت منحنی به‌دست آورند در شکل (۲ـ۲) منحنی توزیع فراوانی مربوط به‌این مثال‌نیزرسم شده است. واضح است که هریک از این نقاط نشان‌دهنده‌فراوانی مطلق مربوط به‌وسط هرطبقه و به‌عبارت بهتر مربوط به‌هرطبقه می‌باشند.

حال که از منحنی توزیع فراوانی صحبت شد بهتر است که اشکال مختلف این منحنی نیز به‌طور مختصر توضیح داده شود.

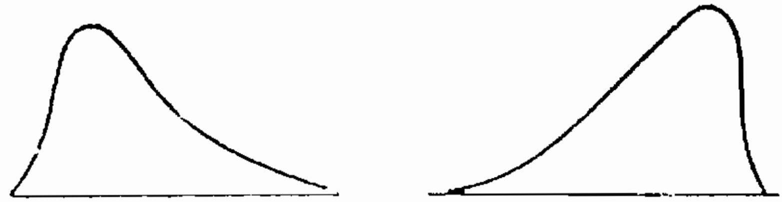


شکل ۲ -۲

ب ــ ۳ــ اشکال مختلف منحنی‌های توزیع فراوانی:

بطور کلی منحنی‌های توزیع به‌دوشکل متقارن و نامتقارن (چوله) وجود دارند .
" هرگاه مقدار وسطی یک متغیر تصادفی بیشترین فراوانی را داشته و وقتی که بطرف "
" دو حد بالا و پائین مقادیر حرکت می‌کنیم از مقدار فراوانی دوطرف به‌صورت نسبتا"
" هماهنگ و متقارنی کاسته گردد توزیع فراوانی این متغیر را متقارن و یا زنگی شکل "
" و در غیر اینصورت توزیع را نامتقارن یا چوله می‌نامند .

شکل (۲ــ۲) یک منحنی توزیع متقارن یا زنگی شکل و دو منحنی موجود در شکل (۳ ــ ۲) منحنی‌های نامتقارن یا چوله می‌باشند . چون در منحنی الف بیشترین فراوانیها در طرف چپ منحنی متمرکز است آن را منحنی چوله به‌راست و به‌دلیلی مشابه منحنی ب را چوله به‌چپ می‌نامند .



الف ب

شکل ۳-۲

ج : محاسبه مشخصه‌های تمرکز:

همانطور که قبلا" گفته شد قضاوت در مورد وضعیت یک متغیر وقتی می‌تواند به‌خوبی انجام شود که در مورد کیفیت تمرکز و پراکندگی آن متغیر اطلاعات کافی در اختیار باشد .

اگر آمارگری بخواهد در مورد وضعیت یک متغیر در نمونه قضاوت کند علی‌الاصول به چیزی بیش از اطلاعات طبقه‌بندی شده نیاز نخواهد داشت و همین میزان پرورش اطلاعات برای تأمین هدف او کفایت می‌کند زیرا از روی یک جدول توزیع فراوانی و یا یک هیستوگرام فراوانی به‌راحتی می‌توان در مورد وضعیت تمرکز و پراکندگی متغیر مورد نظر قضاوت کرد. اما معمولاً" هدف آمارگران تحقیق در مورد نمونه نیست بلکه آنها از نمونه به‌عنوان مبنایی برای قضاوت در جامعه استفاده می‌نمایند و می‌خواهند با توسل به نمونه درمورد وضعیت تمرکز و پراکندگی یک متغیر در جامعه قضاوت نمایند و همانطور که قبلاً" نیز به‌دفعات گفته شده است چنین هدفی وقتی قابل‌تأمین است که در مورد مشخصه‌های تمرکز و پراکندگی جامعه نتایجی هرچند به‌صورت تقریبی بدست آورد.

تنها وسیله‌ای که می‌تواند به یک آمارگر درمورد بدست آوردن تقریبی مقدار مشخصه‌های جامعه کمک کند مقدار همین مشخصه‌ها در نمونه می‌باشد و به‌همین دلیل دانستن چگونگی محاسبه مشخصه‌های تمرکز و پراکندگی درنمونه درراستای تحقیق درمورد یک متغیر در جامعه (نه نمونه) امری لازم و اجتناب‌ناپذیر می‌نماید که محاسبه مشخصه تمرکز یک نمونه موضوع این قسمت از بحث می‌باشد.

بطورکلی سه مشخصه اصلی برای تمرکز وجود دارند این سه مشخصه عبارتند از میانگین(۱) و میانه(۲) و بالاخره نما یا مد(۳).

ج ـ ۱ ـ میانگین:

در بین سه مشخصه نامبرده میانگین به‌دلیل اطلاعات مفیدتری که در بیشتر موارد در اختیار محقق قرار می‌دهد مهمترین آنها می‌باشد و به‌همین‌دلیل نیز بیش از دیگر مشخصه‌های تمرکز مورد استفاده آمارگران قرار می‌گیرد. از این مشخصه گاهی به‌عنوان میانگین حسابی نیز یاد می‌شود وخصوصیات آن نیز برای افرادی که حتی معلومات ریاضی بسیار کم هم دارند مشخص است و لذا در اینجا از بحث پیرامون آن خودداری می‌گردد. علامت اختصاری این مشخصه در کتب آماری ($\bar{x}$) می‌باشد و طرز محاسبه آن نیز بسته به اینکه اطلاعات، طبقه‌بندی شده باشند یا نباشند، کمی متفاوت است. لذا در اینجا ابتدا روش محاسبه میانگین برای اطلاعات طبقه‌بندی نشده و سپس برای اطلاعات طبقه‌بندی شده توضیح داده می‌شود.

---

1 - Mean

2 - Median

3 - Mode

محاسبه میانگین برای مقادیر طبقه‌بندی نشده:

میانگین حسابی در چنین مواردی به‌سادگی قابل محاسبه است . فرض کنید متغیری مثل (X) n مقدار ($x_1, x_2, \ldots\ldots, x_n$) را اختیار کند . برای محاسبه میانگین در این موارد از رابطه زیر استفاده می‌شود .

$$\bar{x} = \frac{x_1 + x_2 + \cdots\cdots\cdots + x_n}{n} \qquad (۲-۱)$$

مثال ۶ ــ ۲ : وضعیت درآمد ماهیانه افراد در یک نمونه ۵ تایی برحسب ریال بصورت زیر است . میانگین درآمد را برای این نمونه محاسبه کنید .

(۵۰٫۰۰۰ ، ۶۵٫۰۰۰ ، ۴۰٫۰۰۰ ، ۵۰٫۰۰۰ ، ۱۰۰٫۰۰۰)

پاسخ: از رابطه (۱ــ۲) استفاده می‌کنیم .

$$\bar{x} = \frac{x_1 + x_2 + \cdots\cdots + x_5}{n} = \frac{50000+65000+40000+50000+100000}{5}$$

$$\bar{x} = \frac{305000}{5} = 61000$$

قبل از آنکه به‌توضیح روش محاسبه میانگین برای اطلاعات طبقه‌بندی شده بپردازیم لازمست مختصرا" بعضی قواعد مربوط به‌علامت مجموع که با حرف یونانی Σ (سیگمای‌بزرگ) نشان داده می‌شود توضیح داده شوند .

علامت مجموع :

فرض کنید متغیر X مقادیر ($x_1, x_2, \ldots\ldots, x_n$) را اختیار کند . در صورتی که بخواهیم این مقادیر را با یکدیگر جمع کنیم مسلما" باید آنها را به‌صورت زیر بنویسیم ؛

$$x_1 + x_2 + \cdots\cdots\cdots + x_n$$

به‌منظوررعایت اختصار در عملیات ریاضی وآمار به‌جای جمله فوق از علامت مجموع‌بامعنای مفروض زیر استفاده می‌کنند .

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots\cdots\cdots + x_n \qquad (۲-۲)$$

معنای عبارت سمت راست مجموع مقادیر x_i است در صورتی که i از ۱ تا n تغییر کند .

مثال ۷ ــ ۲ : ضمن استفاده از علامت مجموع مثال (۶ــ۲) را حل کنید .

پاسخ : در صورت استفاده از علامت مجموع رابطه (۱ــ۲) یعنی رابطه مربوط به محاسبه میانگین به‌صورت زیر تغییر می‌یابد .

(۲ ــ ۳)

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

تنها تفاوت پاسخهای دومثال فوق در همین‌است یعنی در مثال (۷ــ۲) باید به‌جای استفاده رابطه (۱ــ۲) از رابطه (۳ــ۲) استفاده شود .

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{۵۰۰۰۰+۶۵۰۰۰+۴۰۰۰۰+۵۰۰۰۰+۱۰۰۰۰۰}{۵} = ۶۱۰۰۰$$

قضایای زیر در مورد علامت مجموع صادق است .

قضیه : ۱ــ۲ .

(۲ ــ ۴)

$$\sum_{i=1}^{n}(x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

قضیه : ۲ــ۲ . اگر Y مقدار ثابتی داشته باشد داریم

(۲ ــ ۵)

$$\sum_{i=1}^{n} Y_i = nY$$

قضیه : ۳ــ۲ . اگر Y مقدارثابتی داشته باشد بااستفاده از قضیه (۲ــ۲) قضیه (۱ــ۲) به‌صورت زیر تغییرشکل می‌دهد :

(۲ ــ ۶)

$$\sum_{i=1}^{n}(x_i+Y_i) = \sum_{i=1}^{n} x_i + nY$$

قضیه : ۴ــ۲ . اگر C مقدار ثابتی داشته باشد داریم :

(۲ ــ ۷)

$$\sum_{i=1}^{n} Cx_i = C\sum_{i=1}^{n} x_i$$

درمواردی که ازعلامت Σx_i یا Σy_i استفاده‌می‌شود مقصود حاصل جمع همه n عنصرموجود است .

مثال ۸ ــ ۲ : در یک نمونه ۴ تایی مقادیر متغیر تصادفی X به‌ترتیب از چپ به‌راست (۶ ،

۴،۷،۵) و مقادیر متغیر با همان ترتیب (۲،۵،۸،۴) می‌باشد:

الف: مقدار $\sum_{i=1}^{4}(x_i + y_i)$ را محاسبه کنید.

ب: اگر Y در هرچهار مورد مقدارثابت (۴ = Y) را اختیار کند درآن صورت مقادیرعبارات ΣY_i و $\Sigma(x_i + Y_i)$ و $\Sigma Y_i x_i$ را محاسبه کنید.

پاسـخ: الف: از رابطه ۴ـ۲ استفاده می‌کنیم.

$$\Sigma(x_i + y_i) = \Sigma x_i + \Sigma y_i$$

اما داریم:

$$\Sigma x_i = ۴+۷+۵+۶ = ۲۲ \qquad و \qquad \Sigma y_i = ۲+۵+۸+۴ = ۱۹$$

در نتیجه مقدار خواسته شده برابر مقدار زیر است.

$$\Sigma(x_i + y_i) = ۲۲+۱۹ = ۴۱$$

ب: چون نمونه ۴ تایی است بنابراین n = ۴ بوده و به‌ترتیب از روابط(۵ـ۲) و (۶ـ۲) و (۷ـ۲) استفاده می‌کنیم.

$$\Sigma Y_i = nY = ۴(۴) = ۱۶$$

$$\Sigma(x_i + Y_i) = \Sigma x_i + nY = ۲۲+۱۶ = ۳۸$$

$$\Sigma Y_i x_i = Y \Sigma x_i = ۴(۲۲) = ۸۸$$

محاسبه میانگین برای مقادیر طبقه‌بندی شده:

وقتی که اطلاعات طبقه‌بندی شده باشند برای محاسبه میانگین باید از رابطه (۸ـ۲) یعنی رابطه زیر استفاده نمود.

(۲ ـ ۸)

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{h} x_i F_i$$

در این رابطه n نشان دهنده حجم نمونه h معرف تعداد طبقات، x_i بیانگر علامت هر طبقه بوده و بالاخره F_i فراوانی مطلق هرطبقه را نشان می‌دهد. البته علامت طبقات

تا به‌حال با (X') نشان داده می‌شد که از این پس جهت رعایت هماهنگی آنرا در جداول توزیع فراوانی با X نشان خواهیم داد.

رابطه (۲ـ۸) در واقع همان رابطه (۲ـ۳) است که به‌اقتضای طبقه‌بندی اطلاعات تغییرات مختصری پیدا کرده است. برای توضیح بیشتر مجدداً" به‌مثال (۲ـ۶) توجه کنید. مقدار میانگین در این مثال به‌صورت زیر قابل محاسبه است:

$$\bar{x}= \frac{(۵۰۰۰۰)(۲)+(۶۵۰۰۰)(۱)+(۴۰۰۰۰)(۱)+(۱۰۰۰۰۰)(۱)}{۵} = ۶۱۰۰۰$$

اعداد ۲ و ۱ که در صورت کسر در مقادیر متغیر ضرب شده‌اند در واقع همان فراوانی مطلق آن مقادیر می‌باشند که چون مقدار این فراوانیها اکثراً" معادل ۱ بوده است قبلاً" از نشان‌دادن آنها صرف نظر شده و عدد (۵۰۰۰۰) نیز به‌جای آنکه در دو ضرب شود، دوبار تکرارشده است. بدین‌ترتیب در رابطه (۲ـ۳) نیز هرمقدار، در فراوانی خود ضرب می‌شده است. از طرفی چون در اطلاعات طبقه‌بندی شده معمولاً" مقدار اصلی مشاهده شده برای متغیر تنوع زیادی داشته و به‌همین دلیل نیز آنها را طبقه‌بندی می‌کنند استفاده از مقادیر اصلی ممکن نیست. مثلاً" در مثال (۲ـ۲) وزن ۲۰ محصول انتخاب شده ۲۰ مقدار مختلف دارد. واضح است که در چنین مواردی استفاده از علامت هر طبقه به‌عنوان مقدار X درهر طبقه مناسب‌تر از هر مقدار دیگری می‌باشد. رابطه (۲ـ۸) در واقع چیزی نیست جز همان رابطه (۲ـ۳) که در آن دو نکته فوق ملحوظ گشته است. یعنی فراوانی مطلق جهت کوتاه‌کردن جملات صورت، در رابطه ظاهر شده و به‌جای مقدار اصلی تک تک مشاهدات، علامت هر طبقه گذاشته شده است.

مثال ۲ ـ ۸. میانگین قد افراد را در مثال (۲ـ۵) حساب کنید.

پاسخ: از رابطه (۲ـ۸) استفاده می‌کنیم.

$$\bar{x}=\frac{1}{n}\sum x_i P_i = \frac{۱}{۲۰۰}\ (۱۵۲)(۲)+(۱۶۰)(۱۰)+(۱۶۸)(۴۸)+(۱۷۶)(۶۴)+(۱۸۴)(۵۶)+(۱۹۲)(۱۶)+(۲۰۰)(۴) = \frac{۳۵۴۰۸}{۲۰۰} = ۱۷۷$$

ج ـ ۲ـ میانه

دومین مشخصه تمرکز مهم آماری میانه می‌باشد.

"اگر مقادیر یک متغیر به‌ترتیب اندازه مرتب شده باشند عددی که در وسط"
"آنها قرار دارد را میانه نامند."

واضح است که اگر تعداد مشاهدات و یا تعداد مقادیر متغیر فرد باشد یک عدد در وسط قرار می‌گیرد و همان عددمیانه است اما اگر تعدادمشاهدات زوج باشد دو عدددر وسط قرار می‌گیرند .در چنین مواردی میانگین دوعدد وسطی به‌عنوان میانه در نظر گرفته‌می‌شود روش‌محاسبه میانه برای اطلاعات طبقه‌بندی نشده و طبقه‌بندی شده متفاوت است که ذیلاً هر دو روش توضیح داده می‌شوند .

محاسبه میانه برای مقادیر طبقه‌بندی نشده

روش کار در این مواردبسیار ساده است . بدین‌ترتیب که ابتدا مقادیر متغیررابرحسب اندازه مرتب می‌کنند . اگر تعداد مشاهدات فرد باشد عدد وسطی همان میانه است و اگر تعداد آنها زوج باشد میانگین دو عدد وسطی میانه خواهد بود .

مثال ۹ ــ ۲ : مقادیر مشاهده شده X در یک نمونه ۵ تایی به‌قرار زیر است . میانه را برای این مقادیر مشخص کنید .

(۸ ، ۶ ، ۴ ، ۱۲ ، ۵)

پاسخ : مقادیر X را برحسب اندازه مرتب می‌کنیم

(۴ ، ۵ ، ۶ ، ۸ ، ۱۲)

چون تعداد مشاهدات فرد می‌باشد عدد میانه ۶ است .

مثال ۱۰ ــ ۲ : در یک نمونه ۱۰ تایی مقادیر مشاهده شده X به‌صورت زیر است مقدار میانه را برای مقادیر مشخص کنید .

(۲ ، ۱۱ ، ۸ ، ۴ ، ۳ ، ۸ ، ۴ ، ۸ ، ۵ ، ۷)

پاسخ : مقادیر فوق را برحسب اندازه مرتب می‌کنیم .

(۲ ، ۳ ، ۴ ، ۴ ، ۵ ، ۷ ، ۸ ، ۸ ، ۸ ، ۱۱)

چون تعداد مشاهدات زوج است میانگین دو مقدار وسطی یعنی ۵ و ۷ میانه خواهد بود .

$$Me = \frac{۵+۷}{۲} = ۶$$

محاسبه میانه برای مقادیر طبقه‌بندی شده :

وقتی که اطلاعات طبقه‌بندی شده باشند پیداکردن مقدار میانه کمی‌پیچیده‌تراز حالت

قبل است . روش محاسبه میانه را در این حالت در قالب یک مثال توضیح می‌دهیم .

مثال ۱۱ ـ ۲ . مشاهدات حاصل از یک نمونه ۴۱ تایی برای متغیر X مطابق جدول توزیع فراوانی زیر است . میانه این مشاهدات را محاسبه کنید .

حدود طبقه	X_i علامت طبقه	F_i	ΣF_i
۱۰-۲۰	۱۵	۱	۱
۲۰-۳۰	۲۵	۲	۳
۳۰-۴۰	۳۵	۵	۸
۴۰-۵۰	۴۵	۶	۱۴
۵۰-۶۰	۵۵	۱۰	۲۴
۶۰-۷۰	۶۵	۴	۲۸
۷۰-۸۰	۷۵	۶	۳۴
۸۰-۹۰	۸۵	۲	۳۶
۹۰-۱۰۰	۹۵	۳	۳۹
۱۰۰-۱۱۰	۱۰۵	۲	۴۱

جدول ۶-۲

پاسخ : برای محاسبه میانه‌ابتدا باید دید چندمین مشاهده در وسط قرار می‌گیرد اگرتعداد مشاهدات‌زوج‌باشد این‌رقم معادل $\frac{n}{۲}$ و اگر فرد باشد معادل $\frac{n+۱}{۲}$ درنظرگرفته‌می‌شود. چون در مثال فوق ($n = ۴۱$) بوده و فرد می‌باشد پس مشاهده وسطی مشاهده زیر خواهد بود .

$$\frac{n+۱}{۲} = \frac{۴۱+۱}{۲} = ۲۱$$

در مرحله بعدی باید دید مشاهده (۲۱) در کدام طبقه قرار می‌گیرد . ستون فراوانی تجمعی مشخص می‌کند که این طبقه ، طبقه پنجم است زیرا مشاهدات پانزدهم تا بیست و چهارم در این طبقه قرار گرفته‌اند . در مورد مقادیر مشاهدات در این طبقه فقط می‌دانیم که مقدار آنها ($۶۰ > X > ۵۰$) بوده است و رقم دقیق هریک از ۱۰ مشاهده موجود در این طبقه ، درجدول فوق مشخص نشده است . از آنجا که هفتمین مشاهده موجود این‌طبقه میانه می‌باشد (زیرا بیست و یکمین مشاهده در کل مشاهدات است) اگر مقادیر دقیق هریک از مشاهدات معلوم بود مقدار میانه براحتی مشخص می‌گردید . اما در شرایط موجود باید مقدار هفتمین مشاهده‌طبقه پنجم یعنی بیست و یکمین مشاهده در کل مشاهدات را به‌طریقی

تخمین زد .

برای این‌کار فرض می‌شودکه در هریک از طبقات جدول‌توزیع فراوانی مقادیرمشاهدات بطورمساوی توزیع شده‌اند .به‌عنوان مثال طبق این‌فرض طبقه سوم به ۵ قسمت مساوی تقسیم شده و در هریک از این قسمتها یک مقدار قرار می‌گیرد یعنی مقادیر دقیق مشاهدات در طبقه سوم بر اساس این فرض به‌صورت زیر است .

(۳۱ ، ۳۳ ، ۳۵ ، ۳۷ ، ۳۹)

حال براساس این فرض مقدار میانه را می‌توان تعیین نمود زیرا می‌توان مقدار میانه را معادل جمع مقادیر دو جزء رابطه زیر دانست که اینک پیدا کردن مقدار این دوجزء نیز به‌راحتی امکان‌پذیر است

(مقداربیست و یکمین‌مشاهده) = Me = مقدارمشاهده چهاردهم + (عرض طبقه پنجم) $\frac{۷}{۱۰}$)

صحت رابطه فوق به راحتی قابل درک است . چون در جدول توزیع فراوانی مشاهدات به‌ترتیب اندازه و از کوچک به‌بزرگ مرتب شده‌اند بنابراین مقدارمشاهده بیست و یکم قطعاً" ازمقدار مشاهدات قبلی بزرگتر است . اینک بر اساس فرض مزبورمی‌توان مقدار مشاهده چهاردهم یعنی جزء اول رابطه فوق را به‌راحتی مشخص نمودزیرا این مشاهده آخرین عنصر طبقه چهارم بوده و طبق فرض فوق‌الذکر بزرگترین‌مقدار را در طبقه دار است. معمولا" در آمارجهت سهولت ، مقدارآخرین عنصر هرطبقه‌را برابر مقدار حد بالای آن طبقه و یا حد پائین طبقه بعد در نظر می‌گیرند که در این مورد خاص این مقدار برابر حدبالای طبقه چهارم یعنی عدد ۵۰ می‌باشد . برای بدست‌آوردن جزء دوم فرض کنید که مشاهده بیست و چهارم مشاهده میانه باشد که مقدار آن به‌اندازه عرض طبقه پنجم از مقدارمشاهده چهاردهم بیشتر است . چون مشاهده بیست و چهارم آخرین عنصر طبقه پنجم است .

۶۰=۱۰+۵۰ = عرض طبقه پنجم + مقدار مشاهده چهاردهم = مقدار مشاهده بیست و چهارم

اما مشاهده بیست و یکم و به‌عبارت دیگر مشاهده هفتم طبقه پنجم ، میانه بوده و مطلوب مسئله است .برای آنکه بدانیم مقدار این مشاهده چقدر از مقدار مشاهده چهاردهم بزرگتر است کافی است تناسب زیر را ببندیم .

مقدار افزایش	تعداد مشاهدات بعداز مشاهده چهاردهم
۱۰	۱۰
$x = \frac{۲۱-۱۴}{۱۰}$ ۱۰	۷ =۲۱-۱۴

مقدار X برابر جزء دوم رابطه فوق است . از جایگذاری مقادیر هریک از دو جزء مزبور مقدار میانه بدست می‌آید .

$$Me = ۵۰ + \frac{۲۱ - ۱۴}{۱۰} \times ۱۰ = ۵۷$$

در رابطه فوق عدد ۵۰ حد پائین طبقه میانه است و معمولاً" آن را با حرف (L)نشان می‌دهند عدد ۲۱ شماره مشاهده وسطی است که اگر مجموع مشاهدات زوج باشد با $\frac{n}{۲}$ و اگر فرد باشد با $\frac{n+۱}{۲}$ نشان داده می‌شود . عدد ۱۴ فراوانی تجمعی طبقه قبل از طبقه میانه است و معمولاً" با حرف (F)نمایش داده‌می‌شود عدد ۱۰ موجود در مخرج‌کسردر واقع فراوانی مطلق طبقه میانه است و با F_m نشان داده می‌شود و بالاخره عدد ۱۰ بیرون ازپرانتزعرض طبقه میانه است که معمولاً" با حرف(C)نشان داده می‌شود . واضح است که رابطه فوق را می‌توان در تمام موارد ، مورد استفاده قرار داد و بدین ترتیب رابطه کلی بدست‌آوردن میانه برای مقادیر طبقه‌بندی شده رابطه زیر می‌باشد .

$$Me = L + \left(\frac{\frac{n}{۲} - F}{F_m} \right) C \qquad (۲-۹)$$

مثال ۱۲ ــ ۲ : مقدار میانه را در مثال (۵ ــ ۲) تعیین کنید .

پاسخ : چون تعداد عناصر نمونه زوج است عیناً" از رابطه (۹ــ۲) استفاده می‌کنیم . چون $\frac{n}{۲} = ۱۰۰$ است صدمین عنصر میانه خواهد بود . به‌راحتی می‌توان دریافت که طبقه چهارم حاوی صدمین عنصربوده و طبقه‌میانه است و فراوانی تجمعی طبقه قبلی نیز $F = ۶۰$ می‌باشد. مقدار میانه براساس این اطلاعات مقدار زیر است .

$$Me = ۱۷۲ + \left(\frac{۱۰۰ - ۶۰}{۶۴} \right) ۸ = ۱۷۷$$

مثال ۱۳ ــ ۲ : میانه را برای مثال (۲ــ۲) که توزیع فراوانی آن در جدول (۳ــ۲) منعکس شده است پیدا کنید در صورتی که فراوانی مطلق طبقه آخر این جدول ۱۰ باشد .

پاسخ : اگر فراوانی مطلق طبقه آخر $F_{۱۰} = ۱۰$ باشد ، $n = ۲۹$ خواهد شد که عددی فرد است و باید برای محاسبه میانه آن ، از رابطه زیر استفاده نمود .

$$Me = L + \left(\frac{\frac{n+۱}{۲} - F}{F_m} \right) C$$

$۱۵ = \frac{n+۱}{۲}$ بوده و عنصر پانزدهم میانه است که در طبقه هشتم قرار دارد .

$$Me = ۲۷ + \left(\frac{۱۵ - ۱۴}{۲} \right) ۱ = ۲۷/۵$$

ج : ۳ ـ نمایامد (Mo) :

مد یا نما سومین مشخصه مهم تمرکز محسوب می‌گردد .

" نمای یک دسته مشاهدات مقداری است که بیشترین فراوانی را داشته باشد . "

ممکن است در بعضی موارد نما وجود نداشته باشد و یا بیش از یک نما موجود باشد "

واژه مد در اینجا همان مفهومی‌است که در لسان عامه مردم به‌کار می‌رود . وقتی که می‌گویند فلان کالا مد شده‌است یعنی مصرف آن زیاد شده‌است . در اینجا نیز عنصر مد عنصری‌است که بیشترین فراوانی را دارد .

مثال ۱۴ ــ ۲ : در مشاهدات زیر نما را مشخص کنید .

الف : ۲ ، ۳، ۳ ، ۳ ، ۴ ، ۶ ، ۷

ب : ۲ ، ۳ ، ۴ ، ۵ ، ۶ ، ۷

ج : ۲ ، ۳ ، ۳ ، ۳ ، ۴ ، ۵ ، ۵ ، ۵ ، ۶

پاسخ : الف ــ ۳   ب ــ نما ندارد   ج ــ ۳ و ۵

در داده‌های طبقه‌بندی شده معمولا" فراوانی مربوط به هرطبقه وجود دارد و نه فراوانی هر مشاهده و لذا مشخص‌کردن نما مشکل‌است . در چنین مواردی معمولا" مرکز طبقه (علامت طبقه) ای که بیشترین فراوانی را دارد به‌عنوان نما برآورد می‌شود .

مثال ۱۵ ــ ۲ : در مثال (۵ ــ ۲) نما را بدست آورید .

پاسخ : طبقه چهارم نما بوده و عدد ۱۷۶ به‌عنوان نما برآورد می‌شود .

د : مشخصه‌های پراکندگی :

قبلا" دانستیم که یکی از ویژگیها و صفاتی که لازم است در جریان یک مطالعه آماری از سوی آمارگران در مورد متغیر مورد مطالعه آنها مورد بررسی قرار گیرد پراکندگی آن متغیر حول مرکز توزیع می‌باشد .اگر جدول توزیع فراوانی جامعه اصلی در اختیار محقق قرارداشت برای ارزیابی پراکندگی او به‌هیچ مشخصه‌ای احتیاج نداشت زیرا که جدول و منحنی توزیع به‌خوبی این مسئله را توضیح می‌دهند . اما محقق فقط جدول توزیع فراوانی یک یا چند

نمونه را در اختیار دارد و باید بر مبنای اطلاعات این نمونه و یا نمونه‌ها در مورد جامعه به‌نتیجه‌ای دست یابد . در چنین وضعیتی او باید بر اساس اطلاعات نمونه به‌مشخصه‌هائی از جامعه دست یابد که او را در ارزیابی پراکندگی توزیع یاری نمایند . چنانکه قبلا" گفته‌شد این کار فقط وقتی میسر است که او مقادیر مشخصه‌های پراکندگی نمونه را در اختیار داشته باشد و بنابراین در جریان یک کار آماری محاسبه مشخصه‌های پراکندگی نمونه امری لازم و ضروریست . اما سئوال این است که چه مشخصه‌هایی می‌توانند وضعیت پراکندگی یک متغیر را در یک نمونه و یا جامعه مشخص کنند . کوششهای فراوانی به‌منظور وضع چنین مشخصه‌هایی تا کنون صورت‌گرفته است‌که ثمره آنها پیدایش سه‌مشخصه به‌نامهای دامنه تغییرات، میانگین انحرافات و انحراف معیار، می‌باشد . در بین این سه مشخصه ، سومی در آمار از اهمیت بسیار زیاد برخوردار بوده و کاربرد بسیار وسیعی دارد .

د ـ ۱ـ دامنه تغییرات[1]

اولین و ساده‌ترین شکل مشخصه‌های پراکندگی دامنه تغییرات می‌باشد .

" دامنه تغییرات یک مجموعه اعداد عبارت از تفاضل بین بزرگترین و "
" کوچکترین عدد موجود در آن مجموعه می‌باشد . "

وقتی‌که داده‌ها طبقه‌بندی شده‌باشند مقدار دامنه تغییرات برابر با تفاضل حد پائین‌ترین طبقه و حد بالای بالاترین طبقه می‌باشد .

مثال ۱۶ ـ ۲ : دامنه تغییرات را در موارد زیر پیدا کنید .

الف : مجموعه مشاهدات ، مجموعه مقابل می‌باشد .

$$\{24، 22، 21، 23، 12، 8، 15\}$$

ب : مجموعه مشاهدات مطابق جدول (۵ ـ ۲) باشد .

پاسخ : الف : $R = 24 - 8 = 16$

ب : $R = 204 - 148 = 56$

دامنه تغییرات اصولا" مشخصه خوبی برای ارزیابی پراکندگی نیست زیرا از وضعیت پراکندگی دربین دو نقطه حدی‌هیچگونه اطلاعی بدست‌نمی‌دهد . بااین حال‌مواردا ستفاده این مشخصه کم هم‌نیستند . در بسیاری از واحد های صنعتی تا وقتی که نوسانات و تغییرات

1 - The Range

محصول تولید شده از حیث اندازه، شکل، کیفیت و غیره بین دو حد خاص باشد دستگاه، سالم و بی‌عیب محسوب شده و تولید در وضعیت عادی تلقی می‌گردد.

د ـ ۲ـ میانگین انحرافات[1]

دومین مشخصه مهم در مورد پراکندگی یک متغیر میانگین انحرافات است. این مشخصه در مقایسه با دامنه تغییرات از کمال بیشتری برخوردار بوده و اشکال عمده‌ای که برای تغییرات ذکر شده، در اینجا مرتفع گردیده است. زیرا میانگین انحرافات بر خلاف دامنه تغییرات فقط ناظر بر دو نقطه حدی از مقادیر متغیر نیست بلکه در این مشخصه سعی شده است اولاً" تمام مقادیر متغیر منعکس شده و ثانیا" هریک از این مقادیر با همان میزان اهمیتی که در ایجاد پراکندگی دارند در آن ظاهر شوند که این معنا با لحاظ نمودن فراوانی مطلق هریک از مقادیر متغیر محقق می‌شود. به‌عنوان مثال اگر از ۱۰۰ نفر عضو یک نمونه ۹۹ نفر درآمدی درحدود میانگین داشته و فقط درآمد یک نفر از این افراد با میانگین درآمدها اختلاف زیاد داشته باشد نمی‌توان توزیع درآمد را در این نمونه بسیار پراکنده دانست. دادن ضریب ۱ (در مقابل ضریب ۹۹ برای سایرین) در این مورد به مقدار درآمد این شخص خاص سبب می‌شود که میزان پراکندگی زیاد بزرگ نشود در حالی که در دامنه انحرافات به‌خاطر همین یک نقطه، توزیع درآمد بسیار پراکنده جلوه می‌نمود.

مقصود از انحراف یک مقدار خاص از متغیر، اختلاف و تفاضل آن مقدار با میانگین مقادیر متغیر است. میانگین انحرافات بر مبنای این تصور وضع شده است که اگر بتوان در یک مشخصه به‌نحوی از انحاء تمام مقادیر انحرافات متغیر مورد نظر را در یک مشخصه منعکس نمود این مشخصه برای ارزیابی پراکندگی حول محور میانگین یک متغیر، ملاک خوبی خواهد بود. برای این کار ابتدا مجموع انحرافات از میانگین برای مقادیر مختلف یک متغیر با عبارت $\sum_{i=1}^{n}(x_i - \bar{x})$ نشان داده شد. اما این عبارت دارای دو اشکال است. اولاً" مقدار این عبارت در همه موارد معادل صفر است. این یکی از خواص مهم میانگین می‌باشد که انشاءالله در ضمن مسائل این فصل به اثبات خواهد رسید.

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0 \qquad (۲-۱۰)$$

ثانیا" در مواردی که اطلاعات طبقه‌بندی شده باشند همه n مقدار اصلی x در اختیار نبوده و بجای آن فقط به‌تعداد طبقات (h) مقادیر علائم هر طبقه در دسترس هستند و لذا عبارت فوق در چنین مواردی

1 - Mean Deviation

قابل استفاده نیست. برای رفع اشکال دوم عبارت قبلی اصلاح گردیده و به‌صورت $\sum_{i=1}^{h}(x_i-\bar{x})F_i$ تغییر شکل یافته است. این شکل جدید در مورد اطلاعات طبقه‌بندی شده به‌راحتی قابل اعمال است زیرا همه اطلاعات لازم برای آن در دسترس می‌باشند و حتی می‌توان آن را برای اطلاعات طبقه‌بندی نشده نیز با منظور نمودن($F_i = ۱$) به‌کاربرد. اما اشکال اول که در رابطه (۲ـ۱۰) منعکس گردید هنوز به‌قوت خود باقی است. برای رفع این اشکال به جای مقدار اصلی هر انحراف، قدر مطلق آنها را در نظر گرفته شده و جملات قبلی به صورت $\sum_{i=1}^{n}|x_i-\bar{x}|$ برای اطلاعات طبقه‌بندی نشده و $\sum_{i=1}^{h}|x_i-\bar{x}|F_i$ برای مقادیر طبقه‌بندی شده تبدیل شده‌اند.

برای سهولت کار نهایتاً" با تقسیم هریک از دو عبارت فوق بر حجم کل نمونه آنها را کوچک نموده از صور فوق که حاصل جمع انحرافات هستند به‌میانگین[1] انحرافات تبدیل می‌نمایند بدین ترتیب محصول نهائی تلاشهای فوق برای اطلاعات طبقه‌بندی نشده بصورت رابطه (۲ـ۱۱) و برای اطلاعات طبقه‌بندی شده به‌شکل رابطه (۲ـ۱۲) بیان می‌گردد.

$$M.D = \frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}| \quad \text{برای مقادیرطبقه‌بندی نشده} \qquad (۲-۱۱)$$

$$M.D = \frac{1}{n}\sum_{i=1}^{h}|x_i-\bar{x}|F_i \quad \text{برای مقادیرطبقه بندی شده} \qquad (۲-۱۲)$$

مثال ۱۷ ـ ۲ : انحراف متوسط را برای دو مورد زیر پیدا کنید.

الف : مقادیر مشاهده شده، اعداد (۱۱، ۵، ۵، ۴، ۳، ۲) باشند

ب : مقادیر مشاهده شده مطابق جدول (۵ ـ ۲) باشند.

پاسخ : الف :

$$\bar{x} = \frac{۲+۳+۴+۵+۵+۱۱}{۶} = ۵$$

چون اطلاعات طبقه‌بندی شده نیستند از رابطه (۲ـ۱۱) استفاده میکنیم.

$$M.D = \frac{1}{n}\ \Sigma|x_i-\bar{x}| = \frac{1}{۶}\left\{|۲-۵|+|۳-۵|+|۴-۵|+|۵-۵|+|۵-۵|+|۱۱-۵|\right\}$$

$$M.D = \frac{1}{۶}(۱۲) = ۲$$

ب : در مثال (۲ـ۸) مقدار میانگین این مشاهدات را ($\bar{x}$ =۱۷۷) به‌دست آوریم. چون اطلاعات طبقه‌بندی شده هستند از رابطه (۲ـ۱۲) استفاده میکنیم.

$$M.D = \frac{1}{n}\sum_{i=1}^{h}|x_i - \bar{x}|F_i = \frac{1}{200}\Big\{|152-177|(2)+|160-177|(10)+|168-177|(48)$$
$$+|176-177|(64)+|184-177|(56)+|192-177|(16)+|200-177|(4)\Big\}$$
$$= \frac{1}{200}\cdot(1440) = 7/2$$

د ـ ۳ ـ انحراف معیار:[1]

سومین و مهمترین مشخصه‌ای که برای ارزیابی پراکندگی مقادیریک‌متغیر حول محور میانگین وجود دارد انحراف معیار است . در راستای تلاش برای دست یافتن به‌یک مشخصه پراکندگی خوب ، عده‌ای از افراد برای رفع اشکالی که در عبارت $\Sigma(x_i - \bar{x})$ وجود داشت و در قالب رابطه (۱۰ـ۲) قبلا" بیان گردید ، راه استفاده از قدر مطلق را اختیار نمودند و با ادامه این راه میانگین انحرافات را به‌عنوان یک مشخصه پراکندگی وضع نمودند .اما گروه دیگری به‌این سئوال پاسخ دیگری دادند و آن مجذورنمودن همه جملات عبارت فوق بود . بدین ترتیب‌این‌گروه ازعبارت $\Sigma(x_i - \bar{x})^2$ برای دست یافتن به‌مشخصه‌ای درمورد پراکندگی استفاده‌نمودند .خاصیت این عبارت جدید آن است‌که ضمن رفع اشکال فوق‌الذکر انحرافات بسیار کم را به‌مراتب کوچکتر از مقدار واقعی ، انحرافات نه چندان زیاد را کمی بزرگتر از مقدار واقعی و انحرافات خیلی زیادرا به‌مراتب بزرگتراز میزان واقعی نشان می‌دهد .مقصود از انحرافات بسیار کم انحرافاتی هستند که مقدارشان در داخل پرانتز بیان شده است .

(۱>انحراف بسیارکم > ۱-)

واضح است که مجذور چنین اعدادی از خود آنها همیشه کوچکتر است در سایر موارد بجز انحرافات معادل ۱ که مجذورشان مساوی خودشان است بقیه مقادیر انحراف ، مجذورشان بزرگتر از خودشان می‌باشد .

ادامه کار همانند مورد انحراف از میانگین می‌باشد . یعنی پس‌از مجذور نمودن‌مقدار انحرافات و جمع آنها با یکدیگرآنها را بر حجم نمونه تقسیم‌می‌نمایند تا بدین‌ترتیب میانگین مجذور انحرافات را به‌دست آورند . به‌مشخصه حاصله یعنی به‌میانگین مجذور انحرافات ،واریانس و به‌جذرواریانس انحراف معیار گفته می‌شود . معمولا" انحراف‌معیار را (برای‌نمونه) با حرف S وواریانس را به‌دلیل آنکه مجذور انحراف معتیار است بسا S^2 نشان‌می‌دهند . مقدار واریانس برای داده‌های طبقه‌بندی نشده از رابطه (۱۳ـ۲) و بـرای

---

1 - Standard Deviation

طبقه‌بندی شده از رابطه (۱۴ـ۲) بدست می‌آید . البته این دو رابطه فقط برای محاسبه واریانس نمونه هستند و برای واریانس جامعه باید از رابطه دیگری استفاده شود .

برای داده‌های طبقه‌بندی نشده $$S^{۲} = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\bar{x})^{۲} \qquad (۲-۱۳)$$

برای داده‌های طبقه‌بندی شده $$S^{۲} = \frac{۱}{n-۱}\sum_{i=۱}^{h}(x_i-\bar{x})^{۲}F_i \qquad (۲-۱۴)$$

به احتمال قوی اکنون این سئوال در ذهن شما نقش بسته است که چرا در دو رابطه فوق بجای استفاده از n که حجم نمونه است در مخرج کسر از (n ــ۱) استفاده شده است ؟ گرچه دلیل اصلی این عمل در این مقطع از بحث قابل توضیح نیست و باید در فصول بعدی توضیح داده شود اما در اینجا نیز می‌توان یکی از دلایل این کار را مطرح نمود .

همانطوری که قبلا" گفته شد روابط (۱۳ـ۲) و (۱۴ـ۲) فقط برای محاسبه واریانس نمونه می‌باشند . معمولا" اگر به اطلاعات مربوط به جامعه اصلی دسترسی وجود داشته باشد میانگین جامعه را با (σ)، حجم جامعه را با (N) و واریانس آنرا با $\sigma^{۲}$ نشان داده و مقدار واریانس جامعه از رابطه زیر محاسبه می‌گردد .

$$\sigma^{۲} = \frac{۱}{N}\sum_{i=۱}^{h}(x_i-\mu)^{۲}F_i \qquad (۲-۱۵)$$

در بیشتر اوقات اطلاعات موجود از جامعه اصلی نیستند یعنی مقادیر N ، X و μ برای جامعه اصلی در اختیار ما قرار ندارند بلکه اطلاعات در دسترس مربوط به یک نمونه می‌باشند و محقق می‌خواهد از طریق واریانس این نمونه به واریانس جامعه اصلی دست یابد . تجربه نشان داده است که اگر در روابط (۱۳ـ۲) یا (۱۴ـ۲) از n در مخرج کسر استفاده کنیم واریانسهای نمونه‌ای که محاسبه می‌شوند به طور متوسط از واریانس جامعه اصلی مقداری کوچکتر هستند و اگر بجای n در مخرج کسر از (n ــ۱) استفاده شود واریانس نمونه احتسابی کمی بزرگ شده و به واریانس جامعه نزدیکتر می‌شود به طوری که می‌توان آن را به عنوان تخمینی از واریانس جامعه پذیرفت . به عنوان مثال اگر از رابطه (۱۵ـ۲) واریانس یک جامعه ۳۰ عضوی را معادل ۲ محاسبه کنیم و سپس از این جامعه تعدادی نمونه ۵ تائی اخذ نموده و واریانس هریک از این نمونه‌ها را بر مبنای روابط (۱۳ـ۲) یا (۱۴ـ۲) که در مخرج آنها (n) قرار داشته باشد محاسبه کنیم در بیشتر موارد مقدار واریانس کمتر از ۲ و مثلا" حدود ۱/۶ بدست

می‌آید که اگر بجای (n) از ($n-1$) در مخرج استفاده می‌شد این عدد به ۲ بسیار نزدیکتر بوده و می‌توانست به‌عنوان واریانس جامعه مورد استفاده قرار گیرد. به‌دلیل همین استفاده از $n-1$ به‌جای n در دو رابطه فوق واریانس نمونه به‌دست آمده را واریانس بدون انحراف یا بدون تورش می‌نامند که البته گاهی از این مفهوم با عناوین ناتور و نا اریب نیز یاد شده است.

با انجام عملیاتی برروی روابط (۲-۱۳) و (۲-۱۴) می‌توان روابط جدیدی را برای محاسبه واریانس به‌دست آورد که در عمل محاسبه واریانس را سهل تر می‌نمایند. برای نمونه این عملیات بر روی رابطه (۲-۱۳) انجام می‌شود.

$$S^{2}=\frac{1}{n-1}\Sigma\ (x_i-\bar{x})^{2}=\frac{1}{n-1}\Sigma\ (x_i^{2}-2x_i\bar{x}+\bar{x}^{2})=\frac{1}{n-1}(\Sigma\ x_i^{2}-2\Sigma\ x_i\bar{x}-\Sigma\ \bar{x}^{2})=$$

$$\frac{1}{n-1}(\Sigma x_i^{2}-2\Sigma x_i\bar{x}+n\bar{x}^{2})$$

می‌دانیم که مقدار $\bar{x}$ از رابطه زیر بدست می‌آید.

$$\bar{x}=\frac{\Sigma x_i}{n}$$

می‌توان در جمله آخر از جملات فوق به‌جای $\bar{x}$ معادل آن را قرار داد.

$$S^{2}=\frac{1}{n-1}\left[\Sigma x_i^{2}-2\frac{\Sigma x_i\Sigma x_i}{n}+n\frac{(\Sigma x_i)^{2}}{n^{2}}\right]=$$

$$\frac{1}{n-1}\left[\Sigma x_i^{2}-2\frac{(\Sigma x_i)^{2}}{n}+\frac{(\Sigma x_i)^{2}}{n}\right]=\frac{1}{n-1}\left[\Sigma x_i^{2}-\frac{(\Sigma x_i)^{2}}{n}\right]$$

شکل نهائی این رابطه بصورت زیر خواهد بود. واضح است که این رابطه برای داده‌های طبقه‌بندی نشده می‌باشد.

$$S^{2}=\frac{\Sigma x_i^{2}}{n-1}-\frac{(\Sigma x_i)^{2}}{n(n-1)}=\frac{n\ \Sigma x_i^{2}-(\Sigma x_i)^{2}}{n(n-1)} \qquad (۲-۱۶)$$

با عملیات مشابهی برای داده‌های طبقه‌بندی‌شده نیز می‌توان رابطه زیر را برای محاسبه واریانس بدست آورد.

$$S^{2}=\frac{n\ \sum_{i=1}^{h}F_i x_i^{2}-\left(\sum_{i=1}^{h}F_i x_i\right)^{2}}{n(n-1)} \qquad (۲-۱۷)$$

مثال ۱۸ ــ ۲ : مقدار انحراف معیار را در موارد زیر حساب کنید .

الف ــ مشاهدات نمونه مقادیر (۱۱، ۵، ۵، ۴، ۳، ۲) را داشته باشند .

ب ــ مشاهدات بند الف مربوط به‌یک جامعه ۶ نفری باشند .

ج ــ مقادیر مشاهده شده از نمونه مطابق جدول (۵ ــ ۲) باشند .

پاسخ : ابتدا واریانس را حساب نموده و سپس از آن جذر می‌گیریم تا انحراف معیار بدست آید .

الف از رابطه (۱۳ــ۲) استفاده‌می‌کنیم . در مثال (۱۷ــ۲) مقدار میانگین ($\bar{x} = ۵$) بدست آمد .

$$S^{۲} = \frac{\Sigma(x_i - \bar{x})^{۲}}{n-۱} = \frac{(۲-۵)^{۲} + (۳-۵)^{۲} + (۴-۵)^{۲} + \cdots + (۱۱-۵)^{۲}}{۵}$$

$$S^{۲} = \frac{۹ + ۴ + ۱ + ۳۶}{۵} = \frac{۵۰}{۵} = ۱۰$$

$$S = \sqrt{۱۰} = ۳/۱۶$$

ب ، چون در اینجا فرض شده است اطلاعات بند الف از جامعه اصلی اخذ شده‌اند بنابر این از رابطه (۱۵ــ۲) باید استفاده نمود . بازهم میانگین ($\mu = ۵$) می‌باشد .

$$S^{۲} = \frac{۱}{N} \Sigma(x_i - \mu)^{۲} = \frac{۵۰}{۶} = ۸/۳$$

$$S = ۲/۹$$

ج ــ اطلاعات مربوط به‌نمونه و طبقه‌بندی شده هستندلذا از رابطه (۱۴ــ۲) استفاده می‌کنیم . برای کمک به‌حل این بند از پاسخ بند ب مثال (۱۷ــ۲) می‌توانیم استفاده‌کنیم .

$$S^{۲} = \frac{۱}{n-۱} \Sigma(x_i - \bar{x})^{۲} F_i = \frac{۱}{۲۰۰ - ۱}\left[(۱۵۲ - ۱۷۷)^{۲}(۲) + \cdots\cdots + (۲۰۰ - ۱۷۷)^{۲}(۴)\right]$$

$$= \left[(-۲۵)^{۲}(۲) + (-۱۷)^{۲}(۱۰) + (-۹)^{۲}(۴۸) + (-۱)^{۲}(۶۴) + (۷)^{۲}(۵۶) + (۱۵)^{۲}(۱۶) + (۲۳)^{۲}(۴)\right]\frac{۱}{۱۹۹}$$

$$S^{۲} = \frac{۱۶۵۵۲}{۱۹۹} = ۸۳$$

$$S = ۹/۱$$

واضح است که بندهای الف و ج را به‌ترتیب با استفاده از روابط (۱۶ــ۲) و (۱۷ ــ ۲) نیز می‌توانستیم حل کنیم .

برای به‌دست‌آوردن تقریبی مقدارواریانس راه‌غیرمستقیمی نیزوجوددارد و آن‌استفاده از قضیه ریاضی دان روسی چبی‌شف[1] است .

---

1 - Chebysheves, Theorem

قضیه چبی‌شف

مصداق اصلی این قضیه بیشتر توزیعهای فراوانی زنگی شکل یا نرمال هستند اما در سایر توزیعها نیز کم و بیش صادق می‌باشد .

قضیه: ۵ ـ۲ـ احتمال اینکه هریک از مقادیر یک متغیر تصادفی مثل (X) به اندازه K برابر انحراف معیار از میانگین فاصله داشته باشند از رابطه زیر بدست می‌آید .

(۱۸ـ۲)

$$P\left[\mu-K\sigma<X<\mu+K\sigma\right] \geqslant 1-\frac{1}{K^2}$$

بدین‌ترتیب براساس این رابطه حداقل ۷۵٪ احتمال دارد که مقادیر X حداکثر به‌میزان دو برابر انحراف معیار ($K=2$) (2σ) با میانگین فاصله داشته باشند . اگر خاطرتان باشد قبلاً گفتیم که مفهوم احتمال به‌زبان فراوانی همان فراوانی نسبی می‌باشد یعنی در هر توزیع فراوانی و به‌خصوص توزیع زنگی شکل و نرمال حداقل $\frac{3}{4}$ از مشاهدات در فاصله دو انحراف از میانگین قرار می‌گیرند . قضیه فوق برای هر توزیع فراوانی اعم از آنکه مربوط به نمونه و یا جامعه باشد صادق است .

مثال ۱۹ ـ ۲: ثابت‌کنیدکه حداقل $\frac{3}{4}$ از مشاهدات در بندج مثال ۱۸ـ۲ درفاصله‌دوانحراف از میانگین قرار دارند .

پاسخ: در آن مثال حجم نمونه $n=200$ بود و بنابراین $\frac{3}{4}$ این مقداریعنی ۱۵۰ مقدارمتغیر تصادفی می‌باید در فاصله ($\bar{x}\pm 2S$) باشند . داریم .

$$\bar{x}+2S=177+2(9/1)=195/2$$
$$\bar{x}-2S=177-2(9/1)=158/8$$

بنابراین ۱۵۰ تا از مشاهدات باید در فاصله ($158/8<X<195/2$) باشند . با نگاهی به‌جدول (۵ـ۲) می‌توان فهمید که حدود ۱۹۴ مشاهده در فاصله فوق قرار دارند .

مثال ۲۰ ـ ۲: با استفاده‌از قضیه چبی‌شف مقدار تقریبی انحراف‌معیار را برای مثال (۵ـ۲) پیدا کنید .

پاسخ: براساس این قضیه تعدادی از مشاهدات که در فاصله مثلاً ($K=3$) برابر انحراف معیار یعنی حداکثر فاصله آنها تا میانگین ۳ S می‌باشد به‌صورت زیر محاسبه می‌شود .

$$P\left[\bar{x}-3S<X<\bar{x}+3S\right] \geqslant 1-\frac{1}{3^2}=\frac{8}{9}=89\%$$

که اگر این احتمال را در تعداد کل مشاهدات ضرب کنیم تعداد مشاهدات مورد بحث بدست می‌آید .

$$۲۰۰ \times ۹۰ \% = ۱۸۰$$

یعنی حداکثر ، فقط ۲۰ تا از مشاهدات می‌توانند فاصله‌ای بیش از (۳ S) با میانگین داشته باشند .اگر به‌جای ۲۰ عدد ۱۶ را اختیار کنیم ملاحظه می‌شود که ۱۸۴ مشاهده در فاصله‌زیر قرار می‌گیرند .

$$۱۵۶ < X < ۱۹۶$$

حال اگر جملات رابطه چبی‌شف را با عبارت فوق معادل قرار دهیم مقدار انحراف معیار به‌دست می‌آید .

$$\left[(\bar{x} - ۳S) < X < (\bar{x} + ۳S) \right] \geqslant \left[۱۵۶ < X < ۱۹۶ \right]$$

یعنی داریم :

$$(\bar{x} + ۳S) \geqslant ۱۹۶$$

قبلا" مقدار میانگین ($\bar{x} = ۱۷۷$) محاسبه گردید .

$$۱۷۷ + ۳S \geqslant ۱۹۶$$

$$۳S \geqslant ۱۹۶ - ۱۷۷$$

$$۳S \geqslant ۱۹$$

و در نتیجه مقدار انحراف معیار به‌صورت تقریبی به‌دست می‌آید .

$$S \geqslant \frac{۱۹}{۳} = ۶/۳$$

البته همانطور که مشاهده شد مقدار واریانس و یا انحراف معیار که از این طریق‌بدست می‌آید تقریبی بوده و ممکن است با انحراف معیار واقعی اختلاف زیادی داشته باشد ، اما از آنجا که مقدار واریانس و انحراف معیار هرچه کوچکتر باشد معمولا" مطلوب‌تر بوده‌و از توزیع عادلانه‌تری حکایت‌می‌کند همین‌قدر که از قضیه چبی‌شف می‌توان حداقل‌مقدارانحراف‌معیار را بدست‌آورد می‌تواند کمک بسیار بزرگی برای آمارگر محسوب گردد .

موضوعات اصلی مربوط به‌فصل دوم در همین جا تمام می‌شوند اما قبل از آنکه پایان این فصل اعلام گردد لازم است در مورد دو نکته که می‌توانند به‌عنوان مبحث کمکی و یا حتی مکمل مباحث این فصل تلقی گردند توضیحاتی داده شود . اولین نکته ، مربتط کردن مباحث‌این‌فصل با مبحث‌احتمالات و مخصوصا" محاسبه‌میانگین‌در مواردی‌است‌که‌به‌جای‌فراوانی

مطلق یا نسبی ،احتمال حوادث به‌عنوان اطلاعات داده می‌شوند .به‌چنین میانگینی اصطلاحا" امید ریاضی گفته می‌شود . دومین نکته چگونگی ساده کردن اطلاعات طبقه‌بندی شده برای محاسبه ساده‌تر مشخصه‌هائی است که در این فصل از آنها صحبت به‌میان آمد

هـ ـ محاسبه مشخصه ها با استفاده از احتمال حوادث

درمباحث قبلی گفته شد که مقدار احتمال یک حادثه بافراوانی نسبی آن برابرمی‌باشد به‌عبارت دیگر احتمال آنکه متغیر تصادفی X در هریک از h طبقه یک جدول توزیع فراوانی مقدار خود را اختیار کند با فراوانی نسبی آن طبقه مساوی بوده و اگر این فراوانی نسبی را در اختیار داشته باشیم در واقع احتمال آن حادثه در اختیار ما می‌باشد .

$$P_{x_i} = \frac{F_i}{n}$$

حال اگر به‌صفحات قبلی مجددا" مراجعه کرده و روابط مربوط به‌محاسبات مشخصه‌ها را مجددا" مرور کنیم خواهیم دید که در بسیاری از این روابط عبارت $\left(\frac{F_i}{n}\right)$ به‌کار رفته است . از آنجا که در بسیاری ازموارد بجای جداول توزیع فراوانی مربوط به‌یک‌متغیرجداول توزیع احتمال آن متغیر و به‌عبارت بهتر احتمالات مربوط به‌آن متغیر به‌عنوان اطلاعات ، در اختیار محقق قرار می‌گیرد او به‌راحتی می‌تواند به‌جای عبارت $\left(\frac{F_i}{n}\right)$ در هریک از این روابط P_{x_i} را قرار داده و مشخصه مربوطه را حساب کند .

در بین مشخصه‌های مختلف تمرکز و پراکندگی دو پارامتر میانگین و انحراف معیار (یا واریانس) از اهمیت بیشتری برخوردار هستند . به‌همین دلیل روابط مربوط به‌محاسبه‌این دو مشخصه با استفاده از مقادیر احتمالات ذیلا" آورده می‌شوند .

ابتدا در مورد میانگین صحبت می‌کنیم . چنانکه قبلا" گفته شد وقتی میانگین براساس احتمالات مربوط به‌یک حادثه محاسبه می‌شود نام آن تغییر نموده و از آن با عنوان میانگین مورد انتظار یا امید ریاضی[1] یاد می‌شود علت این تغییر نام این است که در مواردی که جدول توزیع فراوانی‌برای مقادیر یک متغیر داده می‌شود در واقع از حادثه‌ای صحبت‌می‌شود که اتفاق افتاده‌است و بنابراین میانگین مقادیر مشاهده‌شده‌به‌صورت‌قطعی‌محاسبه‌می‌شود .اما وقتی‌که‌احتمالات مربوط به‌یک‌حادثه‌داده‌می‌شوند میانگین‌احتسابی‌به‌هیچ وجه قطعیت نخواهد داشت . به‌عنوان مثال وقتی به‌شما می‌گویند میانگین اعداد مشاهده شده در n بار پرتاب یک تاس همتراز را پیدا کنید درصورتی که n عدد بسیار بزرگی‌باشد ،مسلما" با خود می‌گویید

1 - Expected Valve

۶ حالت ممکن است اتفاق بیفتد و احتمال مربوط به‌هریک از این ۶ حالت با دیگرحالات مساوی است بنابراین میانگین اعداد مشاهده شده مقدار زیر را خواهد داشت .

$$\bar{x} = \frac{۱+۲+۳+۴+۵+۶}{۶} = ۳/۵$$

حال از شما می‌پرسیم که آیا می‌توانید با قطعیت بگویید که در صورت انجام n بار آزمایش میانگین اعداد مشاهده شده معادل ۳/۵ خواهد بود . مسلما" خیر ، زیرا اگر چه احتمال وقوع چنین میانگینی بسیار زیاد است اما احتمال بسیار کمی نیز وجود دارد که این حادثه واقع نشود . به‌همین دلیل در اینجا می‌گویند میانگین مورد انتظار و یا امید ریاضی حادثه فوق برابر ۳/۵ است اگر چه ممکن است در عمل خلاف این‌امر اتفاق بیفتد .

هـ ـ ۱ ـ امید ریاضی یا میانگین

این پارامتر با [E (X)] نشان داده‌شده و همانطورکه‌قبلا" گفته شد با قراردادن P_x به‌جای $\frac{F_i}{n}$ در رابطه مربوط به‌محاسبه میانگین یعنی رابطه (۸ـ۲) روش کلی محاسبه آن بدست می‌آید .

$$\mu = \frac{۱}{N}\sum_{i=۱}^{h} x_i F_i = \sum_{i=۱}^{h} x_i \frac{F_i}{N}$$

$$E(X) = \sum_{i=۱}^{h} x_i P_i \qquad (۲-۱۹)$$

واضح است که در رابطه (۱۹ـ۲) به‌جای حرف h می‌توان از هرحرف دیگری از قبیل n و غیره نیز استفاده نمود .

مثال ۲۱ ـ ۲ : احتمالات مربوط به‌انتخاب هریک از طبقات درآمدی در یک‌جامعه درجدول زیر که به‌جدول توزیع احتمال معروف است مشخص شده‌اند از این جامعه می‌خواهیم فردی‌را انتخاب کنیم . امید ریاضی را برای این حادثه تعیین کنید .

x_i	۳۰۰۰۰	۴۰۰۰۰	۵۰۰۰۰	۶۰۰۰۰	۷۰۰۰۰
P_{x_i}	۱۰ ٪	۲۰ ٪	۴۰ ٪	۲۰ ٪	۱۰ ٪

جـــدول ۷ - ۲

پاسخ: از رابطه (۲-۱۹) استفاده می‌کنیم .

$$E(X) = \Sigma x_i P_{x_i} = ۳۰۰۰۰(۰/۱) + ۴۰۰۰۰(۰/۲) + ۰۰۰۰۰۰۰۰۰ + ۷۰۰۰۰(۰/۱)$$

$$E(X) = ۵۰۰۰۰$$

نکته مهمی که در اینجا باید ذکر شود این است که وقتی مشخصه‌های آماری بر اساس احتمالات مربوط به‌حوادث محاسبه می‌شوند طبیعتا" احتمالات داده شده باید مربوط به‌جامعه اصلی باشند و دیگر صحبتی از نمونه نیست زیرا اگر نمونه‌گیری انجام شده باشد به‌جای میانگین مورد انتظار یا امید ریاضی، خود میانگین قابل محاسبه است . امید ریاضی دارای خواصی است‌که دانستن آنها درعملیات آماری کمک مهمی به‌آمارگر می‌نماید .اکنون تعدادی از این خواص ذکر می‌شوند .

خواص امید ریاضی

خواص امید ریاضی را در قالب چند قضیه به‌صورت زیر بیان می‌کنیم .

قضیه: ۲-۶ . اگر a و b مقادیری ثابت باشند داریم

$$E(aX + b) = aE(X) + b \qquad (۲-۲۰)$$

نتیجتا" اگر داشته باشیم ($a = ۰$) خواهیم داشت .

$$E(b) = b \qquad (۲-۲۱)$$

و اگر($b = ۰$) باشد داریم :

$$E(aX) = aE(X) \qquad (۲-۲۲)$$

قضیه: ۲-۷: اگر $g(X)$ و $h(X)$ توابعی از متغیر تصادفی (X) باشند داریم :

$$E\left[g(X) \pm h(X)\right] = E\left[g(X)\right] \pm E\left[h(X)\right] \qquad (۲-۲۳)$$

مثال ۲۲ ــ ۲: مقدار مصرف افراد یک جامعه$g(X)$ و مقدار سرمایه‌گذاری همین افراد$h(X)$ بر اساس دو رابطه زیر از مقدار درآمد آنان تبعیت می‌کند . اگر جدول توزیع احتمال درآمد این افراد مطابق جدول مثال (۲-۲۱) باشد و بخواهیم یک نفر را به‌طور تصادفی از این جامعه انتخاب کنیم امید ریاضی مصرف و سرمایه‌گذاری برای این فرد را حساب کنید .

مصرف $C = g(X) = 1000 + 0/7(X)$

سرمایه‌گذاری $I = h(X) = -1000 + 0/2(X)$

پاسخ: از رابطه (۲۳ـ۲) باید استفاده نمود.

$$E[g(X)+h(X)] = E[g(X)] + E[h(X)] = ?$$

باید دو پارامتر طرف راست معادله را جدا جدا محاسبه کرده و سپس با یکدیگر جمع نمود.

$$E[g(X)] = \Sigma g(x_i)P_{x_i} \quad , \quad E[g(X)] = \Sigma h(x_i)P_{x_i}$$

روشن است که برای محاسبه این دو پارامتر باید مقادیر $g(X)$ و $h(X)$ را داشته‌باشیم، لذا براساس توابع موجود این مقادیر را حساب می‌کنیم.

$x = 30000 \rightarrow g(X) = 1000 + 0/7(30000) = 22000$ و

$h(X) = -1000 + 0/2(30000) = 5000$

به‌همین‌ترتیب سایر مقادیررا حساب‌نموده درجدول توزیع احتمال‌به‌صورت زیر قرارمی‌دهیم

X	۳۰۰۰۰	۴۰۰۰۰	۵۰۰۰۰	۶۰۰۰۰	۷۰۰۰۰
P(X)	۱۰ ٪	۲۰ ٪	۴۰ ٪	۲۰ ٪	۱۰ ٪
g(X)	۲۲۰۰۰	۲۹۰۰۰	۳۶۰۰۰	۴۳۰۰۰	۵۰۰۰۰
h(X)	۵۰۰۰	۷۰۰۰	۹۰۰۰	۱۱۰۰۰	۱۳۰۰۰

جدول ۸ـ۲

حال می‌توان مقادیر دو پارامتر $E[g(X)]$, $E[h(X)]$ را از جدول فوق‌محاسبه‌نمود.

$$E[g(X)] = 22000(0/1) + 29000(0/2) + 36000(0/4) + 43000(0/2) + 50000(0/1) \qquad E[g(X)] = 35000$$

به‌همین ترتیب پارامتر دیگر مقدار زیر را خواهد داشت.

$$E[\ h(X)\ ] = ۹۰۰۰$$

و نهایتا خواهیم داشت :

$$E\ [\ g(X)+h(X)] = E\ [\ g(X)] + E\ [\ h(X)] = ۳۵۰۰۰ + ۹۰۰۰ = ۴۴۰۰۰$$

قضیه : ۸ــ۲ . اگر X و Y دو متغیرتصادفی باشند داریم

$$E(X \pm Y) = E(X) \pm E(Y) \qquad (۲-۲۴)$$

قضیه : ۹ــ۲ . اگر X و Y دو متغیرتصادفی‌مستقل مثل قدافرا دو رنگ پوست‌آنها باشند داریم .

$$E(XY) = E(X)E(Y) \qquad (۲-۲۵)$$

هــ ۲ــ واریانس

قبلا" گفته شدکه وقتی اطلاعات اولیه براساس احتمالات مربوط به‌هریک‌ازمقادیر متغیر تصادفی داده‌می‌شوند غالبا" این‌اطلاعات مربوط به‌جامعه‌می‌باشند .در چنین مواردی‌طبیعی است که واریانس و انحراف معیاری هم که محاسبه می‌شود مربوط به‌جامعه خواهد بود و لــذا برای محاسبه واریانس درچنین مواردی باید دررابطه (۱۵ــ۲) که‌مربوط به محاسبه واریانس جامعه است اصلاحاتی انجام داده وآن را مورد استفاده قرار داد . ذیلا" این رابطه را به‌منظور تطبیق با وضعیت جدید اطلاعات اولیه اصلاح می‌کنیم .

$$\sigma^۲ = \frac{۱}{N}\sum_{i=۱}^{h}(x_i-\mu)^۲F_i = \sum_{i=۱}^{h}(x_i-\mu)^۲\frac{F_i}{N} = \sum_{i=۱}^{h}(x_i-\mu)^۲P_i$$

اگر فرض کنیم $y_i = (x_i-\mu)^۲$ باشد می‌توان عبارت آخر جمله فوق را به‌صورت‌زیرنوشت .

$$\sigma^۲ = \sum_{i=۱}^{h}(x_i-\mu)^۲P_i = \Sigma\, y_iP_i = E(Y) = E(X-\mu)^۲$$

اما گفتیم که وقتی اطلاعات براساس احتمالات حوادث داده می‌شوند باید به‌جای میانگین امید ریاضی یا میانگین مورد انتظار را قرار داد یعنی چون[$\mu = E(X)$]می‌باشــد آخرین عبارت جمله فوق به‌صورت زیر قابل نمایش است .

$$E(X-\mu)^۲ = E\,[X-E(X)]^۲$$

در نتیجه خواهیم داشت :

$$\sigma^۲ = E[X-E(X)]^۲ \qquad (۲-۲۶)$$

به‌وسیله عملیات زیرمی‌توان براساس رابطه فوق ،رابطه دیگری را نیز برای محاسبه واریانس به‌دست آورد .

$$\sigma^۲ = E[X-E(X)]^۲ = E\left\{X^۲ - ۲XEX + [E(X)]^۲\right\} = E(X^۲) - ۲[E(X)]^۲ + E[E(X)]^۲$$

چون $E(X)$ عدد ثابتی است داریم :

$$E\,[\,E(X)]^۲ = [E(X)]^۲$$

در نتیجه دنباله عملیات به‌قرار زیر است .

$$\sigma^۲ = E(X^۲) - ۲[E(X)]^۲ + [E(X)]^۲$$
$$\sigma^۲ = E(X^۲) - [E(X)]^۲ \qquad (۲-۲۷)$$

مثال ۲۳ ــ ۲ . مقدار انحراف معیار را برای مثال (۲۱ــ۲) محاسبه نمایید .

پاسخ : استفاده از رابطه (۲۶ــ۲) برای حل این مثال مناسب‌تر است و لذا آن را مورد استفاده قرار می‌دهیم .

$$\sigma^۲ = E[X-E(X)]^۲ = \Sigma[X_i - E(X)]^۲\, P_i =$$
$$(۳۰۰۰۰-۵۰۰۰۰)^۲(۰/۱) + (۴۰۰۰۰-۵۰۰۰۰۰)^۲(۰/۲) + (۰)^۲(۰/۴) +$$
$$(۶۰۰۰۰-۵۰۰۰۰)^۲(۰/۲) + (۷۰۰۰۰-۵۰۰۰۰)^۲(۰/۲)$$

$$\sigma^۲ = ۱۲۰،۰۰۰،۰۰۰ \quad \text{و} \quad \sigma = ۱۰۹۵۴$$

واریانس دارای خواصی است که چون با استفاده از قوانین امید ریاضی این خواص راحت‌تر قابل اثبات هستند در مباحث قبلی مورد بحث قرار نگرفتند .

خواص ریاضی مهم واریانس

این خواص نیز همچون خواص مهم امید ریاضی در قالب چند قضیه بیان می‌شوند .
قضیه: ۱۰ــ۲ . اگر X متغیر تصادفی و b مقدار ثابت باشد داریم :

$$\sigma^{2}_{(X+b)} = \sigma^{2}_{X} = \sigma^{2} \qquad (۲-۲۸)$$

قضیه: ۱۱ــ۲ . اگر X متغیر تصادفی و a مقداری ثابت باشد داریم :

$$\sigma^{2}_{(aX)} = a^{2}\sigma^{2}_{X} = a^{2}\sigma^{2} \qquad (۲-۲۹)$$

اثبات

$$\sigma^{2}_{(aX)} = E[aX-E(aX)]^{2} = E[aX-aE(X)]^{2} =$$

$$E\left\{a^{2}X^{2}-2a^{2}XE(X)-a^{2}[E(X)]^{2}\right\}=$$

$$a^{2}E[X^{2}-2XE(X)]+ [E(X)]^{2} = a^{2}E[X-E(X)]^{2} = a^{2}\sigma^{2}$$

قضیه: ۱۲ــ۲ . اگر X و Y دو متغیر تصادفی مستقل باشند داریم .

$$\sigma^{2}_{(X+Y)} = \sigma^{2}_{X}+\sigma^{2}_{Y} \qquad (۲-۳۰)$$

اثبات

$$\sigma^{2}_{(X+Y)}=E[(X+Y)-E(X+Y)]^{2}=E[X+Y-EX-EY]^{2} =E\left\{[X-E(X)]+[Y-E(Y)]\right\}^{2}=$$

$$E\left\{[X-E(X)]^{2}+[Y-E(Y)]^{2}+2[X-E(X)][Y-E(Y)]\right\}$$

$$\sigma^{2}_{(X+Y)} =E[X-E(X)]^{2}+E[Y-E(Y)]^{2}+2E[X-E(X)][Y-E(Y)]$$

دوعبارت اول جمله فوق σ^{2}_{X} و σ^{2}_{Y} می‌باشند و ذیلا" ثابت می‌شود که عبارت سوم نیزمساوی صفر است .

$$E[X-E(X)][Y-E(Y)]=E[XY-YE(X)-XE(Y)+E(X)E(Y)] =$$

$$E(XY)-EYEX-EXEY+E[E(X)E(Y)] =0$$

صفرشدن مقدار عبارت آخر سطر بالا براساس قضیه (۹ــ۲) صورت گرفت . بنابراین داریم :

$$\sigma^2_{(X+Y)} = \sigma^2_X + \sigma^2_Y$$

قضیه: ۱۳ــ۲ . اگر X و Y دو متغیر مستقل باشند داریم:

$$\sigma^2_{(X-Y)} = \sigma^2_X + \sigma^2_Y$$

اثبات: براساس قضیه (۱۲ــ۲) می‌توان نوشت:

$$\sigma^2_{(X-Y)} = \sigma^2_X + \sigma^2_{(-Y)}$$

براساس قضیه (۱۱ــ۲) داریم

$$\sigma^2_{(-Y)} = (-1)^2 \sigma^2_Y = \sigma^2_Y$$

و در نتیجه قضیه (۱۳ــ۲) اثبات می‌شود .

$$\sigma^2_{(X-Y)} = \sigma^2_X + \sigma^2_Y$$

بدین‌ترتیب به‌آخرین مبحث فصل‌دوم یعنی مبحث‌ساده‌کردن اطلاعات جهت‌محاسبه پارامترها و مشخصه‌ها می‌رسیم .

و ـ محاسبه مشخصه‌ها از طریق ساده کردن داده‌ها

بحث موضوع این بند را در دو قسمت ساده کردن داده‌ها و محاسبه مشخصه‌ها توضیح می‌دهیم .

و ـ ۱ـ ساده‌کردن داده‌ها

در بسیاری از موارد جداول توزیع فراوانی داده شده دارای ارقام بسیار بزرگی‌هستند به‌طوری که‌محاسبه مشخصه‌ها براساس این‌داده‌ها بسیار دشوارخواهد بود . در چنین‌مواقعی x_i ها را که دارای مقادیر بزرگ می‌باشند براساس رابطه زیر به U_i که مقادیر بسیار کوچکی دارد تبدیل نموده و مشخصه را براساس این مقادیرکوچک محاسبه‌می‌نمایند و سپس با انجام عملیاتی برروی مقادیر محاسبه شده مشخصه‌ها ، مقدار اصلی آنها را به‌دست می‌آورند .

$$U_i = \frac{x_i - a}{c}$$

در این رابطه a مرکز یا علامت یکی از طبقات می‌باشد که معمولا" سعـی می‌شــود از

طبقـات وسطی مقدار آن انتخاب شود و c عرض هریک از طبقات می‌باشد .

مثال ۲۴ ــ ۲ : اطلاعات مربوط به‌مثال (۵ ــ ۲) را که در جدول (۵ ــ ۲) آمده‌اند ازطریـق رابطه فوق ساده کنید .

پاسـخ : مقدار a را از طبقه چهارم انتخاب می‌کنیم یعنی داریم $a = ۱۷۶$

حدودطبقـــــه	x_i	F_i	U_i
۱۴۸-۱۵۶	۱۵۲	۲	-۳
۱۵۶-۱۶۴	۱۶۰	۱۰	-۲
۱۶۴-۱۷۲	۱۶۸	۴۸	-۱
۱۷۲-۱۸۰	۱۷۶	۶۴	۰
۱۸۰-۱۸۸	۱۸۴	۵۶	۱
۱۸۸-۱۹۶	۱۹۲	۱۶	۲
۱۹۶-۲۰۴	۲۰۰	۴	۳

جــــــدول ۹-۲

$$U_۱ = \frac{۱۵۲-۱۷۶}{۸} = -۳$$

$$U_۲ = \frac{۱۶۰-۱۷۶}{۸} = -۲$$

$$U_۷ = \frac{۲۰۰-۱۷۶}{۸} = ۳$$

و ــ ۲ــ محاسبه مشخصه‌ها

ازبین مشخصه‌های‌مختلف تمرکز و پراکندگی دومشخصه مهمتر یعنی‌میانگین و واریانس را از این طریق محاسبه می‌کنیم .

برای محاسبه میانگین ابتدا مقدار $\bar{U}$ را محاسبه‌نموده وسپس از طریق رابطه زیرمقدار $\bar{x}$ را به‌دست می‌آوریم .

$$\bar{x} = C\bar{U} + a$$

اثبات رابطه فوق بسیار ساده است . می‌دانیم که $\bar{U}$ مساوی مقدار زیر است .

$$\bar{U} = \frac{1}{n}\Sigma U_i F_i = \frac{1}{n}\Sigma \frac{x_i - a}{c} F_i$$

$\frac{1}{c}$ به‌عنوان یک عددثابت می‌تواند از داخل سیکما بیرون آمده و عملیات به‌صورت‌زیرادامه یابـد .

$$\bar{U} = \frac{1}{nc}\Sigma(x_i-a)F_i = \frac{1}{nc}\Sigma(x_iF_i-aF_i) = \frac{1}{nc}\Sigma x_iF_i - \frac{1}{nc}a\Sigma F_i = \frac{1}{c}\bar{x} - \frac{na}{nc} = \frac{\bar{x}}{c} - \frac{a}{c} = \frac{\bar{x}-a}{c} = \bar{U}$$

و نهایتا" خواهیم داشت :

$$\bar{x} = c\bar{U}+a$$

به‌همین ترتیب برای محاسبه S_x^2 ابتدا S_u^2 را محاسبه‌نموده و سپس آنرا براساس رابطه زیر به S_x^2 تبدیل می‌کنیم .

$$S_x^2 = c^2S_u^2$$

اثبات رابطه فوق بسیار ساده و به‌صورت زیر است .

$$S_u^2 = S^2_{\left(\frac{x-a}{c}\right)} = S^2_{\left[\frac{1}{c}(x-a)\right]} = \frac{1}{c^2}S^2_{(x-a)}$$

از قضیه (۱۰-۲) داریم :

$$S^2_{(x-a)} = S_x^2$$

با جایگذاری این مقدار در عبارت بالا خواهیم داشت :

$$S_u^2 = \frac{1}{c^2}S_x^2$$

با بردن c به‌آنطرف تساوی رابطه بالا ثابت می‌شود .

$$S_x^2 = c^2S_u^2$$

مثال ۲۵ - ۲ . با استفاده از اطلاعات ساده شده مثال قبل میانگین و انحراف معیار متغیر تصادفی X را محاسبه کنید .

پاسخ : ابتدا میانگین $\bar{U}$ را محاسبه نموده و سپس آنرا به $\bar{x}$ تبدیل می‌کنیم .

$$\bar{U} = \frac{۱}{n}\sum_{i=۱}^{h} U_i F_i = ۲(-۳)+۱۰(-۲)+۴۸(-۱)+۶۴(۰)+۱۵(۱)+۱۶(۲)+۴(۳) \quad \frac{۱}{۲۰۰} =$$

$$\frac{-۶-۲۰-۴۸+۵۶+۳۲+۱۲}{۲۰۰} = \frac{۲۶}{۲۰۰} = ۰/۱۳$$

$$\bar{x} = c\bar{U}+a = ۸(۰/۱۳)+۱۷۶ = ۱۷۷/۴ = ۱۷۷$$

حال $S_u^۲$ را محاسبه نموده و سپس آن را به $S_x^۲$ تبدیل می‌کنیم .

$$S_u^۲ = \frac{۱}{n-۱} \Sigma (U_i-\bar{U})^۲ F_i = \frac{۱}{۱۹۹} \quad (۰۳-۰/۱۳)^۲ ۲+(-۲-۰/۱۳)^۲ ۱۰+$$

$$(-۱-۰/۱۳)^۲ ۴۸+(۰-۰/۱۳)^۲ ۶۴ +(۱-۰/۱۳)^۲ ۵۶+(۲-۰/۱۳)^۲ ۱۶+(۳-۰/۱۳)^۲ ۴$$

$$= \frac{۲۵۹}{۱۹۹} = ۱/۳ \qquad S_x^۲ = c^۲ S_u^۲ = ۸^۲(۱/۳) = ۶۴(۱/۳) = ۸۳ \qquad S_x = ۹/۱$$

جوابهای مزبور با جوابهای دو مثال (۸ــ۲) و (۱۸ــ۲) که در آنها نیز مقادیر همین دو مشخصه محاسبه گردید کاملا" هماهنگ است .

مسائـــل فصل دوم

۱ ــ قضایای (۱ــ۲) ، (۲ــ۲) ، (۴ــ۲) ، (۶ــ۲) ، (۷ــ۲) و نیز رابطهٔ (۱۰ــ۲) را اثبات کنید .

۲ ــ قضایای (۸ــ۲) و (۹ــ۲) و نیز رابطه (۱۷ــ۲) را ثابت کنید .

۳ ــ مجموعه مشاهدات از وزن ۴۰ دانشجوی پسر یک دانشگاه که به‌ترتیب صعودی و از چپ به‌راست مرتب شده‌اند از این قرار است . (ارقام برحسب پوند هستند)

۱۴۲، ۱۴۰، ۱۴۰، ۱۳۸، ۱۳۸، ۱۳۶، ۱۳۵، ۱۳۵، ۱۳۵، ۱۳۲، ۱۲۸، ۱۲۶، ۱۲۵، ۱۱۹،
۱۵۲، ۱۵۰، ۱۵۰، ۱۴۹، ۱۴۸، ۱۴۷، ۱۴۷، ۱۴۶، ۱۴۶، ۱۴۵، ۱۴۵، ۱۴۴، ۱۴۴، ۱۴۲
۱۷۶، ۱۷۳، ۱۶۸، ۱۶۵، ۱۶۴، ۱۶۳، ۱۶۱، ۱۵۸، ۱۵۷، ۱۵۶، ۱۵۴، ۱۵۳،

الف ــ با استفاده از ارقام اصلی میانه را پیدا کنید :

ب : با استفاده از جدول توزیع فراوانی که فاصله طبقات آن ۹ بوده و از عدد ۱۱۸ شروع می‌شود میانه را پیدا کنید .

۴ ــ جدول قانون توزیع صفت X به‌صورت زیر است . مطلوبست میانگین ،میانه و نما برای این صفت .

حدودطبـــــقات	F_i
۴۰ - ۴۵	۲
۴۵ - ۵۰	۳
۵۰ - ۵۵	۷
۵۵ - ۶۰	۱۰
۶۰ - ۶۵	۱۲
۶۵ - ۷۰	۱۵
۷۰ - ۷۵	۱۲
۷۵ - ۸۰	۱۰
۸۰ - ۸۵	۸
۸۵ - ۹۰	۳
	۸۲

۵ ــ میانگین حسابی مقدار X در جامعه ذیل را پیدا کنید .

۴،۵،۲،۸،۴،۵،۴،۳،۸،۴،۵،۶،۸،۲،۵،۴،۵،۶،۳،۵

۶ ــ فروش روزانه چهار فروشنده عبارتست از (۳۰۰۰۰،۶۵۰۰،۶۰۰۰،۵۰۰۰) ریال

الف ــ میانگین حسابی فروش آنها را محاسبه کنید .

ب : آیا این میانگین مشخصه‌ای مناسب برای تمرکز فروش آنها می‌باشد ؟

۷ ــ نمرات دانشجویی در شش امتحان ۸۷،۷۸،۶۸،۷۲،۹۱،۸۴ می‌باشد ، میانه نمرات را پیدا کنید .

۸ ــ در جامعه‌ای جدول توزیع فراوانی صفت X به‌صورت زیر است . مطلوبست واریانس و انحراف معیار این توزیع .

$x_i - x_{i+1}$	F_i
۱/۵ - ۲/۵	۲
۲/۵ - ۳/۵	۳
۳/۵ - ۴/۵	۵
۴/۵ - ۵/۵	۸
۵/۵ - ۶/۵	۱۰
۶/۵ - ۷/۵	۸
۷/۵ - ۸/۵	۷
۸/۵ - ۹/۵	۷
	۵۰

۹ ــ جدول زیر نشان دهنده نمرات آزمایش هوش ۴۸۰ دانش‌آموز یک مدرسه ابتدایی می‌باشد . اولاً میانگین و ثانیاً انحراف معیار این نمرات را محاسبه کنید .

X	۷۰، ۷۴، ۷۸، ۸۲، ۸۶، ۹۰، ۹۴، ۹۸، ۱۰۲، ۱۰۶، ۱۱۰، ۱۱۴، ۱۱۸، ۱۲۲
F	۴، ۹، ۱۶، ۲۸، ۴۵، ۶۶، ۸۵، ۷۲، ۵۴، ۳۸، ۲۷، ۱۸، ۱۱، ۴

۱۰ ــ برای توزیع احتمال ذیل میانگین و واریانس را تعیین کنید .

X	۸	۱۲	۱۶	۲۰	۲۴
P(X)	۱/۸	۱/۸	۳/۸	۱/۸	۳/۱۲

۱۱ ــ برای شخصی در یک معامله ٪۶۰ احتمال یک سود ۳۰۰ ریالی و ٪۴۰ احتمال یک زیان ۱۰۰ ریالی وجود دارد. امید ریاضی اورا برای این سود و زیان تعیین کنید.

۱۲ ــ در کیسه‌ای ۲ توپ سفید و ۳ توپ سیاه وجود دارد افراد A و B و C و D به‌ترتیبی که نام برده‌شدند هریک توپی‌را بدون جایگذاری از این کیسه خارج می‌کنند. به‌اولین‌فردی که توپ سفید را از کیسه خارج کند ۱۰ تومان جایزه داده می‌شود. امید ریاضی هریک از آنها را برای دریافت جایزه ۱۰ تومانی تعیین کنید.

۱۳ ــ متوسط‌انحرافات را برای داده‌های جدول زیر پیدا کنید. در صورتی که $\bar{x}$=۳/۴۱۲۵ باشد.

X	۱/۷	۲/۲	۲/۷	۳/۲	۳/۷	۴/۲	۴/۷	
F	۲	۱	۴	۱۵	۱۰	۵	۳	۴۰

۱۴ ــ جدول ذیل توزیع فراوانی دستمزد ساعتی ۶۵ نفر از کارکنان یک کارخانه را برحسب تومان نشان می‌دهد. تعیین کنید:

الف: حد پائین رده ششم ب: حد بالای رده چهارم

ج: وسط‌رده سوم د: حد بالای رده پنجم

هـ: وسعت رده پنجم و: فراوانی رده سوم

ز: فراوانی نسبی رده سوم

ح: رده‌ای که بیشترین فراوانی را دارد.

ط: درصد کارکنانی که کمتراز ۸۰ تومان در ساعت می‌گیرند.

ی: درصد کارکنانی که از صدتومان کمتر ولی حداقل ۶۰ تومان دستمزد می‌گیرند.

یا: توزیع فراوانی نسبی برای هر طبقه

F_i	$x_i - x_{i+1}$
۸	۵۰/۰۰-۵۹/۹۹
۱۰	۶۰/۰۰-۶۹/۹۹
۱۶	۷۰/۰۰-۷۹/۹۹
۱۴	۸۰/۰۰-۸۹/۹۹
۱۰	۹۰/۰۰-۹۹/۹۹
۵	۱۰۰/۰۰-۱۰۹/۹۹
۲	۱۱۰/۰۰-۱۱۹/۹۹
۶۵	جمع

یب : هیستوگرام فراوانی مطلق و نسبی (درصد)

یج : توزیع فراوانی تجمعی، توزیع تجمعی درصد نسبی، نمودار فراوانی تجمعی و نمودار تجمعی درصد را نشان دهید.

ید : جدول توزیع فراوانی تجمعی " یا بیشتر " و نمودار توزیع فراوانی تجمعی " یا بیشتر " را نشان دهید. مقصود فراوانی تجمعی است که مثلاً " بگوید ۷ نفر بیش از ۱۰۰ تومان دستمزد دریافت می‌کنند.

یه : میانگین و میانه دستمزد کارکنان فوق را محاسبه نموده و انحراف معیار توزیع دستمزدها را از طریق جذر واریانس محاسبه کنید.

پاسخ مسائل فرد فصل دوم

۱ ـ قضیه ۱ـ۲ .

$$\Sigma (x_i+y_i) = \Sigma x_i + \Sigma y_i$$

اثبات ـ طرف اول را بسط داده و سپس دسته‌بندی می‌کنیم .

$$\Sigma (x_i+y_i = (x_1+y_1)(x_2+y_2) + \cdots\cdots + (x_n+y_n) =$$
$$(x_1+x_2+\cdots\cdots+x_n)+(y_1+y_2+\cdots\cdots+y_n) = \Sigma x_i + \Sigma y_i$$

قضیه ۲ـ۲ . اگر Y عدد ثابتی باشد . داریم :

$$\Sigma Y_i = nY$$

اثبات ـ طرف اول را بسط می‌دهیم . چون همه Y ها باهم برابرند داریم .

$$\Sigma Y_i = Y_1+Y_2+\cdots\cdots+Y_n = nY$$

قضیه ۴ـ۲ . اگر c عدد ثابتی باشد داریم :

$$\Sigma c x_i = c\, \Sigma x_i$$

اثبات ـ طرف اول را بسط داده و از عدد ثابت فاکتور می‌گیریم .

$$\Sigma cx_i = cX_1 + cx_2 + \cdots\cdots + cx_n = c(x_1 + x_2 + \cdots\cdots + x_n) = c\Sigma x_i$$

قضیه ۶ـ۲ . اگر a و b مقادیری ثابت باشند داریم .

$$E(aX+b) = aE(X)+b$$

اثبات :

$$E(aX+b) = \Sigma (ax_i+b)P_{x_i} = \Sigma ax_i P_{x_i} + \Sigma bP_{x_i} = a\,\Sigma x_i P_{x_i} + b\,\Sigma P_{x_i} = aE(X)+b$$

چون $\Sigma P_{x_i} = 1 \quad , \quad \Sigma x_i P_{x_i} = E(X)$

قضیه ۷ـ۲ . اگر $g(X)$ و $h(X)$ توابعی از متغیر تصادفی X باشند داریم :

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$$

اثبات :

$$E[g(X) \pm h(X)] = \Sigma [g(x_i) \pm h(x_i)]P_{x_i} = \Sigma [g(x_i)P_{x_i} \pm h(x_i)P_{x_i}] =$$

$$E[g(X)] \pm E[h(X)]$$

قضیه ۱۰ـ۲ . اگر X متغیر تصادفی و b مقداری ثابت باشد داریم :

$$\sigma^2_{(X+b)} = \sigma^2_X = \sigma^2$$

اثبات :

$$\sigma^2_{(X+b)} = E[(X+b)-E(X+b)]^2 = E[X+b-E(X)-b]^2 = E[X-E(X)]^2 = \sigma^2_X = \sigma^2$$

رابطه ۱۰ـ۲ .

$$\Sigma (x_i - \bar{x}) = 0$$

اثبات :

$$\Sigma (x_i - \bar{x}) = \Sigma x_i - \Sigma \bar{x} = \Sigma x_i - n\bar{x} = \Sigma x_i - n\frac{\Sigma x_i}{n} = \Sigma x_i - \Sigma x_i = 0$$

۳ ـ الف : دو عدد میانی این سری هردو ۱۴۶ بوده و لذا میانه که میانگین آنها باشد نیز ۱۴۶ است .

ب : ابتدا جدول توزیع فراوانی را به‌صورت خواسته شده می‌نویسیم و سپس از طریق رابطه مربوط به‌میانه این مشخصه را محاسبه می‌کنیم . از جدول ذیل مقادیر $L = ۱۴۵$ ، $\frac{n}{۲} = ۲۰$ ، $F = ۱۷$ ، $F_m = ۱۲$ و $c = ۹$ به‌دست می‌آید که آنها را در رابطه فوق قرار داده و مقدار میانه را به‌دست می‌آوریم :

$$Me = L + \left(\frac{n/۲ - F}{F_m}\right)c = ۱۴۵ + \frac{۲۰ - ۱۷}{۱۲}\, ۹ = ۱۴۷/۳$$ پوند

$x_i - x_{i+1}$	F_i	ΣF_i
۱۱۸ - ۱۲۷	۳	۳
۱۲۷ - ۱۳۶	۵	۸
۱۳۶ - ۱۴۵	۹	۱۷
۱۴۵ - ۱۵۴	۱۲	۲۹
۱۵۴ - ۱۶۳	۵	۳۴
۱۶۳ - ۱۷۲	۴	۳۸
۱۷۲ - ۱۸۰	۳	۴۰
	۴۰	

۵ -

$$\mu = \frac{\Sigma F_i x_i}{N} = \frac{(۶)(۵) + (۳)(۲) + (۶)(۲) + (۴)(۵) + (۲)(۲) + (۸)(۳)}{۲۰}$$

$$\mu = ۴/۸$$

۷ -

$$Me = \frac{۷۸ + ۸۴}{۲} = ۸۰$$

۹ ــ صفت X را به U تبدیل‌نموده و نتیجه محاسبات را به‌صورت جدول توزیع فراوانی نشان می‌دهیم . برای این کار ($a = ۹۴$) را اختیار می‌نمائیم .

$$U_i = \frac{x_i - a}{c}$$

$x_1 = ۷۰$ $\quad U_1 = \frac{۷۰-۹۴}{۴} = -۶$ $\quad F_1U_1 = -۲۴$ $\quad F_1 . U_1^۲ = ۱۴۴$

$x_۲ = ۷۴$ $\quad U_۲ = -۵$ $\quad F_۲U_۲ = -۴۵$ $\quad F_۲ . U_۲^۲ = ۲۲۵$

$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$

x_i	U_i	F_i	U_iF_i	$U_i^۲F_i$
۷۰	-۶	۴	-۲۴	۱۴۴
۷۴	-۵	۹	-۴۵	۲۲۵
۷۸	-۴	۱۶	-۶۴	۲۵۶
۸۲	-۳	۲۸	-۸۴	۲۵۲
۸۶	-۲	۴۵	-۹۰	۱۸۰
۹۰	-۱	۶۶	-۶۶	۶۶
۹۴	۰	۸۵	۰	۰
۹۸	۱	۷۲	۷۲	۷۲
۱۰۲	۲	۵۴	۱۰۸	۲۱۶
۱۰۶	۳	۳۸	۱۱۴	۳۴۲
۱۱۰	۴	۲۷	۱۰۸	۴۳۲
۱۱۴	۷ ۵	۱۸	۹۰	۴۵۰
۱۱۸	۶	۱۱	۶۶	۳۹۶
۱۲۲	۷	۷	۴۹	۳۴۳
	۸			
		۴۸۰	۲۳۶	۲۴۰۴

اولا:

$$\mu_x = a + c\bar{U} = a + c\left(\frac{\Sigma F_iU_i}{n}\right) = ۹۴ + ۴\left(\frac{۲۳۶}{۴۸۰}\right) \simeq ۹۶$$

ثانیا:

$$\delta = c\sqrt{\sigma_u^۲} = c\sqrt{\frac{\Sigma F_iU_i^۲}{n} - \left(\frac{\Sigma F_iU_i}{n}\right)^۲} = c\sqrt{\left(\frac{\Sigma F_iU_i^۲}{n}\right) - (\mu_u)^۲}$$

$$= ۴\sqrt{\frac{۲۳۷۴}{۴۸۰} - (/۵)^{۲}} \qquad \delta = ۱۰/۴$$

۱۱ -

$$E\ (X) = ۳۰۰)(۰/۶) + (-۱۰۰)(۰/۴) = ۱۸۰-۴۰ = ۱۴۰ \quad \text{ریال}$$

۱۳ ــ ابتدا جدول زیر را که درآن محاسبات لازم انجام شده است تهیه می‌کنیم .

x_i	F_i	$x_i-\bar{x}$	$F_i\lvert x_i-\bar{x}\rvert$
۱/۷	۲	-۱/۷۱۲۵	۳/۴۲۵۰
۲/۲	۱	-۱/۲۱۲۵	۱/۲۱۲۵
۲/۷	۴	-۰/۷۱۲۵	۲/۸۵۰۰
۳/۲	۱۵	-۰/۲۱۲۵	۳/۱۸۷۵
۳/۷	۱۰	۰/۲۸۷۵	۲/۸۷۵۰
۴/۲	۵	۰/۷۸۷۵	۳/۹۳۷۵
۴/۷	۳	۱/۲۸۷۵	۳/۸۶۲۵
	۴۰		۲۱/۳۵۰۰

بر اساس مقادیر جدول فوق مقدار میانگین انحرافات به‌صورت زیر به‌دست می‌آید .

$$M.D = \frac{\Sigma\lvert x_i-\bar{x}\rvert F_i}{n} = \frac{۲۱/۳۵}{۴۰} = ۰/۵۳۳۸$$

پاسخ مسائل زوج فصل دوم

۴ — $\bar{x}$ = ۶۷/۱   Me = ۶۷/۳۳   Mo = ۶۷/۵

۶ ـ الف : $\bar{x}$ = ۱۱۸۷۵   ب : خیر

۸ — σ^2 = ۳/۶۹ σ = ۱/۹۲

۱۰ — E(X) = ۱۷ σ^2 = ۲۷

۱۲ — E(A) = ۴   E(B) = ۳   E(C) = ۲   و E(D) = ۱

۱۴ ـ الف : ۱۰۰   ب : ۸۹/۹۹   ج — ۷۴/۹۹۹۵   د : ۹۹/۹۹   هـ : ۹/۹۹

و : ۱۶   ز : ۲۴/۶%   ح : سوم   ط : ۵۲/۳%   ی : ۷۶/۹%

یا ، یب ، یج ، ید :   یه : $\bar{x}$ = ۹۵/۹۷   Me = ۷۹/۰۶ σ = ۱۰/۴۷

فصل سوم

توزیع احتمال

مقـدمه :

در این مقدمه در مورد مفهوم توزیع احتمال و ضرورت بحث پیرامون آن صحبت شده و بالاخره چهارچوب کلی مباحث این فصل مشخص می‌گردند .

فرض کنید به‌شما بگویند ده‌عدد سیب را بین ۴ نفر تقسیم‌کنید و شمابراساس‌ضابطه‌ای که در ذهن‌دارید (ضابطه‌ای از قبیل سن ، میزان گرسنگی و غیره) این ده عددسیب‌را بین افراد فوق مطابق جدول زیر تقسیم می‌کنید .

شماره افراد	اول	دوم	سوم	چهارم
تعداد سیبها	۲	۳	۳	۲

جدول ۱–۳

واضح است که‌بهترین نامی که برای جدول فوق می‌توان انتخاب کرد جدول تقسیم یا توزیع سیب می‌باشد و به‌ضابطه‌ای‌نیز که براساس‌آن عمل تقسیم انجام گرفت‌ضابطه تقسیم سیب می‌توان نام نهاد ،احتمال نیز همانند یک سیب ، اعم‌از آنکه تجربی و یا نظری‌باشدبالاخره به‌ترتیبی بین‌حوادث مختلف‌ممکن‌الوقوع در یک آزمایش تقسیم می‌گردد و می‌توان‌به‌عنوان مثال مجموع احتمالات مربوط به‌آزمایش پرتاب یک تاس همتراز را که مساوی واحد می‌باشد بین ۶ حادثه مختلف ممکن‌الوقوع در این آزمایش نمونه‌ای از این توزیع دانست . در ایـن مثال مجموعه احتمالات یعنی عدد (۱) نقش همان ده عدد سیب مثال قبل را ایفا می‌کند و ۶ حادثه ممکن‌الوقوع جایگزین چهارنفری می‌شود که آن سیبها بین آنان تقسیم می‌گردید و بالاخره ضابطه‌توزیع متساوی و علی‌السویه ضابطه‌ای است‌که چگونگی‌توزیع وتقسیم‌احتمال

را بین این ۶ حادثه بیان می‌کند . جدول ۲ نیز جدول توزیع احتمال این آزمایش می‌باشد .

نوع حادثه	۱	۲	۳	۴	۵	۶
مقدار احتمال	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

جدول ۲ـ۳ . توزیع احتمال آزمایش پرتاب یک تاس همتراز

البته جدول ۲ـ۳ . همانطور که ملاحظه می‌شود چگونگی توزیع یک احتمال نظری را نشان می‌دهد . با توزیع احتمال تجربی نیز در فصل دوم به‌قدر کافی آشنا شده‌اید زیرا همانطور که قبلا" گفته شد احتمال تجربی چیزی جز همان فراوانی نسبی نیست و در واقع به یک جدول توزیع فراوانی نسبی جدول توزیع احتمال تجربی می‌گویند . نکته بسیار مهمی که یک جدول توزیع احتمال نظری را از جدول توزیع احتمال تجربی متمایز می‌سازد این است که در توزیع احتمال تجربی احتمالات مختلف به‌صورت نسبتا" تصادفی بین حوادث مربوطه تقسیم می‌گردند . همانطور که در فصل قبل گفته شد هیچ تضمینی وجود ندارد که مقدار احتمال که براساس تجربه برای یک حادثه به‌دست می‌آید میزان واقعی احتمال وقوع همان حادثه باشد چون در وقوع حوادث مختلف در عمل تصادف نقش مهمی ایفا می‌کند و ممکن است نتایج حاصله در هر آزمایش با آزمایشهای مشابه دیگر تفاوت داشته باشد و اگر بخواهیم میزان واقعی احتمال را به‌دست آوریم باید این کار را از طریق نظری انجام دهیم زیرا مقدار احتمال نظری برخلاف احتمال تجربی براساس ضوابط و قواعد مختلفی که در فصل اول به تفصیل پیرامون آنها صحبت گردید به‌دست می‌آید و صدفه و تصادف در آن نقشی ندارد تفاوت بین توزیع احتمال تجربی و نظری نیز از همین جا ناشی می‌شود زیرا به‌دلیل دخالت مسئله تصادف در تعیین مقدار احتمال هر حادثه ، در توزیع احتمال تجربی نمی‌توان از یک ضابطه توزیع قطعی و لایتغیر صحبت نمود درحالی که در توزیع احتمال نظری چنین قاعده و ضابطه‌ای وجود دارد و مقدار احتمال در هر آزمایش بر اساس ضابطه خاصی بین حوادث مختلف تقسیم می‌گردد .وجود قواعد مختلف برای محاسبه مقادیر احتمال نظری آزمایشهای مختلفی که در فصل اول مورد بحث قرار گرفتند برای اثبات قانونمندی توزیع احتمال نظری کفایت می‌کند در حالی که برای محاسبه احتمال تجربی هیچ ضابطه‌ای به‌جز انجام آزمایش در عمل وجود نداشت البته در موردی که حجم نمونه خیلی بزرگ باشد می‌توان گفت که مقدار احتمال تجربی با مقدار احتمال نظری تقریبا" برابر است .

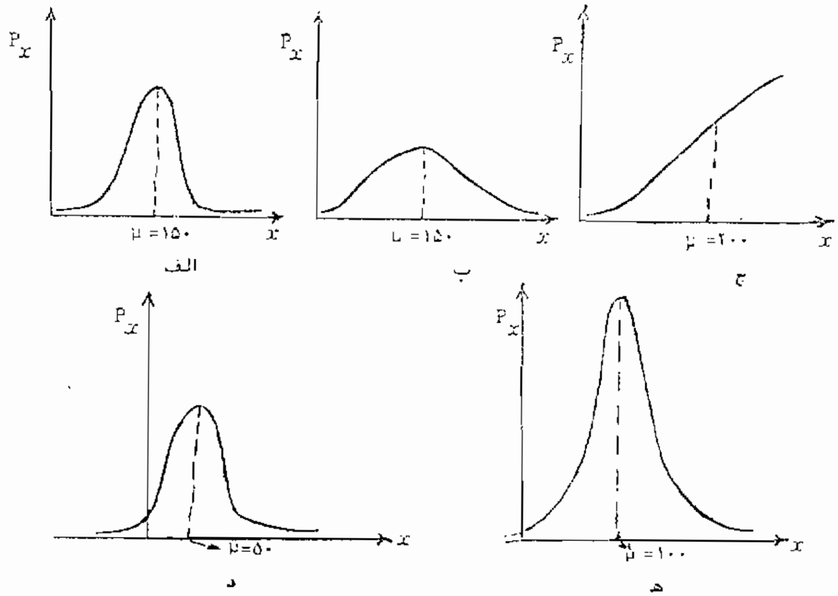
نتیجه بحث فوق این است‌که برخلاف احتمال تجربی درتوزیع احتمال نظری‌ازضابطه و قاعده حاکم بر توزیع می‌توان سخن راند و در واقع مقدار احتمال در هر آزمایش براساس قاعده‌ای خاص مشخص می‌شود که ممکن است این قاعده با قاعده توزیع احتمال نظری‌دریک آزمایش دیگر مشابه باشد و یا اینکه با آن تفاوت داشته باشد .

به‌طور کلی توزیع احتمال سه مشخصه اصلی دارد که اگر وضعیت این سه مشخصه‌روشن باشد می‌توان گفت که آن توزیع دقیقا" مشخص شده است . یکی از این سه مشخصه شکل‌کلی توزیع است که منحنی توزیع احتمال بیانگر آن می‌باشد . در شکل (۱ـ۳) دوشکل الف و ب از نظر شکل کلی توزیع زنگی شکل و شبیه به‌یکدیگر بوده و با شکل ج متفاوت می‌باشند . دومین مشخصه اصلی یک توزیع احتمال امید ریاضی یا میانگین مقادیر متغیری است‌که‌توزیع احتمال ، احتمالات آن مقادیر را نشان می‌دهد . و بالاخره سومین مشخصه یک توزیع احتمال انحراف معیار مقادیر متغیر مورد نظر است . اگر دو توزیع از نظر شکل کلی مشابه یکدیگر بوده و میانگین و انحراف معیار آنها با هم برابر باشد آن دو توزیع کاملا" مشابه یکدیگر می‌باشند . ذیلا" مفهوم میانگین (امید ریاضی) و انحراف معیار بیشتر توضیح داده می‌شود .

اگر یک توزیع احتمال ، احتمالات مربوط به‌مقادیر مختلف متغیر (X) را نشان دهد واضح است که مشخصه‌های تمرکز مربوط به‌مقادیر (X) از قبیل میانگین ، میانه و نمای آنها وضعیت تمرکز آن مقادیر را مشخص می‌سازند بدین معنی که این مشخصه‌ها نشان می‌دهندکه مقادیر مختلف (X) حول چه نقطه‌ای متمرکز شده‌اند . در شکل (۱ـ۳) میانگین به عنوان مشخصه اصلی تمرکز در نظر گرفته شده و همان‌طور که مشاهده می‌شود توزیع‌های الف و ب در این شکل در نقطه ($x = ۱۵۰$) متمرکز شده‌اند . از طرف دیگر بدون شک مشخصه‌های پراکندگی بخصوص انحراف معیار چگونگی پراکندگی مقادیر مختلف متغیر (X)را حول‌محور میانگین نشان می‌دهند چنانکه قبلا" گفته شد هرچه مقدار انحراف معیار بزرگتر باشدفاصله مقادیر مختلف (X) نسبت به‌میانگین بیشتر بوده و بعکس هرقدر این مقدار کوچکتر باشدمقادیر مختلف (X) با میانگین فاصله کمتری داشته و مقدارشان نسبت به‌میانگین از تغییرات جزئی و مختصرتری برخوردار است .

در منحنی‌های توزیع احتمال نشان داده شده در شکل (۱ـ۳) دو شکل (الف و ب)از نظر شکل کلی توزیع و امیدریاضی مشابه یکدیگرهستند اما انحراف معیارشکل (ب) از(الف) بیشتراست . دو شکل (الف و د)از نظر شکل‌کلی توزیع و انحراف معیارتقریبا" شبیه یکدیگر هستند اما میانگین (الف) از (د) بیشتر است . دوشکل (الف و هـ) از نظر شکل‌توزیع شبیه هم هستند اما میانگین و انحراف معیار (الف) از (هـ) بزرگترست و بالاخره شکل ج هم‌از

جهت شکل کلی توزیع هم از نظر میانگین و انحراف معیار با سایرین متفاوت است .



شکل (۱ ـ ۳)

با توضیحات فوق مفهوم توزیع احتمال تا حدودی روشن شد .اکنون زمان بررسی ضرورت بحث توزیع احتمال در مباحث آماری فرا رسیده است . بدون شک پس از روشن شدن مفهوم توزیع احتمال ، در ذهن خواننده محترم این سئوال متبادر شده است که چرا باید چنین مفهومی در آمار مورد بحث و گفتگو قرار گیرد ؟ در این باره قبلا" در مقدمه کتاب به اجمال سخن گفته شده است و در اینجا این مطلب با تفصیل بیشتری مورد بحث قرار می‌گیرد .

همانطور که در مباحث قبلی کرارا" گفته شده است دو گونه احتمال تجربی و نظری وجود داشته و به‌تبع این دوگانگی از دو نوع توزیع احتمال تجربی و نظری نیز می‌توان سخن گفت . در مباحث فصل دوم کاربرد احتمال تجربی مشخص گردیده و نتیجتا" ضرورت این گونه توزیع احتمال نیز اینک واضح و مبرهن است . همانطور که در آنجا ملاحظه گردید محاسبه مشخصه‌های تمرکز و پراکندگی نمونه‌که یکی از ابزارهای اصلی‌کار یک آمارگرمی‌باشد ، بدون وجود توزیع احتمال تجربی ممکن نیست و همین مسئله وجود اینگونه توزیع احتمال در آمار را به‌خوبی تفسیر و توجیه می‌کند . اما نقش احتمال تجربی و توزیع مربوط به‌آن در عملیات آماری اگرچه بسیار مهم و ضروری می‌باشد ، چندان زیاد نیست و این احتمال‌نظری

و به‌خصوص توزیع آن می‌باشد که در جریان کار آماری، نقشی بنیادی ایفا نموده و بارها مورد استفاده آمارگر قرار می‌گیرد و لذا لازم است در اینجا از ضرورت بحث توزیع احتمال نظری و نقش آن در عملیات آماری صحبت شود.

به‌طور کلی توزیع احتمال نظری در آمار دارای دو کاربرد اساسی می‌باشد. یکی از این دو کاربرد آن است که محاسبه مقدار احتمال مربوط به‌حوادث مختلف را ممکن‌می‌سازد.

دومین کاربرد این گونه توزیع احتمال آن است که از طریق آن، آمارگر بین نمونه و جامعه رابطه برقرار نموده و با استفاده از این توزیعها بر مبنای نتایج حاصل از نمونه نتایج مطلوب خود در مورد جامعه را هرچند به‌طور تقریبی به‌دست می‌آورد. به‌جرأت می‌توان گفت که در صورت عدم وجود توزیع احتمال نظری، ممکن نبود که یک آمارگر از طریق نمونه‌گیری و اخذ چند نمونه از جامعه مورد نظر بتواند در باره آن جامعه هیچ‌گونه اظهار نظری بنماید. چون کاربرد اول توزیع احتمال نظری موضوع بحث این فصل بوده و کاربرد دوم، موضوع بحث فصل آینده است، لذا در اینجا از بحث بیشتر پیرامون دومین کاربرد توزیع احتمال نظری در آمار خودداری نموده و این کار را به‌فصل آینده موکول می‌کنیم و در عوض در مورد کاربرد اول توزیع احتمال[1] یعنی ممکن نبودن محاسبه مقدار احتمال توضیح بیشتری می‌دهیم.

برای رعایت اختصار در توضیح کاربرد اول توزیع احتمال و اجتناب از اطاله کلام فقط به‌یکی از موارد استفاده مهم توزیع احتمال پس از ذکر یک مثال غیر آماری در مورد اهمیت نقش توزیع در محاسبات مختلف اکتفا می‌شود.

فرض کنید به‌شما بگویند که ده نفر به‌ترتیب قد و از کوچک به‌بزرگ مرتب شده‌اند. اگر قد نفر اول یعنی قدکوتاه‌ترین فرد ۱۵۰ سانتیمتر باشد اندازه قد نفر پنجم را به‌دست آورید مسلما" با اندک تأملی خواهید گفت با چنین اطلاعات ناچیزی نمی‌توان این مسئله را حل نمود و باید علاوه بر معلومات فوق چیزهای دیگری را نیز بدانیم تا بتوانیم اندازه قد نفر پنجم را محاسبه کنیم. اکنون اگر به‌شما بگویند که مقدار تفاوت قد هر فرد با نفر بعد از خود و همچنین با نفر قبل از خود دو سانتی متر است یعنی از نفر قبل از خود دو سانتیمتر بلندتر و از نفر بعدی دو سانتیمتر کوتاهتر است به‌راحتی می‌توانید مقدار مطلوب مسئله را پیدا کنید. برای این کار اگر قد نفر اول را با y_1 و نفر دوم را با y_2 . . . و نفر پنجم را با y_5 نشان دهیم

---

[1] ــ از این ببعد هرجا لفظ توزیع احتمال بکار می‌رود مقصود همان توزیع احتمال نظری می‌باشد و در صورت صحبت از احتمال تجربی، قید تجربی نیز به عبارت توزیع احتمال افزوده می‌شود.

مقدار y_5 را به‌صورت زیر می‌توان به‌دست آورد .

$$y_2 = y_1 + 2 = 150 + 2 = 152 \quad \text{و} \quad y_3 = y_2 + 2 = 152+2 = 154$$

$$y_4 = y_3 + 2 = 154 + 2 = 156 \quad \text{و} \quad y_5 = y_4 + 2 = 156+2 = 158$$

برای حل مسئله فوق در واقع از دو اطلاع اولیه استفاده گردید . این دو اطلاع عبارت بودند از اینکه اولاً" قد نفر اول ۱۵۰ سانتیمتر است ، ثانیاً" قد هریک از افراد با افراد قبلی و بعدی خود فقط دو سانتیمتر تفاوت دارد . اگر درست دقت کنیم در واقع این دو اطلاع بر رویهم چگونگی توزیع و تقسیم عنصر قد ، بین این ده نفر را بیان می‌کنند . یعنی در واقع محاسبه قد نفر پنجم که مطلوب مسئله بود بر اساس اطلاعات مربوط به‌توزیع قد بین افراد فوق صورت گرفت .

مثال فوق نقش و اهمیت شکل توزیع را در حل مسائل مختلف تا حدودی نشان می‌دهد در محاسبه احتمال مربوط به‌هرحادثه‌ای بدون داشتن اطلاع از چگونگی توزیع احتمال، محاسبه غیر ممکن خواهدبرد . ممکن است از خود بپرسید که در این صورت چطور در فصل اول بدون طرح مسئله‌توزیع احتمال ، مقدار احتمال مربوط به‌حوادث مختلف تعیین می‌شد ؟ در پاسخ باید گفت که در آنجا هم از توزیع احتمال بی‌نیاز نبودیم و برای سهولت بحث فرض کردیم که همه نقاط نمونه متساوی الاحتمال هستند که این خود نوعی توزیع احتمال است . اما در بسیاری از موارد نمی‌توان از فرض فوق استفاده نمود زیرا نقاط نمونه مربوط به‌یک آزمایش متساوی‌الاحتمال نیستند (مثل آزمایش پرتاب کبریت و میلیونها آزمایش دیگر) و لذا باید برای حل مسائل فوق حتماً" نسبت به‌توزیع احتمال آگاهی داشته باشیم و همین مسئله ضرورت بحث توزیع احتمال را به‌خوبی تبیین می‌کند . از این گذشته با استفاده از توزیع احتمال در عملیات آماری می‌توان حل مسائل مربوط به‌احتمال فاصله‌ای[1] مخصوصاً" در مورد متغیرهای پیوسته‌را به‌طرز شگفت‌انگیزی تسهیل نمود که این فایده اخیر در قالب مثال ذیل بیشتر توضیح داده می‌شود .

فرض کنید متغیرگسسته (X) می‌تواند صد مقدار از (۱۰۰–۱) را اختیار کند .احتمال آن که این متغیر در یک آزمایش مقدار($30 > X > 10$)را اختیار کند چقدر است ؟ معنای این مسئله آن است که احتمال آن که متغیر تصادفی ما مقادیر ($X = 11$ ، $X = 12$ ، ...... و $X = 29$) را در آزمایش اختیار کند چقدر است ؟ بر مبنای روشهای محاسبه احتمالی که در فصل اول گفته شد طبیعی است که ابتدا باید ۱۹ مقدار (P_{11} و P_{12} و ...... ، P_{29})را

---

۱ -- احتمال فاصله‌ای و نیز احتمال نقطه‌ای بعداً" توضیح داده خواهند شد .

حساب نموده و سپس با یکدیگر جمع کنیم .

$$P(۱۰< X < ۲۰) = P_{۱۱}+ P_{۱۲}+........+ P_{۲۹}$$

ملاحظه می‌گردد که برای حل مسئله فوق از روشهای قبلی باید ۱۹ احتمال مختلف را جداگانه حساب کرده و سپس باهم جمع کنیم یعنی یک عملیات ۲۰ مرحله‌ای را بـــرای حل این مسئله انجام دهیم . اما چنانکه در ادامه این فصل خواهیم دید اگر بتوان همین‌مسئله را از طریق توزیعهای احتمال‌شناخته شده‌رایج در آمار حل کنیم بیشترمحاسبات فوق قبـــلا" انجام شده و بااستفاده از نتایج آنها می‌توان مقداراحتمال مورد نظررا خیلی ساده محاسبه نمـــود .

اهمیت کاربرد اول توزیع احتمال وقتی بیشتر روشن می‌شود که فرض‌کنیم درمثال فوق متغیر تصادفی (X)متغیری پیوسته بوده و می‌تواند بین دو عدد (۱۰۰-۱) بینهایت‌مقدار را اختیار کند . در این صورت محاسبه احتمال فاصله‌ای فوق امری محال و غیرممکن می‌شود زیرا متغیر ما در فاصله (۳۰-۱۰)نیز بینهایت مقدار می‌تواند اختیار کند و بـــرای محاسبـــه احتمال‌آنکه (X) در این فاصله مقداربگیرد باید بینهایت‌احتمال‌مختلف جداجدا محاسبـه شده و سپس با یکدیگر جمع شوند و واضح است که محاسبه احتمال بینهایت مقدارمختلفی‌که متغیر ما می‌تواند در فاصله مورد نظر اختیار کند امری غیر ممکن است در حالی که اگربتوان از قاعده توزیع احتمال به‌شکلی که بعدا" گفته می‌شود استفاده کرد . محاسبه احتمـال نا ممکن فوق امری بسیار ساده وآسان خواهد شد .

اکنون مفهوم‌توزیع احتمال روشن شده‌است و دانستیم‌که بحث در مورد توزیع‌احتمال ازآن جهت ضرورت داردکه اولا" بدون وجود توزیع‌احتمال ، محاسبه‌احتمال نقطه‌ای‌به‌هیچ وجه ممکن نیست و ثانیا" بدون استفاده مستقیم از توزیع احتمال ، محاسبه احتمال‌فاصله‌ای درموردمتغیر پیوسته هم مطلقا" امکان‌پذیر نبوده و در مورد متغیر گسسته این‌کار با سختی انجام می‌گیرد . اینک پس از بحث در این دو مورد نوبت به‌بیان کلی چهار چـــوب مباحـث بعدی این فصل فرا می‌رسد .دراینجا با بیان‌جریان‌کلی‌بحث و معرفی عناوین‌کلی‌موضوعاتی که در این فصل مورد بحث قرار می‌گیرند مقدمه فصل سوم را به‌پایان می‌بریم . روشن اســت که قبل از هربحث دیگری باید درمورد چگونگی استفاده از شکل توزیع احتمال در محاسبات مربوط به‌احتمال صحبت شود . در ضمن این بحث متوجه خواهیم شد که استفــاده از توزیــع احتمال برای محاسبه احتمال حوادث مختلف نیز گرچه از روشهای گفته شده قبلـی بهتـــر است ، کاری چندان ساده‌نیست زیرا محاسبات ریاضی نسبتا" مشکلی را لازم دارد که چه‌بسا

از حوصله یک آمارگر خارج باشد. برای حل این مشکل به‌صورت زیر چاره‌جوئی شده است. اولا" از بین اشکال بسیار زیاد و متنوع توزیع احتمال که ممکن است وجود داشته باشند تعدادی را که از سایرین بیشتر مورد استفاده قرار می‌گیرند مشخص نموده و ثانیا" برای این تعداد شکل‌توزیع همه‌محاسبات ریاضی مختلفی‌را که ممکن است در جریان یک تحقیق آماری ضرورت پیدا کنند انجام داده و نتایج این محاسبات را به‌صورت جداولی که به‌این نوشته ضمیمه شده‌اند در اختیار محققین قرار داده‌اند.

اشکال توزیع احتمال مهم فوق‌الذکر را می‌توان در دو گروه اصلی به‌صورت توزیعهای احتمال مربوط به‌متغیر تصادفی گسسته و توزیعهای مربوط به‌متغیر پیوسته طبقه‌بندی نمود که این کار در اینجا انجام گرفته است. بدین ترتیب عناوین کلی بحث در ادامه این‌فصل به‌صورت زیر خواهد بود.

الف ــ چگونگی استفاده از توزیع احتمال برای محاسبه مقدار احتمال.

ب ــ توزیعهای احتمال مهم

ب ــ ۱ــ توزیعهای احتمال گسسته شامل توزیعهای دوجمله‌ای، فوق هندسی و پواسن

ب ــ ۲ ــ توزیع احتمال پیوسته شامل توزیع نرمال

لازم به‌تذکر است که در بند ب فقط توزیعهایی مورد بحث قرار می‌گیرند که مورد استفاده قابل توجهی در محاسبه مقدار احتمال حوادث مختلف دارند و توزیعهای احتمال دیگری غیر از اینها نیز وجود دارند که چون مورد استفاده اصلی آنها در ایجاد ارتباط بین نمونه و جامعه می‌باشد بحث پیرامون آنها به‌فصل چهارم موکول گردید.

الف ــ چگونگی استفاده از توزیع احتمال برای محاسبهٔ مقدار احتمال

گفتیم که یکی از فواید مهم و عمده توزیع احتمال کمک به‌حل مسائل مختلف و محاسبه مقدار احتمال حوادث گوناگون می‌باشد. اکنون سئوال این است که چگونه از توزیع احتمال‌برای این منظور استفاده می‌شود؟ قبل از پاسخ به‌این سئوال تذکر یک نکته‌لازم است وآن این است‌که احتمال را می‌توان از یک نظر به‌دونوع احتمال نقطه‌ای و فاصله ای تقسیم کرد. مقصود ازاحتمال نقطه‌ای آن است‌که متغیر تصادفی مثل (X)یک مقدارخاص‌رااختیار کند. مثلا" احتمال آنکه قد یک‌فرد که بطور تصادفی انتخاب می‌شود دقیقا" (۱۵۰cm)باشد، یک‌احتمال نقطه‌ای نامیده می‌شود. در حالی که احتمال آنکه قد همین فرد مقدار خود را در فاصله (۱۶۰cm ــ ۱۵۰) اختیار کند، احتمال فاصله‌ای نام دارد. محاسبه احتمال نقطه‌ای برای متغیرپیوسته و گسسته با روش نسبتا" مشابهی انجام می‌گیرد اما محاسبه احتمال‌فاصله‌ای

برای یک متغیر گسسته با روشی متفاوت از متغیر پیوسته انجام می‌گیرد و لذا بحث بند الف را در سه قسمت ذیل دنبال خواهیم کرد .

الف ــ ۱ ــ محاسبه احتمال نقطه‌ای

الف ــ ۲ــ محاسبه احتمال فاصله‌ای برای متغیر گسسته

الف ــ ۳ــ محاسبه احتمال فاصله‌ای برای متغیر پیوسته

الف ــ ۱ــ محاسبه احتمال نقطه‌ای

مثالی را که در مورد توزیع قد ده نفر در مقدمه همین فصل ذکر شد به‌خاطر بیاورید . درآنجا با اطلاعات مربوط به‌توزیع متغیر قد ، قد نفر پنجم به‌مقدار (۱۵۸ cm) تعیین گردید . اگر درست‌دقت شود مجموع اطلاعات بکار گرفته شده در آن مثال‌ضابطه‌توزیع قد را بین افراد مورد بحث توضیح می‌دادند . این ضابطه را به‌راحتی می‌توانستیم در قالب یک رابطه ریاضی به‌صورت زیر بیان نموده و به‌جای محاسبه مطلوب مسئله در ۴ مرحله‌ای که در آنجاانجام‌شد ، آن را در یک مرحله محاسبه کنیم .

$$y = ۲x + ۱۴۸$$

در این رابطه (y) نشان دهنده قد افراد و x نشان دهنده شماره افراد می‌باشد . براساس این رابطه قد نفر پنجم به‌صورت زیر محاسبه می‌شود .

$$y = ۲(۵) + ۱۴۸ = ۱۵۸$$

درست همانند مثال فوق اگر بخواهیم ازتوزیع احتمال نیز برای محاسبه مقداراحتمال استفاده کنیم باید از ضابطه توزیع استفاده کنیم . واضح است که مقصود از ضابطه توزیع احتمال ، قاعده و قانونی است که از طریق آن بتوان مقادیر احتمال مربوط به‌همه حوادث ممکن‌الوقوع در یک‌آزمایش (یا چند آزمایش مشابه) را محاسبه نمود ، همانطور که بر اساس رابطه فوق می‌توانستیم قد هریک از ده‌نفر موردبحث درمثال‌مربوطه را محاسبه‌کنیم . معمولاً" به‌دست‌آوردن ضابطه توزیع احتمال بخصوص ضابطه‌ای که بتوان براساس آن احتمال مربوط به‌کلیه حوادث ممکن‌الوقوع در n آزمایش را به‌دست آورد کار آسانی نیست و در بسیاری موارد کار نسبتا" پیچیده و دشواری می‌باشد .

به‌هرحال وقتی می‌توان از توزیع احتمال در محاسبه احتمال استفاده نمود که ضابطه توزیع به‌صورتی‌که تعریف گردید دراختیار بوده و معمولاً"جهت سهولت محاسبات‌این‌ضابطه باید به‌صورت ریاضی بیان شده باشد . البته معمولاً" در حین یک کارآماری ، محقق محتاج

آن نیست که ضابطه توزیع را برای آزمایش مورد نظرش شخصا" به‌دست آورد زیرا این مهم که یکی از مشکل‌ترین قسمتهای کارآماری می‌باشد برای بسیاری از آزمایشهای مختلفی که‌ممکن است انجام گیرند قبلا"انجام گردیده و ضوابط توزیع‌احتمال مربوط به‌این آزمایشهای‌مختلف به‌صورت روابط ریاضی گوناگونی تهیه شده و به‌همراه محاسبات مورد لزوم در اختیار محقق قرار دارد و از این جهت جایی برای نگرانی نیست . اگر به‌کتب آماری مختلف مراجعه شود ملاحظه می‌گردد که در این کتب هیچ اسمی از ضابطه توزیع احتمال برده نشده و از همین مفهوم با عناوین دیگری یاد شده‌است . ما نیزدر اینجا از این لفظ صرفا" به‌منظور فهم بهتر مطلب استفاده نمودیم و از این پس از همان اصطلاحات رایج استفاده خواهیم کرد .

همانطوری که گفته شد برای سهولت محاسبات باید ضابطه توزیع احتمال را به صورت ریاضی بیان نمود .

" اگر متغیر مورد مطالعه متغیری گسسته باشد بیان ریاضی ضابطه "
" توزیع احتمال را قانون توزیع احتمال گویند . به عبارت‌بهتر برای "
" یک متغیر گسسته به رابطه ریاضی که بر اساس آن می‌توان احتمال "
" مربوط به هریک از حوادث ممکن‌الوقوع در n آزمایش را محاسبه "
" نمود قانون توزیع احتمال گویندکه درآن n می‌تواندهرمقطدار مثبتی "
" را اختیار کند . این رابطه را معمولا" با P_x نشان می‌دهند "

تعریف فوق که به‌متغیر گسسته اختصاص دارد با مختصر تفاوتی شامل متغیر پیوسته هم می‌گردد و آن تفاوت این است که اگر متغیر مورد مطالعه پیوسته باشد رابطه ریاضی که‌اکنون تعریف‌گردید به‌جای قانون توزیع‌احتمال تابع چگالی احتمال نامیده شده و با $\varphi(x)$ نشان داده می‌شود بنابراین برای یک متغیر پیوسته به‌رابطه ریاضی که براساس آن بتوان احتمال مربوط به‌همه حوادث ممکن‌الوقوع در یک یا چند آزمایش را محاسبه نمود تابع چگالی احتمال گفته می‌شود .

به‌عنوان مثال بیان ریاضی ضابطه توزیع احتمال و به‌عبارت‌بهترقانون توزیع احتمال درآزمایش n بار پرتاب یک تاس همتراز ، وقتی که آمدن عدد ۳ مطلوب ما باشد و سایر پیشامدها برای مانا مطلوب تلقی گردد رابطه زیر می‌باشد .

$$P_x = \binom{n}{x} p^x q^{n-x}$$

در این رابطه (n) تعداد آزمایشها ، x تعداد وقوع حادثه مورد نظر یعنی آمدن عدد ۳ می‌باشدکه متغیر ما در این آزمایش بوده و متغیری گسسته است ، p احتمال وقوع حادثه‌مورد

نظر در یک آزمایش و q احتمال عدم وقوع حادثه مورد نظر در یک آزمایش می‌باشد . اگر فرض کنیم ($n = ۵$) باشد در آن صورت آزمایش ما ۶ حادثه ممکن‌الوقوع خواهد داشت که عبارتند از ($x=۰$ ، $x=۱$ ، ۰۰۰۰ ، $x=۵$) که احتمال مربوط به‌هریک از این ۶ حادثه از رابطه فوق قابل محاسبه می‌باشد .

مثال ۱ ـ ۳ : احتمال آنکه در ۵ بار آزمایش فوق ۳ بار عدد ۳ بیاید چقدر است ؟

پاسخ : اطلاعات زیر در اختیار است .

$$x = ۳ ، \quad P = \frac{۱}{۶} ، \quad q = \frac{۵}{۶} \quad و \quad n = ۵$$

این اعداد را در رابطه فوق قرار داده و با حل معادله جواب مسئله را به‌دست می‌آوریم .

$$P_{(۳)} = \binom{۵}{۳} \left(\frac{۱}{۶}\right)^{۳} \left(\frac{۵}{۶}\right)^{۲} = (۱۰)\,(۲۵)\left(\frac{۱}{۶}\right)^{۵} = \%۳۲$$

همانند مثال (۱ـ۳) می‌توان احتمال مربوط به ۵ حادثه ممکن‌الوقوع دیگر را نیز محاسبه نموده و جدول توزیع احتمال را برای آزمایش فوق مشخص کرد .

آنچه تا کنون گفته شد مربوط به‌چگونگی محاسبه احتمال نقطه‌ای با استفاده از قانون توزیع احتمال یا تابع چگالی احتمال بود . یعنی تا کنون چگونگی محاسبه احتمال هریک از نقاط موجود در فضای نمونه مربوط به‌یک آزمایش با استفاده از توزیع احتمال بیان گردید اما در بسیاری از موارد محاسبه احتمال نقطه‌ای مطلوب نیست بلکه محقق به دلایل مختلفی به‌محاسبه احتمال فاصله‌ای احتیاج دارد . همان طور که در مقدمه این فصل گفته شد محاسبه احتمال فاصله‌ای برای یک متغیر پیوسته به‌روش مستقیم و بدون استفاده از توزیع احتمال (تابع چگالی احتمال) امری غیر ممکن است زیرا در فاصله مورد نظر بینهایت نقطه وجود دارد که باید احتمال هریک از این نقاط محاسبه شده و با احتمال سایر نقاط جمع شود تا احتمال فاصله‌ای محاسبه شود و بسیار واضح است که محاسبه احتمال بینهایت نقطه امری غیر ممکن است . البته اگر متغیر گسسته باشد قانون توزیع احتمال فی‌نفسه کمک چندان زیادی به‌محاسبه احتمال فاصله‌ای نمی‌کند و کمک اصلی آن در این موارد فقط همان محاسبه احتمال نقطه‌ای می‌باشد بدین معنی که راه حل جدید و کوتاه‌تری را پیش پای فرد نمی‌گذارد اما به‌دلیل آنکه امکان استفاده از محاسبات انجام شده توسط دیگران را چنانکه بعداً خواهیم دید فراهم می‌آورد ، باز هم کمک بسیار بزرگی محسوب می‌گردد . به‌هرحال ذیلاً محاسبه احتمال فاصله‌ای برای هر دو نوع متغیر مورد بحث قرار می‌گیرد .

الف ــ ۲ ــ چگونگی محاسبه احتمال فاصله‌ای برای متغیر گسسته

اگر در یک آزمایش متغیری‌گسسته مورد مطالعه باشد معمولا" احتمال‌نقطه‌ای احتیاجات محقق را تأمین می‌کند و برای متغیرهای گسسته احتمال فاصله‌ای چندان مورد احتیاج‌نیست اما در صورتی که محاسبه چنین احتمالی ضرورت پیدا کند ، چنانکه گفته شداستفاده‌ازقانون توزیع احتمال در اشکال شناخته شده‌ای که محاسبات آنها قبلا" انجام گرفته و بـــه صورت جداولی دراختیاراست مارا از بسیاری از محاسبات بی‌نیاز می‌کند . به‌عنوان مثـــال مقادیـــر مختلف P_x برای قانون توزیعی که در مثال (۱ــ۳) مورد استفاده قرار گرفت برای مقادیر مختلف(P, n, x)درجدولی‌که‌ضمیمهٔ همین‌کتاب می‌باشد محاسبه گردیـــده است و محقـــق می‌تواند جواب همان مثال را بدون انجام هیچ گونه محاسبه‌ای از این جداول بـــدست آورد بهمین ترتیب اگر یک احتمال فاصله‌ای مطلوب مسئله باشد کار را بسیار آسان می‌کنند .

مثال ۲ ــ ۳: احتمال آنکه درآزمایش ۵ بار پرتاب یک سکه دوبار یا سه بار و یا ۴ بار شیـــر بیاید چقدر است؟

$$P(۲U۳U۴)=P_{(۲ \text{ یا } ۳ \text{ یا } ۴)}= P_{(۲)}+P_{(۳)}+P_{(۴)}$$

برای حل مسئله باید سه احتمال $P_{(۲)}$و$P_{(۳)}$و $P_{(۴)}$ به‌طور جداگانه محاسبه شوند . ایـــن سه احتمال را نیز می‌توان از همان قانون توزیع مورد استفاده در مثال (۱ــ۳) محاسبه نمـــود . در اینجا $n=۵$ ، $p=\frac{۱}{۲}$ ، $q=\frac{۱}{۲}$ و x سه مقدار متفاوت را اختیار می‌کند . اما بااستفاده از جدول مقدار این سه احتمال را می‌توان به‌راحتی به‌دست آورد . این مقادیر به‌صورت‌زیر می‌باشند .

$$P_{(۲)}=۰/۳۱۲۵ \quad ، \quad P_{(۳)}=۰/۳۱۲۵ \quad ، \quad P_{(۴)}=۰/۱۵۶۳$$

و در نتیجه احتمال مورد نظر مقدار زیر را خواهد داشت .

$$P(۲ U ۳ U ۴) = ۰/۳۱۲۵ + ۰/۳۱۲۵ + ۰/۱۵۶۳ = ۰/۷۸۱۳$$

برای همین توزیع جدول دیگری تهیه شده است که استفاده از آن حل مثال فـــوق‌را بازهم آسانتر می‌کند . در این جدول مقادیر احتمال تجمعی برای‌هریک‌ازمقادیر x محاسبه شده است . بدین ترتیب که برای مثلا" $(x=۳)$ مقدار احتمال $[P(۰U۱U۲U۳)]$ و

برای ۵ = x مقدار احتمال [$P(0 \cup 1 \cup 2 \cup 3 \cup 4 \cup 5)$] محاسبه گردیده است. واضح است که مثال (۲-۳) را از رابطه زیر به‌صورت خیلی ساده‌تری نیز می‌توان حل کرد.

$$P(2 \cup 3 \cup 4) = P(0 \cup 1 \cup 2 \cup 3 \cup 4) - P(0 \cup 1)$$

مقدار هریک از دو جمله سمت راست در جدول به‌قرار زیر داده شده است.

$$\sum_{x=0}^{4} P_x = P(0 \cup 1 \cup 2 \cup 3 \cup 4) = 0/9688 \quad \text{و} \quad \sum_{x=0}^{1} P_x = P(0 \cup 1) = 0/1875$$

با قرار دادن این دو مقدار احتمال مورد نظر به‌دست می‌آید.

$$\sum_{x=2}^{4} P_x = P(2 \cup 3 \cup 4) = 0/9688 - 0/1875 = 0/7813$$

همان‌طور که ملاحظه می‌شود جواب از هر دو روش یکسان می‌باشد. این مثال به‌خوبی نشان داد که استفاده از قوانین توزیعی که جداول آنها تهیه شده است محاسبه احتمال نقطه‌ای و فاصله‌ای را بسیار آسان می‌کند، نکته‌شایان‌توجه در اینجا این است که در روش اول محاسبه احتمال فاصله‌ای اگرچه به‌دلیل استفاده از جداول در تعیین مقادیر احتمال نقطه‌ای بسیاری از محاسبات حذف گردیدند اما محاسبه مجموع احتمالات نقطه‌ای بازهم کار زیادی طلب می‌نمود. اما روش دوم که بسیار ساده‌تر است (مخصوصا" اگر فاصله‌موردنظر نقاط بسیار زیادی را شامل شود) با استفاده از مقادیر احتمال تجمعی در جدول و حذف محاسبات خسته‌کننده و وقت‌گیر، حل مسئله را به‌یک راحت‌الحلقوم بسیار شیرین تبدیل می‌نمود. واضح است که استفاده از این روش فقط وقتی مفید و مقرون به‌صرفه است که از قوانین توزیعی استفاده شود که جداول توزیع احتمال تجمعی آنها وجود داشته باشد. با توجه به‌مطالب فوق می‌توان گفت که وجود این روش بسیار ساده برای محاسبه احتمال فاصله‌ای یکی از فواید مهم موجود قانون توزیع احتمال است.

الف - ۳ - محاسبه احتمال فاصله‌ای برای یک متغیر پیوسته

قبلا" گفته شد که در متغیرهای پیوسته معمولا" مقدار احتمال نقطه‌ای مطلوبیت چندانی ندارد زیرا چنین احتمالی فقط در مورد یکی از بینهایت حوادثی که ممکن است در آزمایش اتفاق بیفتد صحبت می‌کند و مقدار آن به‌اندازه‌ای کوچک و ناچیز است که معمولا" مساوی صفر فرض می‌شود و واضح است که احتمال وقوع چنین حادثه جزئی در راستای یک کار آماری

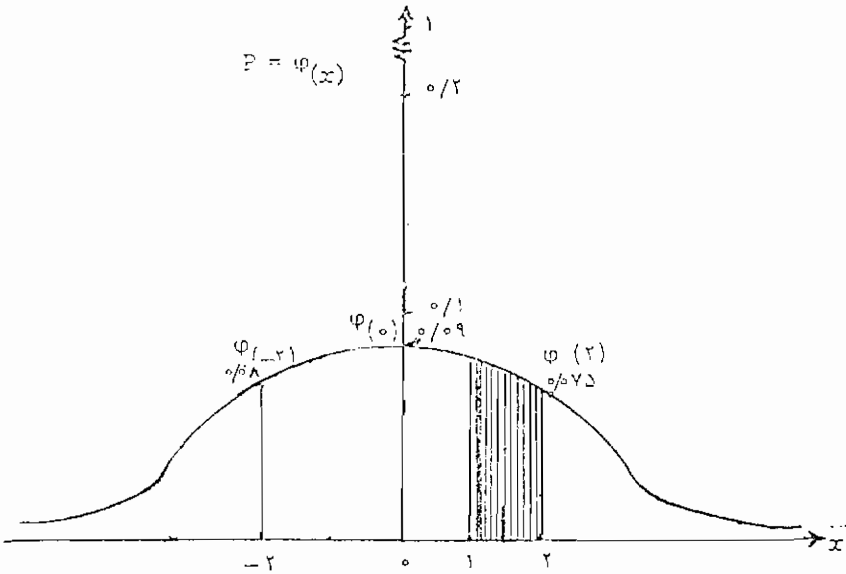
کمتر به‌کار می‌آید و به‌همین دلیل هم احتمال نقطه‌ای برای یک متغیر پیوسته مطلوب کسی نیست و در عمل در متغیر پیوسته از احتمال فاصله‌ای استفاده می‌شود زیرا مقادیر مختلفی که برای حوادث مربوط به‌این احتمال به‌دست می‌آید در خور اعتنا بوده و می‌توانند در محاسبات آماری‌نقشی ایفا کنند. نیاز به‌محاسبه احتمال‌فاصله‌ای برای متغیر پیوسته، ضرورت استفاده از توزیع احتمال را دو چندان و مضاعف می‌کند. چون در اینجا بدون استفاده‌از توزیع احتمال، نه‌تنها محاسبه‌احتمال نقطه‌ای ممکن نیست حتی در صورت عدم وجود این مشکل‌باید‌احتمال هریک از بینهایت نقطه‌ای که در فاصله مورد نظر قرار دارند نیز محاسبه شده و با احتمال سایر نقاط موجود در آن فاصله جمع گردد تا احتمال فاصله‌ای موردنظر به‌دست آید و واضح است که صرفنظر از مشکل اول این کار نیز غیر ممکن بوده و مشکل دومی به‌حساب می‌آید.

اکنون پس از بیان این مقدمات ببینیم چگونه با استفاده از توزیع احتمال، احتمال فاصله‌ای برای یک متغیرپیوسته محاسبه‌می‌گردد. یادآوری می‌شود که چگونگی محاسبه‌احتمال نقطه‌ای برای یک متغیر پیوسته در قسمت قبل گفته شد و دانستیم که‌برای این کار باید خصوصیات نقطه موردنظر را در تابع چگالی احتمال قرار دهیم و احتمال مورد نظرراحساب کنیم و انشاءا... در مبحث توزیع نرمال راه بسیار ساده‌تری نیز برای این کار ارائه خواهد شد. مقدمتا" برای بیان چگونگی محاسبه احتمال فاصله‌ای متغیر پیوسته باید به‌نکته‌ای‌اشاره شود و آن اینکه تابع چگالی احتمال، یک رابطه ریاضی و به‌عبارت دقیقتر یک تابع ریاضی می‌باشد. واضح است که چنین‌تابعی را می‌توان از طریق روشهای ترسیم منحنی، به‌صورت هندسی نشان داد. اکنون فرض کنیم که منحنی یک تابع چگالی احتمال که مربوط به متغیر پیوسته (X) است به‌شکل زیر ترسیم شده باشد.

در این شکل به‌عنوان مثال مقادیر احتمال برای سه‌مقدار متفاوت متغیر (X) یعنی ($x=۲$، $x=۰$ و $x=-۲$) به‌ترتیب برابر مقادیر فرضی و تقریبی ($P_{(۲)}=\varphi_{(۲)}=\%۷۵$، $P_{(۰)}=\%۰۹$) و $P_{(-۲)}=\%۸$ به‌صورت سه خط عمودی نشان داده شده است.

با کمی دقت در می‌یابیم که احتمال آنکه متغیر (X) با تابع چگالی احتمال$\varphi(x)$که منحنی آن در شکل (۱-۳) رسم شده است، بین دو نقطه (۲ و ۱) مقدار خود را در یک آزمایش بطور تصادفی اختیار کند برابر سطح زیر منحنی$\varphi(x)$در فاصله ($x=۲$ و $x=۱$) می‌باشد. واضح است که با محاسبه انتگرال$\varphi(x)$در فاصله فوق سطح زیر منحنی در این فاصله به‌دست می‌آید.

$$P(۱\leqslant x\leqslant ۲)= \text{سطح زیر منحنی } \varphi(x) \text{ درفاصله } (x=۱ \text{ و } x=۲) = \int_{۱}^{۲}\varphi(x)\,d_x$$



شکل (۲-۳)

دلیل این امر نیز بسیار واضح است اگر به‌فرض محال محاسبه مقدار $\varphi(x)$، برای بی‌نهایت مقدار متغیر (X) در فاصله ($x = 1$) تا ($x = 2$) ممکن باشد و این مقادیر یک به‌یک حساب شده و هرکدام به‌صورت یک نمودار میله‌ای روی صفحه رسم شوند[این کار برای تعداد محدودی از مقادیر (X) در شکل (۳-۱) انجام شده است]. بی‌نهایت خط عمودی چسبیده به‌هم به‌دست می‌آید که مجموع آنها سطح زیر منحنی $\varphi(x)$ را در فاصله فوق تشکیل می‌دهد و به‌همین دلیل در بالا گفته شد که سطح زیر منحنی در فاصله ($1 \leqslant x \leqslant 2$) برابر احتمال مورد نظر می‌باشد. بیان کلی رابطه فوق برای متغیر تصادفی پیوسته (X) بصورت زیر است.

$$P(a \leqslant x \leqslant b) = \int_a^b \varphi(x)dx \qquad (۳-۱)$$

قبلا" گفته شد که برای یک متغیر تصادفی پیوسته، احتمال نقطه‌ای معمولا" مطلوب نیست و احتمال فاصله‌ای در محاسبات آماری بیشتر به‌کار می‌آید. نتیجه عملی این

امر با توجه به‌مطالب فوق‌الذکر این است که وقتی متغیر مورد مطالعه، متغیری‌پیوسته باشد خود تابع چگالی احتمال مورد استفاده چندانی در محاسبات آماری نداشته‌بلکه تابع اولیه یا انتگرال آن‌مورد استفاده قرارمی‌گیرد و به‌همین دلیل در آمار به‌انتگرال تابع چگالی به‌مراتب بیش از خود آن اهمیت داده شده و از این مفهوم‌با عنوان تابع‌توزیع احتمال یاد می‌شود و برای متغیر پیوسته (X) به‌صورت F_x نشان داده می‌شود.

" تابع توزیع F_x که تابع اولیه یا انتگرال تابع چگالی $\varphi(x)$ است "
" تابعی است که‌سطح‌زیر منحنی $\varphi(x)$ را برای همه نقاط ($X < x$) "
" نشان می‌دهد. به‌عبارت دیگر مقدار تابع توزیع برای یک مقدار خاص از "
" متغیر تصادفی (X) نشان دهنده احتمال زیر است $F_{(x)}=P(X<x)$ "

به‌عنوان مثال برای تابع چگالی احتمالی که در شکل (۱ـ۳) منحنی آن ترسیم شده است مقدار تابع توزیع $F(-۲)$ نشان دهنده سطح زیر منحنی در سمت چپ نقطه ($x=-۲$) و مقدار $F(۱)$ نشان دهنده‌سطح زیر منحنی درسمت چپ نقطه ($x=۱$)و نیز مقدار $F(۲)$ نشان دهنده سطح زیر منحنی در سمت چپ نقطه ($x=۲$) می‌باشد. واضح است که خواهیم داشت:

$$F(۲) > F(۱) > F(-۲)$$

و چون سطح زیرمنحنی در هرفاصله عبارت‌از مقداراحتمال این حادثه‌است که متغیرتصادفی (X) مقدار خود را درآن فاصله اختیار کند، عبارت زیر را خواهیم داشت.

$$F(۲) = P(X<۲) \quad , \quad F(۱) = P(X<۱) \quad , \quad F(-۲)=P(X<-۲)$$

با استفاده از تابع توزیع می‌توان مقدار احتمال فاصله‌ای را به‌سادگی محاسبه نمود. مجدداً فرض کنید می‌خواهیم احتمال آنکه متغیر تصادفی (X) را که تابع چگالی آن در شکل (۱ـ۳) ترسیم شده است هنگامی که در فاصله($۱ \leqslant X \leqslant ۲$)مقداراختیارکند به‌دست آوریم. می‌توانیم از رابطه زیر برای این کار استفاده کنیم.

$$P(۱<X<۲) = P(X \leqslant ۲)-P(X \leqslant ۱)$$

اما می‌دانیم که دو عبارت سمت راست به‌ترتیب از راست به‌چپ‌مساوی $F(۱)$ و $F(۲)$ می‌باشد و لذا رابطه فوق به‌رابطه زیر تبدیل می‌شود.

$$P(۱<X<۲) = F(۲) - F(۱)$$

رابطه فوق را می‌توان در حالت کلی به‌صورت زیر بیان نمود .

$$P(a < X < b) = F(b) - F(a) \qquad (۲-۳)$$

مسلما" آنچه تا کنون در مورد توزیع احتمال و کاربرد آن در محاسبه مقدار احتمال نقطه‌ای و فاصله‌ای در مورد متغیرهای گسسته و پیوسته گفته شد مشکلی را در ذهن تداعی می‌کند و آن اینست که توزیع احتمال اگرچه در بسیاری موارد از نظر کلی و اصولی حلال بسیاری از مشکلات است اما استفاده از آن به‌صورتی که گفته شد هنوز مشکلات زیادی به‌همراه دارد که معمولا" از حوصله یک کارآماری خارج است ، اولا" در جهان ، بینهایت آزمایش با بینهایت شکل توزیع مختلف وجود دارد که طبعا" هریک از این اشکال قانون توزیع احتمال یا تابع چگالی احتمال مخصوص به‌خود و متفاوت با دیگراشکال توزیع داشته و نتیجتا" بینهایت قانون توزیع احتمال یا تابع چگالی احتمال مختلف برای این آزمایشها وجود خواهد داشت و محقق برای هر آزمایشی که می‌خواهد انجام دهد باید ابتدا قانون توزیع یا تابع چگالی احتمال آن را به‌دست آورد و این کاری بس سنگین و طاقت فرسا می‌باشد و ثانیا" پس از آنکه قانون توزیع یا تابع چگالی احتمال مورد نظر مشخص گردید باید برای احتمال هر حادثه‌ای محاسبات نسبتا" پیچیده‌ای ازقبیل حل توابع مختلف ریاضی و یا انتگرال‌گیری و غیره انجام گیرد که این نیز کاری بسیار مشکل است .

در پاسخ باید گفت که هیچیک از دو اشکال فوق چندان نگران کننده به‌نظر نمی‌رسند زیرا اگر چه بینهایت شکل توزیع و نتیجتا" بینهایت قانون توزیع و تابع چگالی احتمال وجود دارند اما چنانکه در قسمتهای بعدی خواهیم دید روشهایی وجود دارند که می‌توان با استفاده از آنها اشکال توزیع مختلف را به‌یکدیگر تبدیل نموده و نهایتا" همه آنها رادرچند شکل توزیع مهمتر و پرمصرف‌تر خلاصه نمودکه این چند شکل توزیع مهم درآمار موردبررسی قرارگرفته و قانون توزیع و یا تابع چگالی احتمال و تابع توزیع احتمال آنها نیزمشخص شده است و آزمایش مورد نظر آمارگر هرچه باشد او می‌تواند برای محاسبات خود از همین چند شکل مهم توزیع استفاده کند .

ثانیا" اکثریت‌قریب به‌اتفاق محاسباتی که ممکن است آمارگر در راستای کارآماری خود به‌نحوی محتاج آنها باشد برای این چند شکل توزیع مهم محاسبه شده و در جداول‌مختلفی که بعضا" ضمیمه این کتاب شده‌اند ، ارائه شده است و آمارگر می‌تواند با انتخاب یکی ازاین چند شکل مهم توزیع احتمال که با هدف‌او تناسب بیشتری دارد ، از این جداول استفاده نموده و محاسبات خود را به‌راحتی انجام دهد . در قسمت بعدی این اشکال توزیع در حد ضرورت مورد بحث قرار می‌گیرند .

اگر خاطرتان باشد قبلاً مباحث این فصل را به‌دو بخش محاسبه احتمال با استفاده‌از توزیع احتمال و نیز توزیعهای احتمال مهم تقسیم نمودیم. بحث اول یعنی چگونگی محاسبه احتمال در اینجا به‌پایان می‌رسد و از این پس توزیعهای مهم احتمال مورد بحث و بررسی قرار می‌گیرند.

ب ـ توزیعهای مهم احتمال

همانطور که در مقدمه این فصل گفته شد، در فصل سوم فقط توزیعهای احتمال مهمی مورد بحث قرار می‌گیرند که در آمار از آنها به‌منظور محاسبه احتمال استفاده می‌شود. در همانجا گفته شد که این توزیعهای احتمال را می‌توان به‌دو دسته کلی تقسیم نمود. یکی از این دو دسته،توزیعهائی هستند که متغیر مورد بحث آنها گسسته بوده و دومین گروه‌آنهائی هستند که در مورد متغیر پیوسته کاربرد دارند که به‌همین ترتیب نیز در اینجا مورد بررسی قرار می‌گیرند.

ب ـ ۱ـ توزیعهای احتمال گسسته

این توزیعها چنان که از اسمشان پیداست وضعیت احتمالات مربوط به متغیرهای جدا و گسسته‌را توضیح می‌دهند. مهمترین توزیعهای این گروه توزیع دو جمله‌ای[۱]، توزیع پواسن[۲]، توزیع فوق هندسی[۳] می‌باشند که ذیلاً مورد بررسی قرار می‌گیرند.

ب ـ ۱.۱. توزیع دوجمله‌ای

این توزیع که از آن با عناوین توزیع دو نقطه‌ای و برنولی نیز یاد می‌شود مهمترین و پرکاربردترین توزیع گسسته است. وجه تسمیه توزیع دوجمله‌ای یا دو نقطه‌ای آن است‌که این توزیع فقط در مورد آزمایشهایی به‌کار می‌رود که فضای نمونه آنها در یک مرتبه آزمایش فقط دو نقطه داشته باشد. به‌عبارت دیگر در یک بار آزمایش فقط دو حادثه امکان وقوع داشته باشند. به‌عنوان مثال آزمایش پرتاب یک سکه چنین حالتی دارد و فقط دو نتیجه‌شیر یا خط ممکن است‌ازآن عاید گردد. البته این مسئله دامنهٔ استفاده ازاین توزیع را محدود نمی‌کند زیرا به‌راحتی‌می‌توان بسیاری از آزمایشهائی‌را که درحالت طبیعی چندین حالت‌ممکن‌الوقوع

1- Binomial Distribution

2- Poisson Distribution

3- Hypergeometric Distribution

دارند به‌آزمایشهائی‌تبدیل‌نمود که‌فقط دوحادثه ممکن‌الوقوع‌داشته باشند وبدین‌ترتیب‌این آزمایشهارادردایره‌شمول‌توزیع دوجمله‌ای‌قرارداد. مثلا" درآزمایش پرتاب یک تاس در حالت طبیعی ۶حالت‌امکان‌وقوع دارند. امامعمولا" از بین‌تمام‌حالات‌ممکن‌الوقوع فقط یک‌حالت مثل وقوع عدد ۴ موردنظر آزمایش کننده می‌باشد. در اینجا او می‌تواند تمام نتایج ممکن‌الوقوع آزمایش خودرا به‌دوحادثه‌مطلوب(آمدن عدد ۴) و نامطلوب (آمدن عدد غیر۴)تقسیم نماید و با این طبقه‌بندی‌جدیدازقانون توزیع برنولی(دوجمله‌ای)برای آزمایش خوداستفاده کند.

اگرحادثه‌مطلوب را با A نشان‌دهیم‌حادثه نامطلوب یا غیر A را دربسیاری ازنوشته‌ها با ($\bar{A}$) نشان می‌دهند که براساس این‌شکل نمایش حوادث دوحادثه ممکن‌الوقوع درآزمایش پرتاب تاس(۴ و $\bar{۴}$) می‌باشند.

چنان که قبلا" گفته شد هر توزیع احتمال سه مشخصه اصلی دارد که عبارت از شکل کلی توزیع، میانگین و انحراف معیار آن می‌باشند. معمولا" بهترین معرف برای شکل کلی‌یک توزیع احتمال، ضابطه آن توزیع می‌باشد. یعنی بهترین معرف برای شکل کلی توزیع‌احتمال برای یک متغیر گسسته قانون توزیع آن و برای یک متغیر پیوسته تابع چگالی آن‌می‌باشد زیرا از طریق آنها شکل کلی توزیع به‌خوبی مشخص می‌گردد. در متغیرهای پیوسته به‌دلیل‌وجود محاسبات انتگرالی و غیره همان‌طور که قبلا" گفته‌شد به‌جای تابع‌چگالی احتمال، تابع‌اولیه آن یعنی تابع توزیع احتمال را معرفی‌می‌کنندوبعلاوه برای فهم بهتر مطلب شکل کلی توزیع را به‌صورت هندسی و از طریق رسم منحنی نیز توضیح می‌دهند‌مقدارمیانگین‌وانحراف‌معیار یک توزیع نیز با استفاده از همان روابطی که در فصل دوم گفته شد محاسبه می‌گردد. البته در بعضی موارد نیز این روشها به‌اقتضای ضرورت، تغییراتی پیدا می‌کنند که در جای خود به‌آنها اشاره می‌شود لذا در بررسی اشکال مهم توزیع روال کار بدین‌صورت خواهد بود که ابتدا ضابطه ریاضی توزیع (قانون توزیع یا تابع چگالی احتمال و تابع توزیع) مورد بحث قرار گرفته و در صورتی که روشهای محاسبه میانگین و انحراف معیار در موردی دستخوش تغییر شده باشند به‌این روشها نیز اشاره می‌شود و در نهایت در صورتی که توضیحات‌دیگری لازم بود آنها نیز به‌بحث اضافه خواهند شد. بنابراین بحث توزیع دوجمله‌ای با قانون‌این توزیع باید دنبال شود.

اگر حادثه مطلوب را در یک آزمایش که فقط دو حادثه ممکن‌الوقوع دارد، با (A) و حادثه نامطلوب را با ($\bar{A}$) نشان دهیم قانون توزیع دوجمله‌ای وقتی مورد استفاده قرار می‌گیرد که بخواهیم بدانیم در n بار آزمایش چندبار حادثه مطلوب (A) رخ می‌دهد[1] از

۱ ــ در واقع قانون توزیع دوجمله‌ای روشی برای محاسبه احتمال، در حوادث دو یا
←

آنجا که در یک بار آزمایش فقط دو حادثه ممکن‌الوقوع دارد ، داریم :

$$P_A + P_{\bar{A}} = 1$$

که این رابطه را به‌صورت زیر نیز می‌توان نوشت .

$$P_{\bar{A}} = 1 - P_A$$

معمولا" جهت اختصار در قانون توزیع دوجمله‌ای (P_A)را با p و ($P_{\bar{A}}$) را با q نشان می‌دهند . بنابراین در قانون توزیع برنولی (p) نشان دهنده احتمال وقوع حادثه مطلوب در یک بار آزمایش و (q) احتمال عدم وقوع حادثه مطلوب در یک بار آزمایش می‌باشد .با این توضیحات احتمال‌آنکه در (n) بار آزمایش (x) مرتبه‌حادثه دلخواه (A) وقوع پیدا کند از رابطه (۳ــ۳) به‌دست می‌آید که آن را قانون توزیع دو جمله‌ای یا برنولی می‌نامند .

$$b(x,n,P) = \binom{n}{x} P^x q^{n-x} \qquad (۳-۳)$$

مقدار احتمال توزیع دوجمله‌ای برای مقادیر مختلف (x و n و P) محاسبه شده و در جدولی که به‌آخر همین نوشته ضمیمه شده ، نشان داده شده است . همچنین مقدار احتمال تجمعی برای تعدادی از مقادیر مختلف (x و n و P) محاسبه شده و در جدول دیگری که‌آن هم ضمیمه است نشان داده شده‌است ولذا درمسائل‌مختلف بدون‌استفاده از رابطه (۳ــ۳) می‌توان جوابهارااز‌جدول‌بدست‌آورد .

مثال ۴ ــ۳ : سکه‌ای همتراز را ۸ مرتبه پرتاب می‌کنیم . مطلوبست احتمال آنکه :

الف ــ ۵ مرتبه شیر بیاید .

ب ــ حداکثر ۵ مرتبه شیر بیاید .

ج ــ حداقل ۵ مرتبه شیر بیاید .

←

چند مرحله‌ای می‌باشد . واضح است که در بسیاری موارد می‌توان مثلا" بجای n بار پرتاب یک سکه ، n سکه‌را در یک بار پرتاب‌نمودکه‌نتیجه‌یکسان بوده و قانون برنولی درهردوشکل این آزمایش می‌تواند مورد استفاده قرار گیرد .

پاسخ: مسئله را با استفاده از رابطه (۳ـ۳) حل می‌کنیم .

الف

$$P(x=۵) = b(۵ ، ۸ ، \frac{۱}{۲}) = \binom{۸}{۵}(\frac{۱}{۲})^{۵}(\frac{۱}{۲})^{۳} = (۵۶)(\frac{۱}{۲})^{۸} = ۰/۲۱۸۸$$

ب :

$$P(x \leqslant ۵) = P(۰) + P(۱) + P(۲) + \cdots\cdots + P(۵)$$

$$P(۰) = b(۰ ، ۸ ، \tfrac{۱}{۲}) \text{ و } P(۱) = b(۱ ، ۸ ، \tfrac{۱}{۲}) \text{ و}$$

$$P(۲) = b(۲ ، ۸ ، \tfrac{۱}{۲}) \cdots\cdots\cdots \text{ و } P(۵) = b(۵ ، ۸ ، \tfrac{۱}{۲})$$

مقدار $P(۵)$ را در بند الف محاسبه نمودیم و سایر مقادیر خواسته شده نیز محاسبه شده‌اند و نتایج به‌قرار زیر است .

$$P(۰)=۰/۰۰۳۹ \;,\; P(۱)=۰/۰۳۱۳۱ ,\; P(۲) = ۰/۱۰۹۴ \;,\; P(۳)=۰/۲۱۸۸ \;,\; P(۴)=۰/۲۷۳۴$$

با جمع‌نمودن این مقادیر ، احتمال مورد نظر به‌دست می‌آید .

$$P(x \leqslant ۵) = ۰/۰۰۳۹+۰/۰۳۱۳+۰/۱۰۹۴+۰/۲۱۸۸+۰/۲۷۳۴+۰/۲۱۸۸=۰/۸۵۵۵$$

ج ــ برای حل این بند احتیاجی به‌روش مستقیم نیست و می‌توان مسئله را از روش ساده‌تر زیر حل نمود . می‌دانیم که مجموع احتمالات در یک آزمایش برابر ۱ می‌باشد .

$$P(x \leqslant ۵) + P(x > ۵) = ۱$$

اما مطلوب بند ج آمدن حداقل ۵ مرتبه شیر است یعنی داریم :

$$P(x \geqslant ۵) = P(x > ۵) + P(۵) = ۰/۱۴۴۵+۰/۲۱۸۸=۰/۳۶۳۳$$

هر سه جواب بند الف و ب و ج را به‌راحتی می‌توانستیم از جدول به‌دست آوریم .برای بند الف از جدول مقادیر احتمال توزیع دوجمله‌ای و برای بند ب و ج از جدول مقادیراحتمالات تجمعی قانون توزیع دوجمله‌ای باید استفاده نمود .

نکته‌ای که در مورد توزیع دوجمله‌ای قابل ذکر است این است که توزیع دو جمله‌ای در

مورد آزمایشهائی به‌کار می‌رود که مقدار p و q در طول انجام n بار آزمایش ثابت باشد . معنای این حرف این است که توزیع دوجمله‌ای فقط در مورد آزمایشهای مستقل از یکدیگر می‌تواند مورد استفاده قرار گیرد .

از سه مشخصه لازم برای معرفی دقیق یک توزیع احتمال دو جمله‌ای یا دو نقطه‌ای ، مشخصه‌مربوط به‌شکل کلی توزیع در قالب قانون توزیع دونقطه‌ای گفته شد . روابط اصلی برای محاسبه میانگین و انحراف معیار این توزیع همان روابطی هستند که در فصل دوم بحث شد اما به‌اقتضای شرایطی که قانون توزیع دوجمله‌ای ایجاب می‌کند می‌توان آنها را به‌روابط ساده‌تری نیز تبدیل نموده و میانگین و واریانس (انحراف معیار) متغیری را که توزیع احتمال دو جمله‌ای دارد با استفاده از این روابط ساده‌تر محاسبه نمود . دانستن چگونگی تبدیل روابط اصلی محاسبه‌میانگین و انحراف معیار به‌روابطی که اختصاصا" در چهارچوب یک‌توزیع احتمال خاص قابل استفاده است به‌دلیل محاسبات نسبتا" پیچیده ریاضی که به‌همراه دارد، معمولا" برای دانشجویانی که رشته تخصصی آنان آمار نیست ضرورتی ندارد و به‌همین‌دلیل نیز در این‌نوشته فقط به‌اثبات‌رابطه ریاضی محاسبه‌میانگین توزیع دوجمله‌ای به‌عنوان نمونه بسنده نموده و در سایر توزیعها روابط فوق را بدون اثبات ارائه خواهیم کرد .

قضیه ۱ـ۳: میانگین (امید ریاضی) مقادیر متغیری که مقدار احتمال آن با استفاده‌از قانون توزیع دوجمله‌ای محاسبه می‌شود و به عبارت بهتر میانگین توزیع دوجمله‌ای با استفاده از رابطه زیر محاسبه می‌شود .

$$E(X)= \mu = nP \qquad (۳-۴)$$

اثبات : می‌دانیم که رابطه کلی محاسبه امید ریاضی رابطه زیر است .

$$\mu = E(X)=\sum_{x=0}^{n} x\,P(x)$$

اما از رابطه (۳-۳) برای متغیری که قانون توزیع آن دوجمله‌ای است داریم :

$$P(x) = b(x,n,P)= \binom{n}{x} P^{x} q^{(n-x)}$$

بنابراین رابطه اولی را می‌توان به‌صورت زیر نوشت :

$$\mu = E(X) = \sum_{x=0}^{n} x \left[\binom{n}{x} P^{x} q^{(n-x)}\right]$$

مقدار فوق به‌ازاء ($x=0$) معادل صفر است و لذا رابطه فوق به‌صورت زیر نیزنوشته‌می‌شود .

$$\mu=E(X)=\sum_{x=1}^{n} x\left[\binom{n}{x} P^{x} q^{(n-x)}\right]=\sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} P^{x} q^{(n-x)}$$

$$=\sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} P^{x} q^{(n-x)}$$

n و p اعداد ثابتی هستند که می‌توان یکی از آنها را به‌بیرون از (Σ) انتقال داد و لذا داریم .

$$\mu=E(X)=n.P \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)![(n-1)-(x-1)]!} P^{(x-1)} q^{[(n-1)-(x-1)]}$$

جهت سهولت فهم عملیات بعدی فرض کنیم m = ($n-1$) و y = ($x-1$) باشد . داریم.

$$E(X)=nP \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} P^{y} q^{m-y}=nP \sum_{y=0}^{m}\binom{m}{y} P^{y} q^{(m-y)}$$

از دو جمله‌ای خیام ــ نیوتن که در ریاضیات اثبات می‌شود می‌توان در مورد p و q چنین نوشت .

$$(P+q)^{m}=\sum_{y=0}^{m}\binom{m}{y} P^{y} q^{(m-y)}$$

با جایگذاری طرف چپ رابطه فوق در رابطه قبلی داریم :

$$\mu=E(X)=nP(P+q)^{m}$$

اما می‌دانیم که در توزیع دوجمله‌ای ($p+q=1$) می‌باشد و لذا داریم :

$$(P+q)^{m}=(1)^{m}=1$$

و بنابراین در نهایت ، میانگین توزیع دوجمله‌ای به‌صورت زیر یعنی همان رابطه (۳ ــ ۳) قابل محاسبه است .

$$\mu=E(X)=nP$$

به‌همین‌ترتیب می‌توان ثابت‌کرد که واریانس متغیری‌که توزیع دوجمله‌ای‌دارد به‌راحتی با استفاده از قضیه زیر قابل محاسبه است.

قضیه ۲-۳: واریانس متغیری که توزیع دوجمله‌ای دارد و به‌عبارت ساده‌تر واریانس توزیع دوجمله‌ای از رابطه زیر قابل محاسبه است.

$$\sigma^{2}=nPq \qquad (۳-۵)$$

این رابطه نیز انشاءالله در مسائل این فصل اثبات خواهد شد. با ذکر یک مثال در مورد میانگین و انحراف معیار توزیع دوجمله‌ای این بحث را به‌پایان می‌بریم.

مثال ۵-۳: میانگین و انحراف معیار آمدن شیر در آزمایش مثال (۴-۳) را محاسبه کنید.

پاسخ: میانگین از رابطه (۴-۳) و انحراف معیار جذر واریانس است که از رابطه (۵-۳) به‌دست می‌آید.

$$\mu = E(X) = nP = (۸)\left(\tfrac{۱}{۲}\right) = ۴$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{nPq} = \sqrt{(۸)\left(\tfrac{۱}{۲}\right)\left(\tfrac{۱}{۲}\right)} = \sqrt{۲} = ۱/۴$$

ب - ۱-۲. توزیع فوق هندسی

همان طور که قبلاً گفتیم یکی از شرایطی که برای استفاده از توزیع دوجمله‌ای لازم است وجود داشته باشد، مستقل بودن آزمایشهای مختلف از یکدیگر می‌باشد. به عبارت دیگر در توزیع دوجمله‌ای ملاحظه گردید که این توزیع وقتی می‌تواند مورد استفاده قرار گیرد که مقدار (p) در طول آزمایشهای مختلفی که انجام می‌گیرد، ثابت باشد. اما اگر در عمل با آزمایشهای وابسته به‌یکدیگر سروکار داشته باشیم که بالطبع واجد شرط فوق نیستند مسلماً نمی‌توان از قانون توزیع دوجمله‌ای در چنین مواردی استفاده نمود. آزمایشهائی از قبیل انتخاب بدون جایگذاری مهره از درون یک کیسه، انتخاب بدون بازگردانی اشیاء و افراد مختلف دیگر از مجموعه‌هایشان چنین حالتی دارند. وجود چنین حالتی یعنی عدم ثبات مقدار (p) در طول آزمایشها در شرایطی که سایر شرایط حاکم بر توزیع دوجمله‌ای موجود باشند شرط اصلی استفاده از توزیع احتمال دیگری به‌نام توزیع فوق‌هندسی را فراهم می‌آورد. با توجه به‌این مطلب که قانون توزیع دوجمله‌ای فقط وقتی قابل استفاده است که آزمایش ما فقط دو حادثه ممکن‌الوقوع مطلوب (موفقیت) و یا نامطلوب (شکست) داشته باشد، شرط

اصلی فوق به‌طور منطقی شرایط دیگری را به‌دنبال خواهد آورد ، بدین ترتیب که اولا" وقتی می‌توان از تعداد موفقیتها در (n) آزمایش غیر مستقل و وابسته به‌یکدیگر (که نتیجتا" در آنها مقدار p ثابت نیست) صحبت نمود که (n) حجم نمونه‌ای از یک جامعه اصلی به‌حجم (N)باشد و درواقع ما از نمونه‌گیری‌استفاده کرده باشیم در غیراین صورت تعداد موفقیتها امری واضح و بدیهی بوده و هیچ گونه محاسبه‌ای را لازم نخواهد داشت . به‌عنوان مثال‌اگر از یک کیسه حاوی ده مهره که سه تای آن سفید است و انتخاب همین مهره سفید نیزمطلوب ما بوده و موفقیت تلقی شود ، ده مهره را بدون بازگردانی انتخاب کنیم پرسش در مورد تعداد موفقیتها بی‌معنی خواهد بود زیرا دیگر در کیسه ، مهره‌ای باقی نمانده و یقینا" هرسه موفقیت‌ممکن اتفاق خواهد افتاد . ثانیا" تعداد موفقیتها در چنین آزمایشی‌وقتی‌قابل‌محاسبه خواهد بود که تعداد کل موفقیتها و بالنتیجه تعداد کل شکستهای ممکن‌الوقوع ازقبل‌مشخص باشند ، همانند مثال فوق‌که ما از ابتدا می‌دانستیم‌که حداکثر ۳ موفقیت و همچنین حداکثر ۷ شکست در ده انتخاب امکان وقوع دارند . لازمه این امر آن است که از ابتدا تعداد افراد جامعه‌که انتخاب‌آنها برای ما موفقیت تلقی می‌گردد مشخص باشد . به‌همین دلیل درمسائلی که به‌توزیع فوق هندسی مربوط می‌شود گفته می‌شود که مثلا" تعداد افراد جامعه‌ای (N) می‌باشد که انتخاب (K) فرد از این افراد برای ما حادثه مطلوب بوده و موفقیت تلقی می‌گردد و انتخاب $(N - K)$ نفر باقیمانده به‌منزله شکست تلقی خواهد شد .

شرایط لازم‌برای توزیع‌فوق‌هندسی یعنی شرایط فوق را یک‌بار دیگربه‌صورت جمع‌بندی شده و مختصر بیان می‌کنیم . با توجه به‌مطالبی که پیش‌از این ذکر شد ، این شرایط به قرار زیرند :

۱ ـ ازجامعه‌ای به‌حجم (N) نمونه‌ای‌به‌حجم (n) بدون جایگذاری انتخاب می‌شود .

۲ ـ انتخاب (K) فرد از (N) فرد موجود در این جامعه برای ما مطلوب می‌باشد ، یعنی هریک از این (K) فرد که انتخاب گردند برای ما یک موفقیت محسوب می‌شوند و در مقابل‌هریک از $(N - K)$فرد باقیمانده‌که در نمونه انتخاب شوند ازنظر ما یک شکست تلقی خواهد گردید .

اگر خاطرتان باشد قبلا" گفته شد که برای معرفی دقیق یک توزیع احتمال باید شکل کلی‌آن توزیع ، میانگین آن و بالاخره انحراف معیار (واریانس) آن مشخص گردند. همچنین قبلا" گفته شد که مشخص نمودن قانون توزیع در متغیرهای گسسته برای معرفی‌شکل‌کلی‌توزیع کفایت می‌کند ،بنابراین باارائه قانون توزیع فوق هندسی بحث‌پیرامون این توزیع رادنبال می‌کنیم . اگر شرایط دوگانه فوق در مورد آزمایشی صادق باشند یعنی وقتی که از جامعه‌ای به‌حجم (N) ، (K) فرد به‌عنوان موفقیت و$(N - K)$فرد به‌عنوان شکست در نظر گرفته

شده‌باشند وازاین‌جامعه نمونه‌ای به‌حجم (n) انتخاب کنیم ، احتمال آنکه در این نمونه (X) بار موفقیت نصیب ما گردد از رابطه زیر به‌دست می‌آید .

$$P(X)=h(X,N,n,K) = \frac{\binom{K}{X}\binom{N-K}{n-X}}{\binom{N}{n}} \quad , \ X=0,1,2,\ldots\ldots\ldots,n \qquad (۳-۶)$$

روش محاسبه میانگین و واریانس متغیری که توزیع احتمال آن براساس توزیع فوق ـ هندسی‌می‌باشد نیزبه‌اقتضای شرایط موجود دراین‌توزیع به‌روشهای ساده‌تری تبدیل می‌گردد که ذیلاً" این روشهای ساده بدون اثبات ذکر می‌شوند .

قضیه ۳ ـ ۳ : میانگین و واریانس توزیع فوق هندسی به‌ترتیب از روابط (۷ـ۳) (۸ـ۳) محاسبه می‌شوند .

$$\mu = \frac{nK}{N} \qquad (۳-۷)$$

$$\sigma^2=\left(\frac{N-n}{N-1}\right)\left(\frac{nK}{N}\right)\left(1-\frac{K}{N}\right) \qquad (۳-۸)$$

با حل یک مثال در مورد توزیع فوق هندسی ، بحث در باره این توزیع را پایان‌داده و به‌بررسی توزیع احتمال ناپیوسته مهم بعدی یعنی توزیع پواسن می‌پردازیم .

مثال ۶ ـ ۳ : از ۸ نفر افراد شاغل در یک کارگاه ۳ نفر با سواد و بقیه بی‌سواد می‌باشند . اگر از این کارگاه ۵ نفر را به‌طور تصادفی برای استفاده از یک مزیت شغلی انتخاب کنیم .

الف ـ احتمال آنکه ۳ نفر باسواد انتخاب شده باشند چقدر است ؟

ب ـ میانگین و واریانس‌افراد باسواد در این آزمایش چقدر است ؟

پاسخ : شرایط برای استفاده از قانون توزیع فوق هندسی فراهم است و لذا برای انتخاب بند الف از رابطه (۶ـ۳) استفاده می‌کنیم .

الف
$$P(۳) = h(۳،۸،۵،۳) = \frac{\binom{۳}{۳}\binom{۸-۳}{۵-۳}}{\binom{۸}{۵}} = \frac{۱۰}{۵۶}$$

ب ـ از روابط (۷ـ۳) و (۸ ـ ۳) استفاده می‌کنیم .

$$\mu = \frac{nk}{N} = \frac{(۵)(۳)}{۸} = \frac{۱۵}{۸} = ۱/۸۷۵$$

$$\sigma^{۲} = (\frac{N - n}{N - ۱})(\frac{nK}{N})(۱ - \frac{K}{N}) = (\frac{۳}{۷})(۱/۸۷۵)(۱ - \frac{۳}{۸}) = ۰/۵$$

ب ـ ۱ ـ ۲: توزیع پواسن :

در عملیات آماری مواردی پیش می‌آیند که در آنها میانگین و یا متوسط وقوع یک‌حادثه خاص‌که می‌توان آنرا موفقیت نیز تلقی نمود برای آمارگراز ابتدا مشخص بوده‌و او می‌خواهد بداند که احتمال وقوع (X) بار موفقیت چقدر است؟ به‌عنوان مثال محقق‌می‌داند که‌به‌طور متوسط همه‌ساله دبستانها به‌علت بارش برف ۴ روز در زمستان تعطیل است و او می‌خواهد بداند احتمال آنکه امسال (x =۶) روز به‌همین علت مدارس تعطیل گردند چقدر است؟در چنین مواردی از توزیع احتمال گسسته‌ای به‌نام توزیع پواسن می‌توان استفاده نمود.

شکل کلی این توزیع را نیز همانند توزیعهای ناپیوسته قبلی با ارائه قانون توزیع احتمال آن بیان می‌کنیم. اگر در یک آزمایش وقوع موفقیت (X) یک متغیر گسسته بوده و میانگین آن (μ)معلوم باشد احتمال وقوع (x) بار موفقیت در آن آزمایش ازرابطه (۹ـ۳) به‌دست می‌آید که این رابطه همان قانون توزیع پواسن می‌باشد.

$$P(x,\mu) = \frac{e^{-\mu}\mu^{x}}{x!} \qquad (۳-۹)$$

در این رابطه (μ) میانگین وقوع موفقیتها و (e) عددی ثابت بوده و مقدار آن تقریبا" (e = ۲/۷) می‌باشد، درتوزیع پواسن‌چنان‌که دیدیم میانگین از قبل مشخص بوده و مقدار واریانس نیز مساوی مقدار میانگین می‌باشد.

مثال ۷ـ۳: میانگین مصرف‌گوشت هر خانوارساکن تهران ۴ کیلوگرم در ماه می‌باشد. مطلوب است احتمال آنکه مصرف خانوار در یک ماه به‌۶ کیلوگرم برسد.

پاسخ: (x =۶ و μ =۴) می‌باشد. به‌راحتی می‌توان از رابطه (۹ـ۳) مسئله را حل نمود.

$$P(۶,۴) = \frac{(e^{-۴})(۴^{۶})}{۶!} = ۰/۱۰۴۲$$

مقادیر احتمال مطلق و تجمعی توزیع پواسن برای بسیاری از مقادیر (X ، μ)که‌بیشتر مورد نیاز هستند محاسبه گردیده و در جداولی منعکس شده است که جدول مربوط به‌مقادیر

احتمال تجمعی توزیع پواسن به‌آخر کتاب ضمیمه شده است . واضح است‌که‌مقادیر مطلق رانیز به‌راحتی می‌توان با انجام عمل تفریق ساده‌ای از همین جدول به‌دست آورد . به‌عنوان‌مثال پاسخی را که برای مثال (۷ـ۳) به‌دست آوردیم یک بار دیگر با استفاده از جدول به دست می‌آوریم .

$$P(۶\text{و}۴) = \sum_{x=0}^{۶} P(x\ \text{و}\ ۴) - \sum_{x=0}^{۵} P(x\ \text{و}۴) = ۰/۸۸۸۳ - ۷۸۵۱ = ۰/۱۰۴۲$$

تذکر نکته‌ای در مورد توزیع پواسن ضرورت دارد و آن این است که هیستوگرام توزیع دوجمله‌ای وقتی که مقدار n بزرگ بوده‌ومقدار (p) بسیار کوچک و نزدیک به صفر است (مقادیری از قبیل $p = ۰/۰۰۵$ و . . . .) ، به‌هیستوگرام توزیع پواسن شباهت زیادی‌داشته و به همین دلیل در چنین مواقعی می‌توان مسائل توزیع دوجمله‌ای را از طریق توزیع پواسن به صورت ساده‌تری حل نمود . با حل یک مثال دراین موردبحث توزیع پواسن و نیز توزیعهای گسسته‌را به‌پایان رسانده و به‌بررسی‌توزیعهای‌پیوسته‌خواهیم پرداخت

مثال ۸ ــ ۳ . احتمال آنکه در جامعه تهران ، افراد درآمد ماهیانه بیش از یک میلیون تومان داشته باشند ($p = ۰/۰۰۱$) می‌باشد مطلوب است احتمال آنکه در یک نمونه (۸۰۰۰)نفری تعداد افراد با درآمد مزبور کمتراز ۷ نفر باشد ؟

پاسخ : توزیع اصلی برای مثال فوق توزیع دوجمله‌ای می‌باشد اما چون (n) بزرگ بوده و مقدار (p) نیز خیلی کوچک است ابتدا میانگین را از توزیع دوجمله به‌دست آورده و بقیه مسئله را با استفاده از توزیع پواسن حل می‌کنیم .

$$\mu = nP = ۸۰۰۰\ (۰/۰۰۱) = ۸$$

$$P[(x < ۷)\ \text{و}\ ۸] = \sum_{x=0}^{۶} P(x\ \text{و}۸) = ۰/۳۱۳۴$$

ب ــ ۲ . توزیع احتمال پیوسته

در قسمتهای اولیه این فصل ملاحظه گردید که محاسبه احتمال نقطه‌ای برای هریک از مقادیر یک متغیر پیوسته جز با استفاده از تابع چگالی احتمال و محاسبه احتمال فاصله‌ای برای همین متغیر جز بااستفاده ازتابع توزیع‌احتمال که مقادیر احتمال تجمعی را برای‌هریک از مقادیر آن متغیر به‌دست می‌دهد ، ممکن‌نیست . همین مسئله ضرورت بررسی توزیعهای احتمالی را که به‌متغیر پیوسته مربوط می‌شوند و می‌توان از آنها برای محاسبه مقادیراحتمال

استفاده نمود ثابت می‌کند . همچنین قبلا" دیدیم که مقدار احتمال نقطه‌ای برای یک متغیر پیوسته در عملیات آماری معمولا" مشکلی را حل نکرده و چندان مطلوب نیست زیرا مقدار آن بسیار ناچیز بوده و لذا قابل اغماض می‌باشد و به‌همین دلیل تابع چگالی احتمال نیز که مورد استفاده اصلی آن محاسبه احتمال نقطه‌ای می‌باشد اهمیت خود را تا حدودی از دست می‌دهد از طرف دیگر درست به‌همان دلایلی که احتمال نقطه‌ای را برای یک متغیر پیوسته غیر ضروری می‌نمودند، دربیشتر مسائل مقدار احتمال فاصله‌ای برای چنین متغیری مورد نیاز است . چنانکه می‌دانیم مقدار احتمال فاصله‌ای معمولا" با استفاده از مقادیر احتمال تجمعی به‌دست می‌آید که در یک متغیر پیوسته، تابع توزیع احتمال نقش مهمی را در محاسبات ایفا می‌نماید و مهمترین خاصیت تابع چگالی احتمال نیز این است که از طریق آن می‌توان تابع توزیع احتمال را به‌دست آورد .

توضیحات فوق یادآوری مطالبی بودند که قبلا" پیرامون توزیع احتمال پیوسته گفته بودیم و حال نوبت به‌بررسی توزیعهای احتمال پیوسته‌ای می‌رسد که در آمار حائز اهمیت می‌باشند. تنها توزیع احتمالی که برای محاسبه مقدار احتمال یک متغیر پیوسته اهمیت داشته و در سطح وسیعی در آمار مورد استفاده قرار می‌گیرد ، توزیعی به‌نام توزیع نرمال می‌باشد . این توزیع نه تنها در محاسبه احتمال برای متغیر پیوسته نقش اصلی را ایفا می‌کند بلکه در محاسبه احتمال بسیاری از متغیرهای گسسته نیز نقش مهمی را به‌عهده دارد زیرا چنانکه خواهیم دید دربسیاری از موارد آمارگران توزیعهای احتمال گسسته ازقبیل توزیع دوجمله‌ای را به‌توزیع نرمال تبدیل نموده و محاسبات خود را با استفاده ازتوزیع نرمال انجام می‌دهند . به‌این دلایل مبحث توزیع احتمال پیوسته را در بررسی توزیع نرمال خلاصه می‌کنیم .

ب ـ ۲ـ۱ : توزیع نـرمال

همانطورکه قبلا" گفته شد دربین توزیعهای گسسته و پیوسته‌ای که در آمار مورد استفاده قرار می‌گیرند توزیع نرمال مهمترین آنها می‌باشد که البته این توزیع مربوط به‌متغیر پیوسته می‌باشد . همانند توزیعهای احتمالی که قبلا" مورد بررسی قرار گرفتند در اینجا نیز درمورد شکل کلی توزیع ومیانگین و واریانس آن توضیحاتی داده و سپس روش محاسبه احتمال حوادث مختلف را با استفاده از توزیع نرمال موردبررسی قرار خواهیم داد . زیرا برخلاف توزیعهای احتمال گذشته ، روش محاسبه احتمال از طریق توزیع نرمال ویژگیهایی دارد که به‌یک بحث تفصیلی مستقل در این مورد نیاز دارد .

شکل (۳ـ۳) یک منحنی توزیع نرمال را نشان می‌دهد . چنانکه ملاحظه می‌گردد در توزیع نرمال مقدار احتمال نقطه ($x=\mu$) از هر نقطه دیگری بیشتر است و به‌تدریج که از نقطه μ به‌دو طرف حرکت می‌کنیم مقادیر احتمال به‌صورت کاملا" متقارنی مرتبا" کم می‌شوند

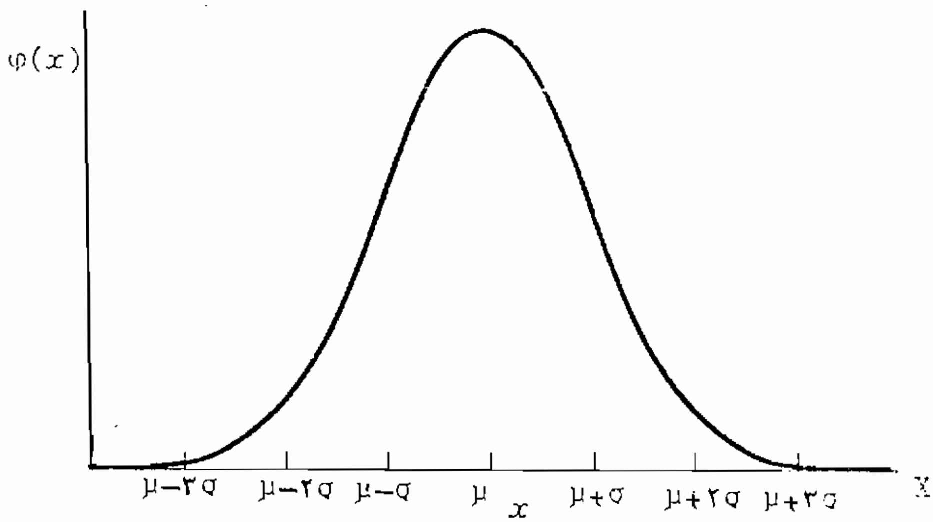
به‌نحوی که به‌عنوان مثال داریم :

$$P(\mu+\sigma)=P(\mu-\sigma) \quad , \quad P(\mu+2\sigma) = P(\mu-2\sigma),..........$$

چون شکل کلی منحنی نرمال به‌کاسه زنگ اخبار شباهت دارد از این توزیع گاهی با عنوان توزیع زنگی شکل نیز یاد می‌شود . بنابراین منحنی نرمال یک منحنی زنگی شکل است که‌سه خصوصیت زیر را دارا می‌باشد .

۱ ــ این منحنی که نقاط مختلف آن مقادیر احتمال را برای نقاط متناظر آن روی‌محور (x) ها نشان می‌دهد ، در ابتدا یک منحنی صعودی بوده و در نقطه $x = \mu$ به حداکثر مقدار خود می‌رسد و سپس در نقاط بعداز میانگین به‌صورتی کاملاً" متقارنی نسبت به‌ماگزیمم منحنی نزولی می‌شود .

۲ ــ نتیجه طبیعی خصوصیت فوق این است که در توزیع نرمال مقدار میانگین ، میانه و نما باهم مساوی خواهند بود . زیرا وقتی که دوطرف صعودی و نزولی منحنی نسبت‌به‌نقطه ($x = \mu$)قرینه باشند ، طبیعتاً" این نقطه درست وسط مقادیر (x) قرار خواهد داشت . از طرفی نمای مقادیر یک متغیرآن مقدار است‌که فراوانی مطلق و نسبی‌آن بیش از سایر مقادیر باشد که وقتی مسئله‌ای به‌صورت‌نظری حل می‌شود ، فراوانی نسبی جای خــود را به احتمال نظری می‌دهد . بنابراین نما در توزیع نرمال نقطه‌ای است که بیشترین احتمال‌را دارا باشد که در این صورت این نقطه همان میانگین می‌باشد .



شکل (۳ــ۳) یک منحنی توزیع احتمال نرمال

۳ ــ چنانکه قبلاً گفته شد ، سطح زیر منحنی مربوط به‌توزیع احتمال یک متغیر پیوسته نشان دهنده مجموع احتمالات مقادیر مختلف آن متغیربوده و بنابراین برابربا (۱) می‌باشد .
می‌دانیم که منحنی توزیع احتمال برای یک متغیر پیوسته ،بر اساس تابع چگالی احتمال آن به‌دست می‌آید و به‌علاوه با استفاده از تابع چگالی احتمال مقدار احتمال نقطه‌ای آن متغیرنیز به‌دست می‌آید .اگرچه برای یک متغیر پیوسته احتمال نقطه‌ای چندان مورد احتیاج نیست ، با این وجود معرفی تابع چگالی توزیع نرمال ، مفید به‌نظر می‌رسد .این تابع به‌صورت زیر می‌باشد .

$$f(x) = n(x;\mu,\sigma) = \frac{۱}{\sigma\sqrt{۲\pi}}\, e^{-\frac{۱}{۲}\left(\frac{x-\mu}{\sigma}\right)^2}$$

چنانکه ملاحظه‌می‌شود این تابع‌از متغیر (x)و نیز از مقدار پارامترهای μ ، σ تبعیت می‌کند .محاسبه مقدار میانگین وواریانس متغیری که توزیع احتمال آن نرمال است بااستفاده از همان روشهایی که در فصل دوم آمد انجام می‌گیرد و ویژگی مهمی ندارد .

مثال ۱۰ ــ۳ : متغیر تصادفی (x) با میانگین (μ = ۵) و انحراف معیار (σ =۲)دردست است . احتمال آنکه در یک آزمایش مقدار (x =۴) انتخاب شود چقدر است ؟

پاسخ : این یک احتمال نقطه‌ای است که مقدار آن را می‌توان از رابطه (۱۰ــ۳) یعنی تابع چگالی احتمال نرمال محاسبه نمود .

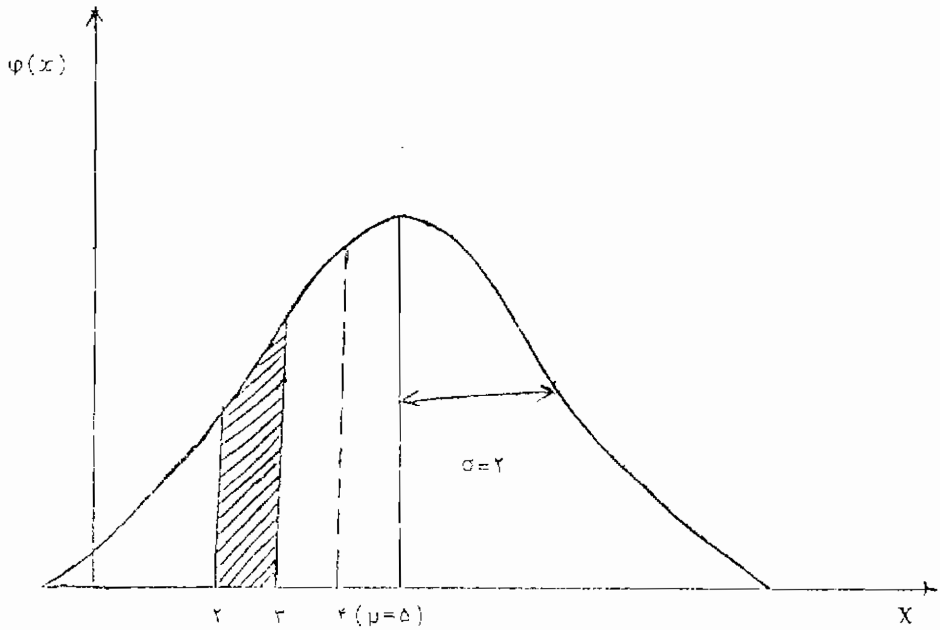
$$P(x=۴) = n(۴و۵و۲) = \frac{۱}{۲\sqrt{۲\pi}}\, e^{-\frac{۱}{۲}\left(\frac{۴-۵}{۲}\right)^2} = ۰/۱۶۹$$

این مقدار در شکل (۴ــ۳) به‌صورت خط چین نشان داده شده است .

اما همان طورکه می‌دانیم برای یک متغیر پیوسته معمولاً به‌احتمال فاصله‌ای نیازمندیم و این گونه احتمال با استفاده از تابع اولیه یا انتگرال تابع چگالی که تابع توزیع نام دارد به‌دست می‌آید . به‌عنوان مثال اگر به‌خواهیم $P(۲ < X < ۳)$ را در مثال (۱۰ــ۳)به‌دست آوریم باید به‌صورت کلی زیر عمل کنیم .

$$P(۲<x<۳) = \int_۰^۳ n(x،۵،۲) - \int_۰^۲ n(x و۵و۲) = F(۳) - F(۲)$$

سپس مقادیر هریک از دو جمله سمت راست را محاسبه نموده و مقدار احتمال مورد نظر را بدست آوریم اگر منحنی توزیع نرمال فوق به‌صورت زیر باشد مقدار احتمال محاسبه شده برابر سطح هاشور زده زیر منحنی خواهد بود .

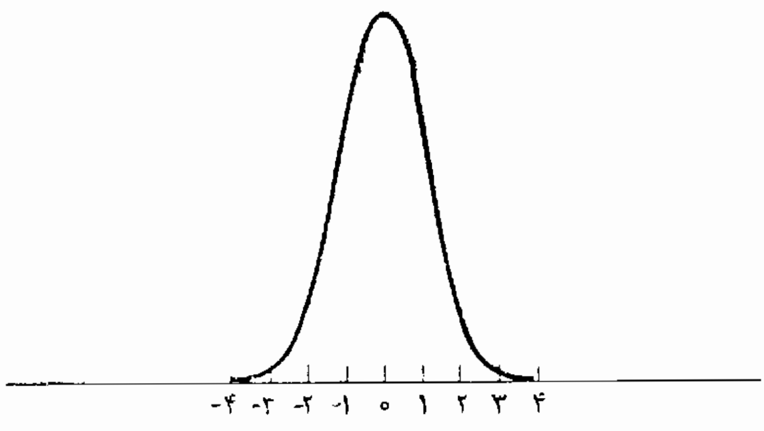


شکل (۴-۳)

بدون شک انجام محاسبات فوق برای به‌دست آوردن احتمال فاصله‌ای در هر مسئله‌کاری مشکل است و شخص آرزو می‌کند که همانند توزیعهای دوجمله‌ای و پواسن جداولی در اختیار داشته باشد که مقدار احتمال تجمعی برای هریک از مقادیر (X) را نشان دهد . اما تأمین این خواسته برای همه توزیعهای نرمال امری غیر ممکن و نا مقدور می‌باشد زیرا هریک از دو پارامتر میانگین (μ) و انحراف معیار (σ) بینهایت مقدار را می‌توانند اختیار کنند و به ازاء هریک از این مقادیر ، منحنی نرمال خاص وجود دارد . بنابراین بینهایت توزیع نرمال موجود است و مسلم است که نمی‌توان برای همه آنها جدول تابع توزیع (انتگرال تابع چگالی) را تنظیم نمود . برای حل معضله فوق ، یعنی برای آنکه بتوان در حل مسائل مربوط به‌توزیع نرمال از جدول استفاده نموده و بدین ترتیب حجم محاسبات و عملیات لازم را کاهش داد روشی ابداع شده است که ذیلاً توضیح داده می‌شود .

ب ــ ۲ ــ ۲ . روش ساده محاسبه احتمال فاصله‌ای با استفاده از توزیع نرمال

همان‌طور که قبلاً گفته شد خاصیت اصلی روشی که ذیلاً توضیح داده می‌شود این است که استفاده از جدول را در حل مسائل مربوط به‌توزیع نرمال میسر می‌سازد . برای رسیدن به‌این منظور مقدار تابع احتمال $n(x ، ۰ ، ۱)$ به‌ازاء مقادیر مختلف (X) محاسبه شده و در جدولی به‌نام جدول (Z) که ضمیمه این کتاب می‌باشد ثبت شده است. چون تابع توزیع (Z) یعنی تابع توزیع نرمالی که میانگین آن ($\mu = ۰$) و انحراف معیار آن ($\sigma = ۱$) می‌باشد بنام نرمال استاندارد معروف است به‌جدول مزبور نیز جدول مقادیر نرمال استاندارد گویند. با داشتن همین یک جدول در واقع هدف مزبور تأمین شده است زیرا با داشتن جدول مقادیر Z به‌راحتی می‌توان تعداد احتمال فاصله‌ای را برای توزیعهای نرمال با میانگین و انحراف معیارهای گوناگون بدون آنکه عملیات مطروحه قبلی ضرورت یابند محاسبه نمود . دلیل این امر و کیفیت انجام این عمل ذیلاً توضیح داده می‌شود . ضمناً شکل (۵ ــ ۳) یک توزیع نرمال استاندارد (Z) را نشان می‌دهد که به‌صورت کلی $n(Z ، ۰ ، ۱)$ نشان داده می‌شود .



شکل (۵ ــ ۳) منحنی نرمال استاندارد

مسئله‌ای که باعث می‌شود بتوانیم برای محاسبه احتمال فاصله‌ای مقادیر مختلف یک متغیر تصادفی مثل (X) با توزیع احتمال نرمال و هر مقدار میانگین و انحراف معیار دلخواه از جداول مربوط به‌توزیع Z استفاده کنیم این است که مقدار احتمال فاصله‌ای بر اساس مقدار سطح زیر منحنی نرمال محاسبه می‌شود که این مقدار به‌نوبهٔ خود از طریق محاسبه مقدار تابع

توزیع (انتگرال تابع چگالی) $F(x)$ قابل‌اندازه‌گیری می‌باشدوچنانکه‌بعدا"ثابت‌می‌شود هر مقدار (X)درهرتوزیع‌نرمال بایک‌مقدار (Z)درتوزیع‌نرمال‌استاندارد متناظراست‌به‌طوری‌که اگر آن‌مقدارخاص X را با (x_1)و آن‌مقدارخاص (Z)را با (Z_1) نشان دهیم داریم :

$$\int_{-\infty}^{x_1}\varphi(x)dx = F(x_1) = F(Z_1) = \int_{-\infty}^{Z_1}\varphi(Z)dz$$

یعنی اینکه سطح زیر منحنی نرمال تا قبل از نقطه (x_1) در توزیع نرمال (X) با سطح‌زیر منحنی نرمال تا قبل از نقطه (Z_1) در منحنی Z باهم برابر بوده و لذا داریم .

$$P(X < x_1) = P(Z < Z_1)$$

نتیجه‌ای‌که از ایـن بحث گـرفتـه می‌شود این است که اگر بتوان برای هرمقدارمتغیر تصادفی باتوزیع نرمال و باهرمقدار (μ و σ)دلخواه نقطه‌متناظرش را در توزیع Z به‌دست آوریم به‌راحتی می‌توان مقدار احتمال فاصله‌ای در آن توزیع نرمال را با استفاده از جدول (Z) محاسبه نمائیم .[1] می‌دانیم که دو توزیع از سه جهت شکل کلی توزیع ، میانگیـــن و انحراف معیار می‌توانند باهم تفاوت داشته باشند اگر شکل کلی دو توزیع نرمال باشد ، در نتیجه می‌توان گفت که آن دو توزیع نرمال فقط از دو جهت می‌توانند باهم اختلاف داشته باشند بدین ترتیب که یا فقط میانگین این دو توزیع باهم تفاوت دارد و یا انحراف معیـار آنها و نهایتا" ممکن است دو توزیع نرمال هم از نظر میانگین و هم از نظر انحراف معیـار باهم متفاوت باشند . بطور طبیعی نتیجه این مطلب آن است که مقدار (X) در یک توزیـع نرمال نیز به‌همین دو دلیل بوده است که با مقدار (Z) در توزیع نرمال استانداردتفاوت پیدا کرده است (این مطلب در قسمتهای بعدی بیشتر توضیح داده خواهد شد) . چه اگر هر توزیع نرمالی دارای مشخصه‌هایی با مقادیر ($\mu = 0$ و $\sigma = 1$) باشد مقادیر X در محور افقی آن دقیقا" بر مقادیر Z منطبق خواهند بود .بنابراین برای آنکه‌مقدارمتناظر X را در توزیع (Z) به‌دست آوریم باید ببینیم مقدار متغیر (X) که دارای توزیع نرمال با مشخصه‌هـای ($\mu \neq 0$ و $\sigma \neq 1$)است ، در صورتی که فرضا" مقدار ($\mu = 0$ و $\sigma = 1$)باشد چقدر خواهد شد.

---

۱ ـ تناظر فوق فقط برای احتمال فاصله‌ای وجود دارد و مقدار احتمال نقطه‌ای چنین تناظری پیروی‌نمی‌کند یعنی‌بعنوان مثال اگر در یک‌توزیع نرمال $n(x ، ۱ ، ۵)$ نقطه متناظر با نقطه ($x = ۵$) در توزیع Z ، نقطه ($Z = ۱$) باشد رابطه زیـر الزامـا" وجـود نـدارد . $P(x = ۵) = P(Z = ۱)$.

برای این کار ابتدا روش محاسبه مقدار (X) را برای حالت فرضی ($\mu=0$ و $\sigma=1$)توضیح داده و سپس طرز محاسبه این مقدار را در حالت ($\mu=0$ و $\sigma=1$) شرح می‌دهیم که نتیجه به‌دست آمده پس از مرحله دوم همین مقدار متناظر X در توزیع (Z) می‌باشد .

بحث را از محاسبه مقدار (X) که توزیع نرمال داشته و مشخصه‌های آن ($\mu \neq 0$ و $\sigma \neq 1$) است برای حالت فرضی ($\mu=0$ و $\sigma \neq 1$) شروع می‌کنیم . می‌خواهیم بدانیم که اگر به‌جای مقدار میانگین فعلی مقدار ($\mu=0$) بود و انحراف معیار ، همین مقدار فعلی را می‌داشت ، مقدار (X) چه تغییری می‌کرد . به‌شکل (۶-۳) نگاه کنید . در شکل الف یک منحنی نرمال با مشخصه‌های ($\mu=0$ و $\sigma=2$)و در شکل ب یک منحنی‌نرمال با مشخصه‌های ($\mu=5$ و $\sigma=2$) رسم شده است و مصداق بحث مادر این شکل خاص این است که‌آیا مقدار متغیر (X)درشکل ب (فرضاً $x=-3$)چه مقداری باید باشد درصورتی‌که میانگین منحنی به‌جای ($\mu=5$) مقدار ($\mu=0$) را دارا بوده و سایر شرایط ثابت باشد همان‌طور که‌درشکل ملاحظه می‌شود در واقع شکل ب همان منحنی الف است که همه نقاط‌آن ۵ واحد به راست منتقل شده و نتیجتاً" میانگین آن از ($\mu=0$) به ($\mu=5$) تغییر یافته است . بنابراین اگر از هریک از مقادیر (X) در منحنی ب همین ۵ واحد افزوده شده را که با مقدار میانگین این منحنی برابر است ، کم کنیم در واقع مقدار میانگین ۵ واحد کم شده و نتیجتاً" نقطه متناظر آن در شکل الف به‌دست خواهد آمد و این همان مقدار مطلوب ماست . برای مقدار ($x=-3$) این مقدار متناظر ، به‌قرار زیر است .

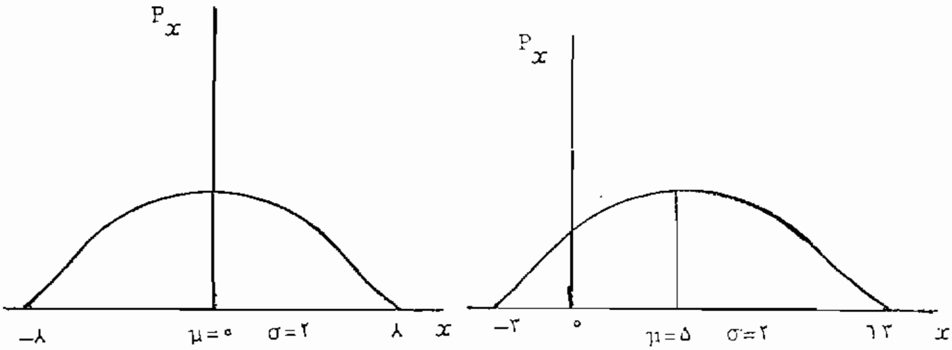
$$-8 = -5 - 3 - = \mu - x = \text{مقدار متناظر}(x=-3) \text{ درمنحنی الف}$$

بنابراین برای آنکه در حالت ثبات سایر شرایط مقدار (x) را با هر مقدار دلخواه و هر توزیع نرمال با میانگین ($\mu \neq 0$)برای حالت فرضی ($\mu=0$) به‌دست آوریم کافی است مقدار میانگین را از آن کم کنیم .

حال زمان آن فرا رسیده است که ببینیم اگر علاوه بر شرط ($\mu=0$) ، شرط دوم یعنی ($\sigma=1$) نیز فرض شود در مقدار X چه تغییری ایجاد می‌شود . به‌راحتی قابل اثبات است که هرگاه (X) متغیری باشد که دارای توزیع نرمال با هر میانگین دلخواه بوده‌ولی انحراف معیار آن ($\sigma \neq 1$) باشد اگر هریک از مقادیر آن را بر انحراف معیار آن (σ_x)تقسیم‌کنیم مقدار متناظر همان نقطه در یک توزیع نرمال دیگر با همان میانگین ولی با انحراف معیار ($\sigma=1$)به دست‌می‌آید و بنابر این برای‌آنکه Z متناظر با یک مقدار (X)باتوزیع نرمال و با خصوصیات $\mu \neq 0$ و $\sigma \neq 1$) را به‌دست آوریم کافی است ابتدا مقدار میانگین توزیع مورد نظر را از آن مقدار خاص x کم کنیم تا میانگین توزیع ($\mu=0$) شود و سپس

نتیجه را به‌انحراف معیارآن توزیع تقسیم کنیم تا انحراف معیارآن نیز ($\sigma = 1$) شده و نتیجتا" مقدار z متناظر با نقطه مورد نظر به‌دست آید . یعنی داریم :

$$Z = \frac{x-\mu}{\sigma} \quad \text{متناظر با} \quad (X=x)$$



الف　　　　　　　　ب

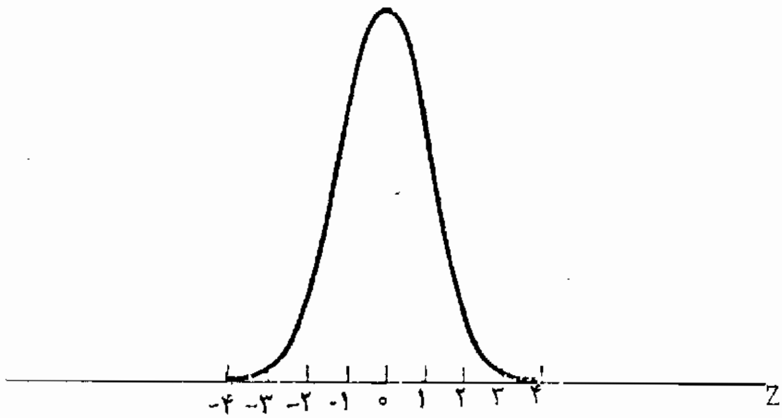
شکل (۶ــ۳)

علت این امر نیز روشن‌است . می‌دانیم‌که سطح زیر منحنی ، در همه منحنی‌های‌توزیع احتمال نرمال برابر ۱ می‌باشد (چون مجموع احتمالات مربوط به همه حوادث یک آزمایش برابر ۱ است)ازطرفی در یک منحنی میانگین مرکز توزیع و انحراف معیار ، طول محور(X) در بین دو نقطه آغاز و پایان منحنی را مشخص می‌سازد .[1] با توجه به مطالب فوق با کمی دقت در می‌یابیم که کشیدگی یا پهنی منحنی نرمال فقط به‌مقدار طول محور (X)ها در زیر منحنی (بین دو نقطه آغاز و پایان آن) بستگی دارد . زیرا منحنی مساحت ثابت و مشخصی‌را روی‌نقاط این‌طول‌تقسیم و توزیع می‌نماید . بنابراین هرچه فاصله محور(X)ها در بین دو نقطه آغاز و پایان منحنی نرمال بیشتر باشد منحنی پهن تر می‌شود و بعکس هرچه این فاصله کمتر باشد منحنی درازتر و کشیده‌تر می‌شود . دو شکل (۵ ــ ۳) و (۷ ــ ۳) را باهم مقایسه کنید . در شکل (۵ ــ ۳) که یک منحنی نرمال استاندارد است منحنی‌ازنقطه ($x = -4$)شروع‌شده و به نقطه ($x = 4$)ختم می‌شود یعنی‌طول محور (x)هادرزیرمنحنی

---

۱ ــ بر طبق قضیه‌چبی شیف‌فاصله $[(\mu - 3\sigma) - (\mu + 3\sigma)]$ بیش از ۹۹% مقادیر (X) را در بر می‌گیرد . این مسئله در شکل (۲ــ۳) نشان داده شده است .

فقط ۸ سانتی متر است و به‌همین دلیل این منحنی نسبت به‌منحنی (۷–۳) که طـول محور (x)ها در زیر منحنی آن ۲۴ سانتیمتر است ، کشیده‌تر می‌باشد . منحنی (۷–۳) نیز یـک منحنی نرمال با $\mu = ۰$ اما $\sigma \neq ۱$ می‌باشد . از طرفی چنانکه گفته شد طول محور (x)هادر زیر منحنی فقط توسط انحراف معیار (σ) مشخص می‌گردد یعنی هرچه مقدار σ بیشتر باشد این طول هم بیشتر بوده و بالعکس. نتیجه‌ای که گرفته می‌شود این است که کشیدگی و پخی یک منحنی فقط به‌مقدار σ بستگی دارد . در دو شکل (۵ – ۳) و (۷ – ۳) نیز این جریـان به وضوح مشهوداست . دوشکل زیر هر دو نرمال بوده و میانگین‌هردو نیز ($\mu = ۰$)می‌باشد و تنها تفاوت این دوشکل درانحراف معیار آنهاست زیرا مقدار انحراف معیارمنحنی (۵–۳) همانگونه که ملاحظه می‌گردد ($\sigma = ۱$) است در حالی که در منحنی (۷–۳) این‌مقدار معادل ($\sigma = ۳$) می‌باشد و بنابراین همین بزرگی مقدارانحراف معیار علت پخ تربودن‌شکل (۷–۳) نسبت به (۵ – ۳) می‌باشد .



شکل (۵ – ۳) یک توزیع نرمال استاندارد

شکل (۷ – ۳) یک توزیع نرمال با ($\mu = ۰$ ، $\sigma = ۳$)

مطلب فوق دلیل اصلی نکته‌ای است که قبلاً" به‌عنوان روش به‌دست‌آوردن نقاط متناظر در دو توزیع (Z) و یک توزیع نرمال با ($\mu = ۰$ و $\sigma \neq ۱$) تشریح گردید . در این روش گفتیم که اگر مقدار متغیر (X) را بر σ_x تقسیم کنیم مقدار (Z) متناظر آن به‌دست می‌آید. بدین ترتیب که به‌عنوان مثال اگر تنها تفاوت دو منحنی (۵ ــ ۳) و (۷ــ۳) در انحراف معیار آنها باشد که در این مثال (σ_x) سه برابر (σ_Z) می‌باشد بنابراین اگر فاصله [(۴) ــ(۴ــ)] را که‌دو حدمتغیر (Z) می‌باشد سه‌برابر کنیم‌این فاصله به‌فاصله[(۱۲)ــ(۱۲ــ)] تبدیل شده و متغیر Z نیز به‌متغیر (X) که منحنی آن ، منحنی (۷ــ۳) می‌باشد تبدیل می‌گردد . عکس این جریان نیز صادق است یعنی اگر فاصله زیر منحنی (۷ــ۳)را برانحراف معیار آن یعنی عدد (۳) تقسیم کنیم منحنی (۷ــ۳) تغییر شکل داده و کاملاً" بر (۵ ــ ۳) منطبق می‌گردد . می‌توان این مسئله را به‌این صورت نیز بیان نمود که یک سطح ثابت مثل یک سانتیمتر مربع ، یک اینچ مربع و یا هر مقدار ثابت دیگری ، وقتی بخواهد در یک فاصله ۸ سانتیمتری روی محور x ها قرار بگیرد یعنی ضمن اینکه شکل کلی آن مثل منحنی‌نرمال است انحراف معیار آن ($\sigma =۱$) باشد مطمئناً" مشابه منحنی (۴ــ۳) خواهد بود و همین سطح ثابت اگر بخواهد به‌شکل منحنی نرمال و با انحراف معیار ($\sigma =۳$) توزیع گردد شکل آن قطعاً" مثل شکل (۵ ــ ۳) می‌باشد .

قبلاً" گفته شد که علت آنکه سطح زیر منحنی در همه منحنی‌های توزیع احتمال ثابت فرض می‌شود این است که همه آنها مجموع احتمالات مربوط به‌همه حوادث ممکن‌الوقوع در یک آزمایش را نشان می‌دهند که این مقدار برابر ۱ می‌باشد . اگر سطح زیر منحنی در هریک از دو شکل (۵ــ۳) و (۷ــ۳) را یک‌سانتیمتر مربع فرض کنیم در شکل (۵ــ۳) این یک سانتیمتر مربع بین ۸ سانتیمتر بطور نرمال تقسیم می‌شود در حالی که در شکل (۷ــ۳) همین یک سانتی‌مترمربع بین ۲۴ سانتیمتر به‌صورت کلی نرمال توزیع می‌گردد . این مطلب را به‌صورت زیر نیز می‌توان نوشت :

$$\text{سطح زیر منحنی} = \underbrace{\int_{-۴}^{۴} \varphi(x)d_x}_{\text{شکل (۳ــ۴)}} = \underbrace{\int_{-۱۲}^{۱۲} \varphi(x)d_x}_{\text{شکل (۳ــ۵)}} = \text{سطح زیر منحنی} = ۱\ \text{cm}^2$$

از این‌گونه روابط برای اجزاء و قسمتهای کوچکتر محور (x) ها نیز می‌توان نوشت . به‌عنوان مثال داریم :

$$\text{درشکل (۳-۵)} \quad \int_{-۴}^{۰}\varphi(x)d_x = \int_{-۱۲}^{۰} \varphi(x)d_x \quad \text{درشکل (۳-۷)}$$

همان‌طور که ملاحظه می‌شود اگر بخواهیم مقدار Z متناظر با هریک ازمقادیرمتغیری مثل (X) را که دارای توزیع نرمال غیر استاندارد اما با ($\mu = ۰$) که در شکل (۷ــ۳)نشان داده شده است ، به‌دست آوریم کافی است آن را برانحراف معیار خود تقسیم کنیم . این قاعده برای هر مقدار دلخواه انحراف معیار صادق است . چون اگر به فرض ($\sigma >۱$) باشد درآن صورت مقدار محور (x) ها در زیر منحنی نرمال مربوط بزرگتر از همین مقداردر زیر منحنی Z است و با عمل تقسیم ، طول محور x ها در زیرآن منحنی نرمال و منحنی (Z) مساوی می‌شوند . بعکس اگر ($\sigma < ۱$) باشد طول محور (x)ها در زیر منحنی مربوط به خود کوتاهتر از مقدار مشابه آن در منحنی (Z) بوده و با عمل تقسیم بازهم این دو طول مساوی هم خواهند شد و بدین ترتیب رابطه (۱۱ــ۳) اثبات می‌گردد .

$$Z = \frac{x - \mu}{\sigma}$$

چون در هر منحنی توزیع احتمال پیوسته مقدار سطح زیر منحنی تا قبل از هر نقطه مثل (x) نشان دهنده مقدار احتمال $P(X < x)$ بوده و کلیه نقاط روی محور x ها در یک توزیع احتمال نرمال قابل تبدیل به‌نقاط متناظر خود در توزیع (Z) می‌باشد (با استفاده‌از رابطه فوق) ،لزومی ندارد مقدار توابع توزیع احتمال نرمال را برای کلیه توزیعهای نرمال‌با میانگین‌ها و انحراف معیارهای مختلف محاسبه نمود بلکه با داشتن مقادیر (μ و σ) برای یک توزیع نرمال ،به‌راحتی می‌توان بااستفاده ازجدول مقادیرتابع توزیع (Z)مقداراحتمال فاصله‌ای را حساب نمود . بدین ترتیب که ابتدا مقادیر (Z) متناظر با دو حد فاصله مورد نظر را به‌دست‌آورده و سپس سطح زیر منحنی نرمال استاندارد (Z) را در این فاصله‌جدید با استفاده از جدول (Z) به‌دست می‌آوریم .

مثال ۱۱ــ۳ . فرض‌کنیم درشکل (۵ ــ ۳) مقدار $F(Z=۱) = ۰/۶$ و $F(Z=۳) = ۰/۸$ باشد، احتمال‌آنکه یک متغیر تصادفی مثل (X) مقدار ($۳ < X < ۹$) را در منحنی (۷ــ۳) اختیار کند چقدر است ؟

پاسخ . اگر دو نقطه مورد نظر در شکل (۷ــ۳) با نقاط($Z=۱$ و $Z=۳$) در شکل (۵ــ۳) متناظر باشند مسئله به‌راحتی قابل حل است بنابراین ابتدا نقاط‌متناظر این دو نقطه‌را در شکل (۵ ــ ۳) به‌دست می‌آوریم ، داریم .

$$Z = \frac{x - \mu_x}{\sigma_x} \qquad Z_1 = \frac{۳ - ۰}{۳} = ۱$$

$$Z_2 = \frac{۹ - ۰}{۳} = ۳$$

تناظر بین نقاط مورد نظر وجود دارد و لذا احتمال فاصله‌ای مورد نظر مقدار زیر را خواهد داشت .

$$P(۳ < x < ۹) = \int_{۳}^{۹} \varphi(x)d_x = \int_{۱}^{۳} \varphi(z)d_z = F(Z=۳)-F(Z=۱)$$

$$P(۳ < x < ۹) = F(۳) - F(۱) = ۰/۸ - ۰/۶ = ۰/۲$$

چنان که قبلاً گفته شد مقدار احتمال نقطه‌ای در یک توزیع پیوسته مساوی صفر فرض می‌شود و به‌همین دلیل رابطهٔ زیر را داریم :

$$P(X < x) = P(X \leqslant x) \qquad (۳-۱۳)$$

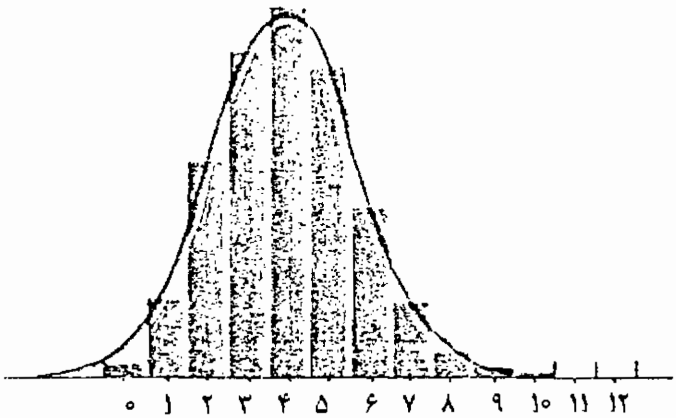
زیرا تفاوت دوطرف چپ و راست معادله فقط یک احتمال نقطه‌ای $P(X = x)$ می‌باشد که نادیده گرفته می‌شود .

سابق براین گفته شد که در بسیاری از موارد می‌توان مسائل مربوط به‌توزیعهای گسسته را نیز با استفاده از توزیع نرمال حل کرد . مهمترین این موارد حل مسائل تقریب توزیع دوجمله‌ای از طریق توزیع نرمال می‌باشد که ذیلاً مورد بحث قرار می‌گیرد .

ب ــ ۲ــ۳ . تقریب دوجمله‌ای با استفاده از توزیع نرمال :

اگر در مسئله‌ای که با استفاده از قانون توزیع دوجمله‌ای حل می‌شود احتمال فاصله‌ای مطلوب‌بوده و مقدار n نیززیاد و بزرگ باشد ،حل مسئله دشوار خواهد شد زیرا برای nهای بزرگ جدول در اختیار نبوده و حل مستقیم نیز مستلزم عملیات ریاضی مفصل و پیچیده می‌باشد . از طرفی به‌راحتی می‌توان اثبات کرد که وقتی مقدار (n) افزایش می‌یابد و عدد نسبتاً" بزرگی می‌شود ، اگر احتمالات مربوط به‌همه مقادیر ممکن موفقیتها که با (χ) نشان داده می‌شود به‌صورت هیستوگرام نشان داده شده‌و از آن یک‌منحنی بگذرانیم ، این منحنی تقریباً" مشابه منحنی توزیع نرمال بوده و هرچه مقدار بیشتر شود این شباهت بیشتر و بیشتر خواهد شد . در شکل (۸ــ۳) یکی از این منحنی‌هاکه‌مربوط به‌توزیع‌دوجمله‌ای ($b(\chi$ ، ۱۵ ، $\frac{۱}{۲}$) می‌باشد ،نشان داده شده‌است . البته شرط بزرگ بودن مقدار (n)وقتی مطرح‌است که مقادیر p و q با $\frac{۱}{۲}$ ، فاصله نسبتاً" زیادی داشته باشند اما اگر این مقادیر تقریباً"مساوی $\frac{۱}{۲}$ باشند (از قبیل $p = ۰/۴$ و یا $p = ۰/۶$ و . . .) ، حتی اگر مقدار n کوچک هم باشد بازهم‌توزیع آن تقریباً" نرمال است، ضمناً"فراموش‌نشود که‌مقدار این‌دونباید‌خیلی‌به (۰ و یا ۱)

نزدیک باشد چون همانطور که قبلا" گفته شد اگر مقدار p یا q به (۰ یا ۱) نزدیک باشد برای حل مسئله از توزیع پواسن استفاده می‌شود.



شکل (۸-۳) هیستوگرام توزیع دوجمله‌ای

$b(X, ۱۵, \frac{۱}{۲})$

با توجه به‌مطلب فوق یعنی با توجه به‌اینکه در n های بزرگ منحنی توزیع دوجمله‌ای و نرمال تقریبا" شبیه یکدیگر هستند و اگر p و q مقداری نزدیک به $(\frac{۱}{۲})$ داشته باشند در n های کوچک نیز چنین شباهتی برقراراست ، در بسیاری از مسائل مربوط به‌توزیع دوجمله‌ای ترجیح داده‌می‌شود که‌از توزیع نرمال‌که منحنی آن مشابه منحنی دوجمله‌ای می‌باشد استفاده شده و با استفاده‌از جدول (Z) مسئله‌به‌راحتی حل شود. برای این کار ابتدا فاصله‌متناظر با فاصله موردنظر را در منحنی (Z)مشخص نموده و سپس با استفاده از جدول (Z)مسئله را حل می‌کنند . از آنجا که در اصل ، توزیع مربوط به‌مسئله توزیع دوجمله‌ای‌بوده و نمودار مربوط به‌آن یک نمودار هیستوگرام است که به‌منحنی نرمال تقریب شده است باید در هنگام تعیین حدود فاصله موردنظر روی این منحنی ، به‌ویژگی هیستوگرام بودن‌آن‌توجه داشته و بر مبنای‌این خصوصیت ،فاصله موردنظر را تعیین نمود . به‌عنوان مثال اگربخواهیم احتمال آنکه (X) درآزمایشی که منحنی آن در شکل (۸-۳) رسم شده است $P(۳< X<۶)$ را به‌دست آوریم ، این فاصله در روی منحنی نرمال به‌صورت $P(۳/۵< x<۵/۵)$ نشان‌داده می‌شود . چون داریم .

$$P(۳ < x <۶)= P(۴)+ P(۵)$$

مقدار هریک از دوجمله سمت راست معادله فوق ، برابر سطح مستطیلی است که برروی همان عدد قرار دارد . و چون قاعده‌اصلی مستطیلهای فوق‌به‌ترتیب (۴/۵ − ۳/۵)و (۵/۵−۴/۵) می‌باشد اگر بخواهیم مقدار این دوجمله را از روی منحنی که از این مستطیل‌ها تقریب شده است به‌دست آوریم خواهیم داشت :

$$P(۴) = P(۳/۵ < x < ۴/۵)$$

$$P(۵) = P(۴/۵ < x < ۵/۵)$$

و در نتیجه صحت ادعای فوق ثابت می‌گردد و فاصله مورد نظر در توزیع دوجمله‌ای ، روی منحنی نرمال به‌صورت زیر به‌دست می‌آید .

$$P(۳ < x < ۴) = P(۴) + P(۵) = P(۳/۵ < x < ۴/۵) + P(۴/۵ < x < ۵/۵) = P(۳/۵ < x < ۵/۵)$$

پس از آنکه فاصله موردنظر در روی منحنی نرمالی‌که از تقریب هیستوگرام دوجمله‌ای‌به‌دست آمده است مشخص شد ، فاصله متناظر با آن را روی منحنی نرمال استاندارد (Z) به دست می‌آوریم و مسئله را حل می‌کنیم . برای به‌دست آوردن فاصله متناظر برروی منحنی (Z)با همان روشی که قبلا" گفته شد عمل می‌کنیم یعنی دو نقطه (Z) متناظر با دو نقطه حدی فاصله مورد نظر را از رابطه زیر به‌دست می‌آوریم .

$$Z = \frac{x - \mu_x}{\sigma_x}$$

چون (X) متغیری با توزیع دوجمله‌ای است بنابراین داریم :

$$\mu_x = nP \quad \text{و} \quad \sigma_x = \sqrt{nPq}$$

اگر این مقادیررا در رابطه قبل قراردهیم رابطه (۱۳−۳) که خاص توزیع دوجمله‌ای می‌باشد به‌دست می‌آید .

$$Z = \frac{x - nP}{\sqrt{nPq}} \qquad (۳−۱۳)$$

پس از آنکه فاصله مورد نظر، روی منحنی Z به‌دست آمد با استفاده از جدول Z مسئله را حل می‌کنیم.

در بسیاری از موارد این شبهه برای افراد به‌وجود می‌آید که آیا شرایط تقریب توزیع دونقطه‌ای به‌نرمال وجود دارند یا خیر؟ به‌عبارت دیگر فرد مردد می‌شود که آیا n به قدر کافی بزرگ و یا p و q به‌قدرکافی به‌عدد $(\frac{1}{2})$ نزدیک هستند که توزیع دوجمله‌ای‌را به‌نرمال تقریب نموده و مسئله را با استفاده از جدول (Z) حل کند؟ یا اینکه شرایط مورد نظر فراهم نبوده و باید علیرغم همه مشکلات مسئله را بازهم از همان قانون توزیع دوجمله‌ای حل کند. برای رهائی از این‌مشکل گفته می‌شود که‌اگر دو شرط زیر وجود داشته‌باشند تقریب توزیع دوجمله‌ای به‌نرمال بلامانع است و در غیر این صورت نمی‌توان‌از توزیع نرمال‌استفاده نمود.

$$nP > ۵ \quad \text{و} \quad nq > ۵$$

مثال ۱۳-۳: ۶۰% توپهائی که یک بازیکن‌بسکتبال پرتاب می‌کند واردحلقه می‌شود. مطلوب است احتمال آنکه از ۱۰۰ پرتاب آینده کمتر از نصف آنها وارد حلقه شوند.

پاسخ: $p = ۰/۶$ بوده، $nP=۶۰$ و $nq=۴۰$ بوده‌و شرایط‌برای‌تقریب‌دوجمله‌ای‌به‌نرمال فراهم است و ضمناً" $P(x < ۵۰)$ نیز مطلوب مسئله می‌باشد. داریم:

$$\mu = nP = (۱۰۰)(۰/۶) = ۶۰$$

$$\sigma = \sqrt{n\,P\,q} = \sqrt{(۱۰۰)(۰/۶)(۰/۴)} = ۴/۹$$

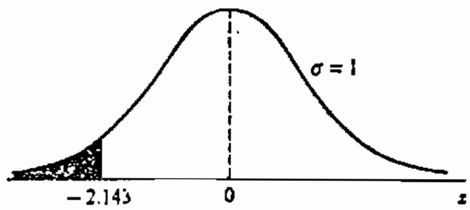
فاصله مورد نظر برروی منحنی نرمال همه نقاط قبل از ($x = ۴۹/۵$) می‌باشد. که متناظر آن را روی منحنی به‌دست می‌آوریم:

$$Z = \frac{x - n\,P}{\sqrt{nPq}} = \frac{۴۹/۵ - ۶۰}{\sqrt{۴/۹}} = -۲/۱۴۳$$

بنابراین پاسخ مسئله مقدار زیر خواهد بود.

$$P(x < ۵۰) = \sum_{x=۰}^{۴۹} b(X \text{،} ۱۰۰ \text{ و } ۰/۶) = P(Z < -۲/۱۴۳) = ۰/۰۱۶$$

این احتمال در شکل (۹-۳) نشان داده شده است.



شکل (۹–۳) مساحت مطلوب در مثال (۱۳–۳)

مسائل فصـــل ســوم

۱ ــ ثابت کنید که در یک توزیع دوجمله‌ای $\sigma^2 = n\,p\,q$ می‌باشد.

۲ ــ اگر (X) یک متغیر تصادفی پیوسته با میانگین (μ) و انحراف معیار (σ) باشد و متغیر نرمال استاندارد (Z) را به‌صورت زیر داشته باشیم:

$$Z = \frac{x - \mu_x}{\sigma_x}$$

ثابت کنید که $\mu = 0$ و $\sigma_z = 1$ می‌باشد.

۳ ــ تابع چگالی متغیر تصادفی پیوسته X که مقادیر آن فقط بین صفر تا ۴ تغییر می‌کند می‌کند عبارتست از $f(x) = \frac{1}{2} - aX$ در این تابع a عدد ثابتی است الف) a را محاسبه کنید. ب) $p(1 < X < 2)$ را پیدا کنید.

۴ ــ تابع چگالی متغیر پیوسته تصادفی X که فقط مقادیر بین $X=2$ تا ۸ را اختیار می‌کند عبارت است از $a(X+3)$ که در آن a عدد ثابتی می‌باشد.

الف) a را محاسبه کنید　　　　ب) $p(3 < X < 5)$ را پیدا کنید.

ج) $p(X > 4)$ را پیدا کنید.　　　　د) $p(|X - 5| < 0/5)$ را پیدا کنید.

۵ ــ ۵۰۰ لاستیک در انبار شرکتی وجود دارد که ۵۰ عدد آن معیوب است. یک مشتری ۱۰ عدد از این لاستیکها را می‌خرد احتمال اینکه وی ۸ لاستیک سالم را خریده باشد چقدر است.

۶ ــ اگر احتمال وجود پسر و دختر مساوی باشد احتمال وجود پسران و دختران را در خانواده‌هائی که دارای سه فرزند هستند پیدا کنید (به‌صورت جدول توزیع احتمال) و نمودار آن را نیز رسم کنید.

۷ ــ ۲۰۰ مسافر برای هواپیما جارزو کرده‌اند. اگر احتمال نیامدن مسافری کـه جارزو کرده است طبق تجارب گذشته ۰/۰۱ باشد احتمال اینکه ۳ نفر نیایند چقدر است؟

۸ ــ الف) شورای شهری مرکب از شهردار و شش عضو شورا است، راجع به‌هر موضوع طرح‌شده با اکثریت آرای این هفت نفر تصمیم‌گرفته می‌شود. فرض کنید که شهردار می‌خواهد

پیشنهاد معینی را به‌تصویب‌برساند اما ازحمایت شورا ازآن پیشنهاد یقین ندارد . همچنین فرض کنید که شش عضو ، مستقلا" ، هرکدام با احتمال ۴۰٪ به‌نفع پیشنهاد مذکور رای مثبت دهند . چقدر احتمال دارد که این پیشنهاد تصویب شود ؟

ب) فرض کنیم شهر دارد و متحد پا برجا دارد که مطمئن است به‌پیشنهاد رای مثبت می‌دهند این امر شانس تصویب پیشنهاد را چقدر افزایش خواهد داد ؟ (فرض کنید که چهار عضو دیگر مثل قسمت (الف) عمل کنند) .

ج) فرض کنید که همه شش عضو مثل قسمت (الف) رای می‌دهند با این تفاوت‌که دو نفر ازآنها دوستان شهردارهستند و قبل‌از اجلاس با او مذاکره‌و توافق می‌کنند که‌درداخل گروه خودشان رای دهند تا موضع اکثریتشان تعیین گردد و سپس با سه رأی یکپارچه‌و متفق برای حمایت‌ازآن موضع به‌اجلاس شورا بروند . آیا این موضوع شعار " اتحاد ،پیروزی است و تفرقه ، شکست است " را توجیه می‌کند یا نمی‌کند ؟

۹ ــ اگر تعداد مراجعین به‌بانک به‌طور متوسط در ساعت ۷۲ نفر باشد احتمال اینکه چهارنفر در سه دقیقه اول ساعت به‌بانک مراجعه کنند چقدر است ؟

۱۰ ــ ۵۰ لاستیک در انبار شرکتی وجود دارد که ۱۰ تای آن معیوب است .یک‌مشتری ۵ عدد از این لاستیک‌ها را می‌خرد .توزیع احتمال‌را برای تعدادلاستیکهای سالم‌خریداری شده به‌وسیله مشتری به‌دست آورید .

۱۱ ــ نمره امتحان انگلیسی دو دانشجو بر حسب واحدهای استاندارد به‌ترتیب ۰/۸ و ۰/۴ می‌باشد . اگر نمره آنها به‌ترتیب ۸۸ و ۶۴ بوده و توزیع نیز نرمال باشد میانگین و انحراف معیار نمره‌ها را پیدا کنید .

۱۲ ــ در امتحان نهائی‌درس ریاضی ، میانگین‌نمره‌ها ۷۲ و انحراف معیار آن ۱۵شده است .

الف : نمره‌های ۶۰، ۹۳، ۷۲ را بر حسب واحدهای نرمال استاندارد تعیین کنید .

ب : نمره‌های ۱ـ، ۱/۶ را که برحسب استانداردهستند برحسب تابع نرمال فوق تعیین کنید .

۱۳ ــ میانگین قد ۵۰۰ پسر پانزده ساله ۱۵۱ سانتیمتر و انحراف معیارآن ۱۵می‌باشد با فرض اینکه توزیع‌قد آنها نرمال باشد موارد زیررا محاسبه کنید در صورتیکه تمام‌اندازه‌ها به‌نزدیکترین سانتیمتر روند شده باشند .

الف : قد چند دانشجو بین ۱۲۰ و ۱۵۵ سانتیمتر است ؟

ب : قد چند دانشجو بیش‌از ۱۸۵ سانتیمتر است ؟

ج : قد چند دانشجو کمتراز ۱۲۸ سانتیمتر است ؟

د : قد چند دانشجو مساوی ۱۲۸ سانتیمتر است ؟

هـ : قد چند دانشجو مساوی یا کمتراز ۱۲۸ سانتیمتر است ؟

۱۴ ــ میانگین قطرداخلی ۲۰۰ واشر نمونه تولیدی یک ماشین ۰/۵۰۲ اینچ وانحراف معیارآن ۰/۰۰۵ اینچ است .

با توجه به‌استفاده‌ای که برای واشرها در نظر گرفته شده ، ماکزیمم قطر قابل قبـول از ۰/۴۹۶ تا ۰/۵۰۸ اینچ است در غیر این صورت واشرها معیوب تلقی می‌گردد .با فرض این که توزیع قطر واشرها نرمال است درصد واشرهای معیوب را تعیین کنید .

۱۵ ــ یک امتحان چند جوابی دارای ۳۰۰ سئوال است که فقط یکی از چهار جواب هر سئوال صحیح است . دانشجویی می‌خواهد به ۸۰ سئوال این امتحان که هیچ اطلاعی از آن ندارد صرفا" از روی تصادف پاسخ درست دهد . مطلوب است احتمال اینکه این دانشجو از این ۸۰ سئوال تعداد ۲۵ تا ۳۰ سئوال آنرا جواب درست بدهد .

۱۶ ــ مقدار یا مقادیر Z را در هریک از سه‌حالت زیر با توجه به‌سطح زیر منحنی‌نرمال تعیین کنید .

الف) سطح بین صفر و Z عبارت است از ۰/۳۷۷۰

ب) سطح سمت چپ Z عبارت است از ۰/۸۶۲۱

ج) سطح بین ۱/۵ــ و Z عبارت است از ۰/۰۲۱۷

۱۷ ــ در یک جریان تولید ، ۱۰ درصد اقلام دارای نقص فنی می‌شوند . اگر ۱۰۰ قلم کالا به‌طور تصادفی از جریان تولید انتخاب شوند ، احتمال اینکه تعداد معیوبها بیش از ۱۳ قلم باشد ، چقدر است ؟

حل مسائل فرد فصل سوم

۱ـ

$$\sigma_x^2 = E[X-E(X)]^2 = E(X^2) - \mu^2$$

$$E(X^2) = \sum_{x=0}^{n} x^2 P_x = \sum_{x=0}^{n} x^2 P_x = \sum_{x=0}^{n} x^2 \cdot \frac{n!}{x!(n-x)!} \cdot P^x q^{n-x}$$

داریم $[\ x^2 = x(x-1)+x\ ]$ که در رابطه فوق می‌گذاریم .

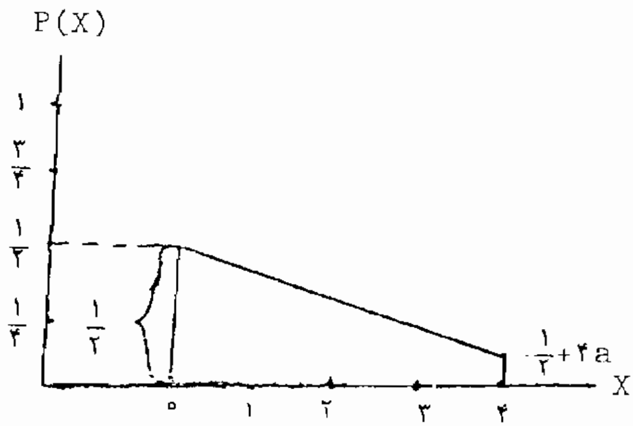
$$E(X^2) = \sum_{x=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} P^x q^{n-x} + \sum_{x=0}^{n} x \cdot \frac{n!}{x(n-x)!} P^x q^{n-x} =$$
$$\sum_{x=2}^{n} x(x-1) \cdot \frac{n!}{x!(n-x)!} P^x q^{n-x} + \mu = n(n-1)P^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)} P^{x-2} q^{n-x} + \mu$$

فرض کنیم $m = n-2$ و $y = x-2$ باشند ، داریم ،

$$E(X^2) = n(n-1)P^2 \sum_{y=0}^{m} \binom{m}{y} P^y q^{m-y} + \mu = n(n-1)P^2 (P+q)^m + \mu = n(n-1)P^2 + nP$$

$$\sigma_x^2 = E(X^2) - \mu^2 = n(n-1)P^2 + nP - (nP)^2 = nP - nP^2$$

۳ ـ الف) نمودار $P_x = \frac{1}{2} - aX$ خط مستقیمی است که به‌صورت تقریبی در شکل نشان‌داده است .



شکل (۱۰ ـ ۳)

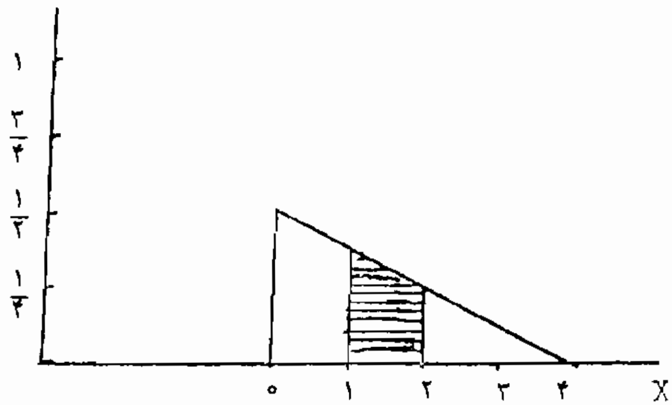
می‌دانیم که در تابع چگالی سطح زیر منحنی باید برابر ۱ باشد و از آنجا که این سطح در شکل فوق ذوذنقه است پس سطح ذوذنقه باید برابر ۱ باشد .

سطح ذوذنقه = (نصف ارتفاع) (مجموع دو قاعده) = $(\frac{1}{2}\times 4)(\frac{1}{2}+\frac{1}{2}-4a)=1$

سطح ذوذنقه = $2\times(1-4a)=1$

$a=\frac{1}{8}$

بنابراین‌قاعده دوم‌ذوذنقه($\frac{1}{2}-4a$) برابرصفر شده و شکل دقیق منحنی‌تابع چگالی‌بصورت زیر در می‌آید .



شکل (۱۱ − ۳)

ب) $P(1<X<2)$ برابر است با سطح زیر منحنی در فاصله بین $X=1$ و $X=2$ که بصورت ذوذنقه‌ای در شکل هاشور زده شده است . برای محاسبه ، بازهم از طریق فرمول تابع چگالی دو قاعده‌را بدست آورده و از طریق محاسبه مساحت ذوزنقه سطح زیر منحنی درفاصله‌مزبور را محاسبه می‌کنیم .

$$P(X)=\frac{1}{2}-\frac{1}{8}X$$

$$P(1)=\frac{1}{2}-\frac{1}{8}=\frac{3}{8}$$

$$P(۲) = \frac{۱}{۲} - \frac{۲}{۸} = \frac{۱}{۴}$$

$$\text{سطح ذوذنقه سایه‌دار} = P(۱ < X < ۲) = \frac{۱}{۲}(۱)\left(\frac{۳}{۸} + \frac{۱}{۴}\right) = \frac{۵}{۱۶}$$

۵ــ چون بین ۵۰۰ لاستیک ۴۵۰ تا سالم است پس احتمال سالم بودن $۰/۹ = \frac{۴۵۰}{۵۰۰}$ است. وفرض‌می‌کنیم این مقدار احتمال طی ده‌بار آزمایش ثابت باشد.

$n = ۱۰ \qquad P = ۰/۹ \qquad q = ۰/۱ \qquad x = ۸$

$$P(X=x) = \binom{n}{x} P^x q^{n-x}$$

$$P(X=۸) = \binom{۱۰}{۸}(۰/۹)^۸(۰/۱)^۲ = \frac{۱۰!}{۸!\,۲!}(۰/۹)^۸(۰/۱)^۲ = ۰/۱۹۴$$

۷ــ در این مسئله "نیامدن" موفقیت است. از طرفی بر طبق حساب سر انگشتی ما مشاهده می‌شود که $n = ۲۰۰$ به‌اندازه کافی بزرگ و $p = ۰/۰۱$ یعنی احتمال موفقیت به‌اندازه کافی کوچک است.

بنابراین داریم:

$$\lambda = \mu$$

$$P(X=x) = \frac{e^{-\mu}\mu^x}{x!} \qquad \lambda = nP = ۲۰۰ \times ۰/۰۱ = ۲$$

$$P(۳) = \frac{(۲/۷۱۸۲۸)^{-۲} \times ۲^۳}{۳!} = \frac{(۰/۱۳۵۳)(۸)}{۶} = ۰/۱۸۰۴$$

۹ــ ابتدا متوسط مراجعات در سه دقیقه را محاسبه و بعد مسئله را بااستفاده از توزیع پواسن حل می‌کنیم.

$$\mu = \lambda = ۷۲$$

$$t = \text{سه‌دقیقه} \Rightarrow t = \frac{۳}{۶۰} = \frac{۱}{۲۰}$$

$$x = ۴$$

$$\lambda t = ۷۲\left(\frac{۱}{۲۰}\right) = ۳/۶$$

$$P(۴) = \frac{e^{-۳/۶}\,۳/۶^۴}{۴!} = ۰/۱۹۱$$

۱۱ ــ با استفاده از فرمول $\mu + \sigma Z = x$ و قرار دادن اطلاعات صورت مسئله‌در این فرمول مسئله به‌صورت دو معادله و دو مجهول درآمده و مشخصه‌ها محاسبه می‌گردند .

$$\begin{cases} ۸۸ = \mu + ۰/۸\ \sigma \\ ۶۴ = \mu + ۰/۴\ \sigma \end{cases} \longrightarrow \begin{cases} \mu = ۴۰ \\ \sigma = ۶۰ \end{cases}$$

۱۳ ــ الف) با فرض روند شدن اعداد ، فاصله ۱۵۵ ــ ۱۲۰ در حقیقت فاصله ۱۵۵/۵ ــ ۱۱۹/۵ بوده است . لذا این دو عدد را برحسب نرمال استاندارد محاسبه می‌کنیم .

$$Z_۱ = \frac{۱۱۹/۵ - ۱۵۱}{۱۵} = -۲/۱۰$$

$$Z_۲ = \frac{۱۵۵/۵ - ۱۵۱}{۱۵} = ۰/۳۰$$

(سطح بین $Z_۱ = -۲/۱۰$ و $Z_۲ = ۰/۳۰$) = نسبت دانشجویان مورد نظر

$= (-۲/۱ < Z < ۰$ سطح مربوط به$) + (۰ < Z < ۰/۳)$ =

$= ۰/۴۸۲۱ + ۰/۱۱۷۹ = ۰/۶۰۰۰$

بنابراین تعداد دانشجویانی که قدشان در فاصله فوق قرار دارد عبارتست از

۵۰۰(۰/۶۰۰۰) = ۳۰۰ نفر

ب) قد بیش از ۱۸۵ به‌معنی بزرگتر از ۱۸۵/۵ می‌باشد لذا ۱۸۵/۵ را تبدیل به‌استاندارد می‌کنیم .

$$Z = \frac{۱۸۵/۵ - ۱۵۱}{۱۵} = ۲/۳۰$$

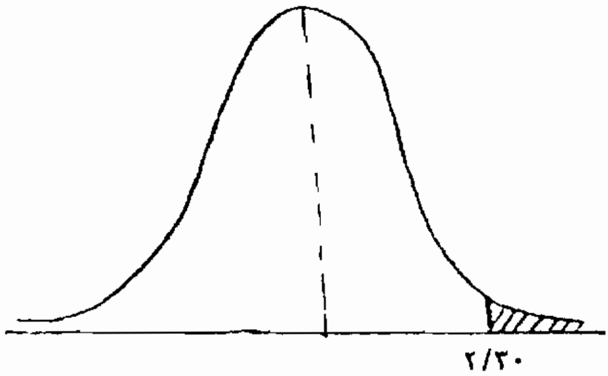
نسبت دانشجویان مورد نظر این فرض برابر سطح هاشور زده زیر منحنی‌خواهدبودکه مساوی سطح زیر منحنی سمت راست نقطه $Z = ۲/۳۰$ می‌باشد .

(سطح مربوط به $Z > ۰$) = (سطح مربوط به $Z > ۲/۳۰$) = نسبت مورد نظر دانشجویان

$-($ سطح مربوط به $۰ < Z < ۲/۳۰)$

$= ۰/۵ - ۰/۴۸۹۳ = ۰/۰۱۰۷$

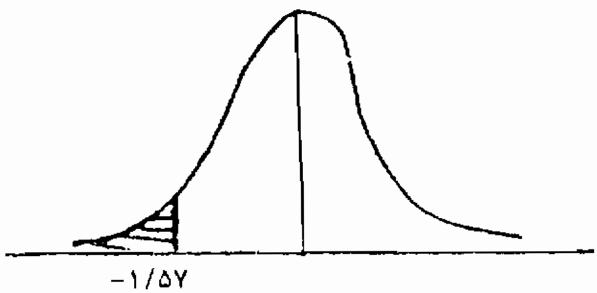
۵ = (۰/۰۱۰۷)۵۰۰ = تعداد دانشجویان با قد بزرگتر از ۱۸۵



شکل(۱۲-۳)

ج) قد کمتر از ۱۲۸ در واقع بمعنی قد کمتر از ۱۲۷/۵ است .

$$Z = \frac{\text{۱۲۷/۵} - \text{۱۵۱}}{\text{۱۵}} = -\text{۱/۵۷}$$



شکل (۱۳-۳)

از جدول داریم :

نسبت دانشجویان مورد نظر

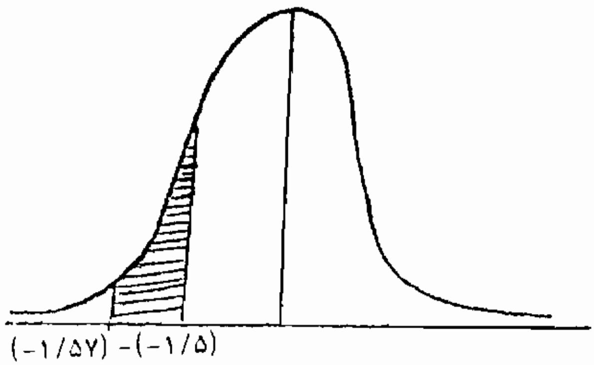
P (Z < −۱/۵۷) = (۰/۵) − (سطح مربوط به ۰ < Z < −۱/۵۷) = ۰/۵ − ۰/۴۴۱۸ = %۵۸۲

تعداد دانشجویان موردنظر ۲۹ = (%۵۸۲) ۵۰۰

د) قد مساوی ۱۲۸ بمعنی کوچکتر از ۱۲۸/۵ و بزرگتر از ۱۲۷/۵ می‌باشد .

$$Z_۱ = \frac{۱۲۸/۵ - ۱۵۱}{۱۵} = -۱/۵$$

$$Z_۲ = \frac{۱۲۷/۵ - ۱۵۱}{۱۵} = -۱/۵۷$$



شکل (۱۴-۳)

=(سطح مربوط به $۰ > Z > -۱/۵$)–(سطح مربوط به $۰ > Z > -۱/۵۷$) =

نسبت دانشجویان مورد نظر $= p\,(-۱/۵۷ < Z < -۱/۵)$

$۰/۴۴۱۸ - ۰/۴۳۳۲ = ۰/۰۰۸۶$

تعداد دانشجویان مورد نظر $۵۰۰\,(۰/۰۰۸۶) = ۴$

هـ) این فرض را از دو طریق می‌توان حل کرد طریق اول بدین ترتیب است که تعداد دانشجویانی را که قدشان کمتراز ۱۲۸/۵ است حساب کنیم (مثل بند ج) و طریق دوم آن است که نتایج بند ج و د را با هم جمع کنیم که جهت سهولت از همین راه عمل می‌کنیم

(تعداد دانشجویان بین ۱۲۷/۵ و ۱۲۸/۵)+(تعداد دانشجویان کمتراز ۱۲۷/۵)

$$۲۹ + ۴ = ۳۳$$

۱۵ ــ احتمال یک جواب صحیح برای هریک از ۸۰ سئوال برابر $p = \frac{۱}{۴}$ است . اگر X تعداد جوابهای صحیح را که برمبنای حدس است ،نشان دهد ،دراین صورت خواهیم داشت .

$$P(۲۵ < X < ۳۰) = \sum_{x=۲۵}^{۳۰} b\left(x : ۸۰ \text{ و } \frac{۱}{۴}\right)$$

$$\mu = np = ۸۰\left(\frac{۱}{۴}\right) = ۲۰$$

$$\sigma = \sqrt{nPq} = \sqrt{۸۰\,(\frac{۱}{۴})(\frac{۳}{۴})} = ۳/۸۷$$

حال سطح بین $X_۱ = ۲۴/۵$ و $X_۲ = ۳۰/۵$ را لازم داریم

$$Z_۱ = \frac{۲۴/۵ - ۲۰}{۳/۸۷} = ۱/۱۶۳$$

$$Z_۲ = \frac{۳۰/۵ - ۲۰}{۳/۸۷} = ۲/۷۱۳$$

$$P\,(۲۵ < X < ۳۰) = P\,(۱/۱۶۳ < Z < ۲/۷۱۳)$$

$$P = (Z < ۲/۷۱۳) - P\,(Z < ۱/۱۶۳) = ۰/۹۹۶۷ - ۰/۸۷۷۶ = ۰/۱۱۹۱$$

۱۷ ـ تعداد کالاهای معیوب دارای توزیع بینومیال یا دوجمله‌ای با پارامترهای $n=۱۰۰$ و $p = ۰/۱$ می‌باشد . اما چون حجم نمونه بزرگ است می‌توان نتایج نسبتاً" دقیقی را با به کاربردن تقریب منحنی نرمال با:

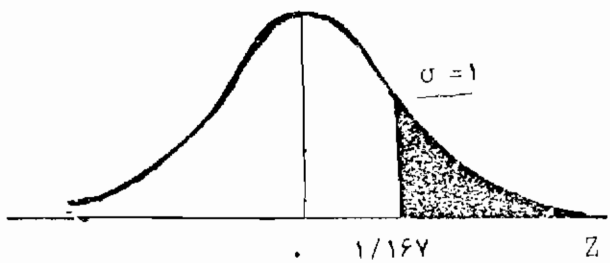
$$\mu = nP = ۱۰۰\,(۰/۱) = ۱۰$$

$$\sigma = \sqrt{nPq} = \sqrt{۱۰۰\,(۰/۱)(۰/۰۹)} = ۳$$

به‌دست آورد . برای به‌دست آوردن احتمال مورد نظر ، باید سطح سمت راست $X=۱۲/۵$ را پیدا کنیم ، که در این صورت Z مربوط به این تعداد عبارت خواهد بود از :

$$Z = \frac{۱۲/۵ - ۱۰}{۳} = ۱/۱۶۷$$

$$P\,(X > ۱۲) = \sum_{X=۱۳}^{۱۰۰} b(X : ۱۰۰ \text{ و } ۰/۱) \approx P(Z > ۱/۱۶۷) = ۱ - P\,(Z < ۱/۱۶۷) = ۱ - ۰/۸۷۸۴ = ۰/۸۲۱۶$$



شکل ۱۵-۳

پاسخ مسائل زوج فصل سوم

۴- الف $a=\frac{۱}{۴۸}$ ب $\frac{۷}{۲۴}$ ج $\frac{۳}{۴}$ د $\frac{۱}{۶}$

۶ ــ چون احتمال وجودپسران بااحتمال وجود دختران مساوی هم می‌باشند بنابراین جدول مربوط به هردو حادثه مشابه بوده و جهت رعایت اختصار جدول مربوط به پسران نوشته می‌شود .

X = تعدادپسران	۰	۱	۲	۳
P_x	$\frac{۱}{۸}$	$\frac{۳}{۸}$	$\frac{۳}{۸}$	$\frac{۱}{۸}$

۸ ــ الف ــ ٪۴۶ ب ــ ٪۸۷ ج ــ توجیه می‌کند .

-۱۰

x =تعدا لاستیکهای سالم	P_x
۰	۰/۰۰۰۱
۱	۰/۰۰۴۰
۲	۰/۰۴۴۲
۳	۰/۲۰۹۸
۴	۰/۴۳۱۳
۵	۰/۳۱۰۶

۱۲ ــ الف ــ به‌ترتیب ۰/۸ ، ۱/۴ و ۰ ب ــ ۵۷ و ۹۶

۱۴ ــ ٪۲۳

۱۶ ــ الف ۱/۹۶± ب ــ ۱/۰۹ ج ــ (۱/۳۵−) و (۱/۶۹−)

فصل چهارم

توزیع احتمال نمونه‌گیری

مقدمه :

مطابق معمول در سه فصل گذشته ،در مقدمه این فصل نیز درمورد مفهوم توزیع احتمال نمونه‌گیری و ضرورت بحث پیرامون آن و نیز روند بحث فصل چهارم گفتگو خواهد شد .

تمامی زحماتی که یک آمارگر متحمل می‌شود برای آن است که بتواند حتی‌الامکان نتایج دقیقی را در مورد جامعه مورد مطالعه‌اش بدست آورد . اگر کمی دقت کنیم در خواهیم یافت که هریک از اهداف سه‌گانه‌ای که در مقدمه کتاب برای عملیات آماری شمرده شدند به نحوی با مشخصه‌های جامعه درارتباط هستند . در مواردی محاسبه مقدار این مشخصه‌ها هدف بوده و در موارد دیگری نیز بدون آنکه محاسبه این مقادیر ضرورت داشته باشد لازم است محقق نتایجی را در مورد این مشخصه درجامعه بدست آورد که آزمون تساوی میانگین و یا انحراف معیار دو یا چند جامعه با یکدیگر نمونه‌هائی از این موارد می‌باشند .

در قسمتهای قبل ملاحظه گردید که برای تحقق اهداف مزبور معمولا" نمونه‌گیری بعمل آمده و پس ازثبت مشاهدات مربوط به‌آن باانجام عملیاتی برروی این مشاهدات ،مشخصه‌های مربوط به نمونه‌را به‌دست می‌آورند . اما تا این مرحله از عملیات هنوز هیچ نتیجه‌ای درمورد جامعه بدست نیامده است و به صرف داشتن مقدار مشخصه‌های یک یا چند نمونه ازیک‌جامعه نمی‌توان در مورد خود جامعه هیچ‌گونه قضاوت قابل اعتمادی را بعمل آورد مگر آنکه برروی مشخصه‌های مربوط به‌نمونه عملیات دیگری انجام گیرد که بر مبنای آن بتوان در موردجامعه نتایجی را بدست آورد . به‌عبارت دیگر باید از اطلاعات مربوط به‌نمونه که اینک در اختیار ما قراردارند به‌عنوان مواد اولیه پلی استفاده‌کرد که این پل نمونه را به‌جامعه‌متصل می‌سازد .

به‌نظر نمی‌رسد که ساختن پل ارتباطی فوق‌الذکر با موادی که اکنون در اختیارند ، مقدور باشد و ظاهرا" برای انجام این کار هنوز همه مواداولیه و به‌عبارت بهتر همه‌اطلاعات لازم در دسترس نیستند . ثمره همه تفکرات و تلاشهائی که برای تکمیل این اطلاعات بعمل

آمده‌است معمولا" در کتب آماری تحت‌عنوان توزیع احتمال نمونه‌گیری بیان می‌شودکه موضوع بحث این فصل می‌باشد و بنابراین توزیع احتمال نمونه‌گیری تکمیل کننده موادی می‌باشدکه برای ساختن پل رابط بین نمونه و جامعه لازم هستند و انشاءالله در فصول بعدی از ترکیب این مواد پلهای مزبور زده شده و خواهیم دید که چگونه با استفاده از این اطلاعات‌نتایج دلخواه در مورد جامعه ، یکی پس از دیگری به‌صورتی نسبتا" قابل اعتماد بدست می‌آیند .

برای بیان مفهوم توزیع احتمال نمونه‌گیری توجه به‌این نکته لازم است‌که از یک‌جامعه معمولا" می‌توان نمونه‌های‌متعددی را اخذ نمود . به‌عنوان‌مثال از یک جامعه ۵ عضوی می‌توان $C_{5}^{2}=10$ نمونه دوتائی را بدون جایگذاری انتخاب نمود . واضح است که مقدار مشخصه‌های هریک از این نمونه‌ها الزاما" با مقادیر مشخصه‌های سایر نمونه‌ها مساوی نبوده و به احتمال بسیارزیاد متفاوت خواهد بود . به‌عنوان مثال مقادیر میانگین ده نمونه مثال فوق می‌توانند به‌صورت زیر باشند.

(۳و ۵/۱و ۵/۳و ۲و ۳و ۱و ۱و ۵/۱و ۲و ۲)

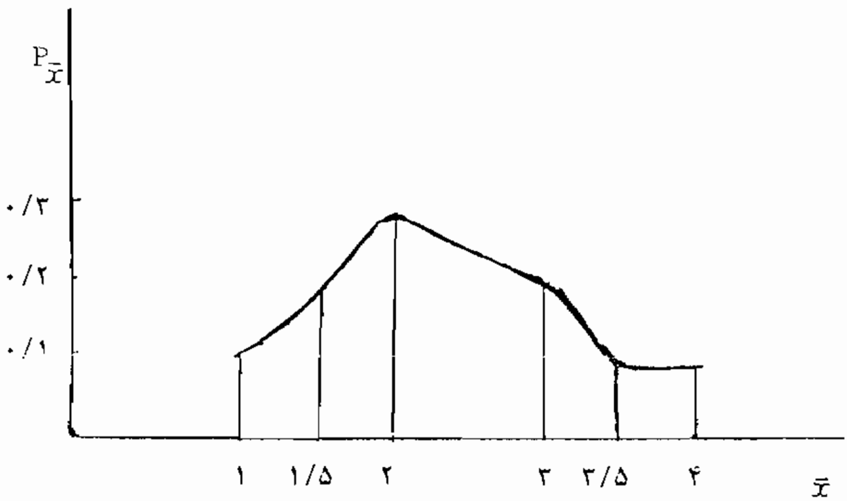
ازآنجا که معمولا" یکی از شرایط اصلی نمونه‌گیری ،انتخاب تصادفی آن می‌باشد مسلما" مقدار مشخصه‌های هر نمونه نیز به‌صورت تصادفی تعیین می‌گردد و به‌عبارت دیگر هریک از مشخصه‌های مربوط به نمونه متغیری تصادفی می‌باشند . در مثال بالا میانگین نمونه‌های دوتائی متغیری بود که ده مقدار مزبور را به‌طور تصادفی اختیار نمود . همان طور کەدرمثال فوق ملاحظه می‌شود مقدار میانگین نمونه ($\bar{X}$) از هر نمونه به‌نمونه دیگر قابل تغییر است و لذا بر آن لفظ متغیر اطلاق شده و چون به تبع انتخاب تصادفی نمونه‌ها ، این مقادیر به‌صورت تصادفی اختیار می‌شوند بر میانگین مذکور عنوان متغیر تصادفی اطلاق می‌گردد .

طبیعتا" اطلاق عنوان متغیر تصادفی بر هریک از مشخصه‌های نمونه ، این نتیجه را در پی خواهد داشت که‌آن مشخصه دارای یک توزیع احتمال خاص باشد زیرا هر متغیرتصادفی دارای یک توزیع احتمال است . از آنجا که توزیع‌های احتمال مربوط به‌مشخصه‌های نمونه ، در اثر نمونه‌گیری ایجاد می‌شوند به‌این توزیع‌ها ، توزیع احتمال نمونه‌گیری گفته می‌شود . برای مثال توزیع احتمال مربوط به‌میانگین نمونه‌های ده گانه فوق ، ذیلا" به‌صورت جدول و منحنی ترسیم می‌شوند .

اگر یک محقق مقدار مشخصه‌های مربوط به نمونه را در اختیار داشته باشد و علاوه بر آن شکل توزیع احتمال آن مشخصه‌ها را نیز بداند می‌تواند با این اطلاعات نتیجه دلخواه خود را در مورد جامعه بدست آورد . معمولا" در عملیات آماری ، از بین مشخصه‌های‌مختلف مربوط به یک جامعه محقق به‌دنبال اخذ نتیجه در مورد میانگین و انحراف معیار آن‌می‌باشد

$\bar{x}_i$	F_i	$P_i = \frac{F_i}{n}$
۱	۱	۰/۱
۱/۵	۲	۰/۲
۲	۳	۰/۳
۳	۲	۰/۲
۳/۵	۱	۰/۱
۴	۱	۰/۱
	=۱۰	

جـــدول ۱-۴



شـــکل ۱-۴

و چنانکه بعداً" خواهیم دید در اکثریت قریب به‌اتفاق موارد رسیدن به‌این هدف‌مستلزم‌آن است که او اولاً" مقدار همین مشخصه‌ها را در مورد حداقل یک یا دو نمونه در اختیار داشته باشد ثانیاً" شکل توزیع احتمال نمونه‌گیری این دو مشخصه را نیز بشناسد . چگونگی محاسبه مقدار مشخصه‌های نمونه از مباحث مربوط به‌فصل دوم بوده و قطعاً" خواننده عزیز اکنون با آن به‌قدرکافی آشنا شده‌است و شناسائی‌اشکال مختلف توزیع نمونه‌گیری دو مشخصه میانگین و انحراف معیار نمونه نیز موضوع بحث این فصل می‌باشد .

خوشبختانه توزیعهای احتمال نمونه‌گیری میانگین و انحراف معیار نمونه از قانونمندی نسبتاً" دقیق و ثابتی برخوردارند بدین معنی که اشکال این توزیع ها محدود بوده‌ودرچهار توزیع نرمال ، t دانشجو[1] ، کای اسکویر (χ^2)[2] و توزیع فیشر (F)[3] خلاصه می‌شوند که از این چهار توزیع ، دو توزیع نرمال و دانشجو مربوط به میانگین نمونه‌ها بوده و دو توزیع دیگر یعنی (χ^2) و (F) نحوه توزیع احتمال واریانس و یا انحراف معیار نمونه‌هارا توضیح می‌دهند .

بنابراین ادامه‌بحث در این فصل به‌بررسی توزیعهای احتمالی که به‌تعمیم نتایج‌نمونه به‌جامعه کمک می‌نمایند اختصاص خواهد داشت . در این بررسی ابتدا هریک از توزیعهایی که پیش ازاین آمده مستقلاً" توضیح داده شده‌و سپس چگونگی‌ارتباط‌آن توزیع‌با مشخصه‌های نمونه در قالب قضایایی بیان می‌گردد . روند دقیق ادامه بحث به‌قرار زیر است .

الف ــ توزیعهای احتمال مربوط به میانگین نمونه (شامل دو توزیع نرمال و t دانشجو یا t استیودنت)

ب ــ توزیعهای احتمال مربوط به انحراف معیار نمونه (شامل دو توزیع χ^2 و F)

الف ــ توزیعهای احتمال مربوط به میانگین نمونه‌ها

میانگین یک یا چند جامعه معمولاً" به‌دو صورت مختلف هدف عملیات آماری قرار می‌گیرد . در یکی از این دو صورت بدست آوردن مقدار تقریبی میانگین جامعه (μ) هدف عملیات است ، در حالی که در صورت دوم مقایسه مقدار میانگین دو یا چند جامعه هدف و غایت عملیات بوده و بدست آوردن مقدار میانگین هریک از آن جوامع ضرورتی ندارد . معمولاً" وقتی که مورد دوم مقصود محقق بوده و او به‌دنبال مقایسه مقادیر میانگین چند جامعه با یکدیگر است تفاضل میانگین های جامعه ($\mu_1 - \mu_2$) را به‌عنوان هدف عملیات در نظر گرفته و سعی می‌کند از طریق تخمین مقدار آن ، بین میانگین دو جامعه مقایسه‌بعمل آورد . چنین تخمینی براساس این استدلال استوار است که اگر ($\bar{X}_1$) که میانگین نمونه مأخوذه از جامعه اول است ، متغیری تصادفی بوده و ($\bar{X}_2$) نیز به‌عنوان میانگین نمونه‌اخذ شده‌از جامعه‌دوم‌متغیری تصادفی باشد قهراً"تفاضل‌این‌دو متغیرتصادفی یعنی ($\bar{X}_1 - \bar{X}_2$) نیز متغیری تصادفی بوده و مطمئناً"این‌متغیر تصادفی جدید نیز دارای یک توزیع احتمال می‌باشد که اگر مقدار متغیر ($\bar{X}_1 - \bar{X}_2$)را نیز با توزیع احتمال نمونه‌گیری آن درآمیزیم

---

1 - Student

۲ ــ به‌این توزیع کای دو ، چی دووخی دو نیز گفته می‌شود .

3 - Fisher

مطمئناً" می‌توان به‌طریقی مقدار ($\mu_{۲} - \mu_{۱}$) را به‌طور تقریبی بدست آورد درست همان‌طور که با استفاده از مقدار میانگین نمونه ($\bar{x}$) توزیع احتمال نمونه‌گیری آن، مقدار میانگین جامعه (μ) به‌صورت تقریبی قابل محاسبه است.

بنابراین کسی که‌می‌خواهد در مورد میانگین جامعه تحقیق کند یا بدست آوردن‌مقدار میانگین جامعه (μ) هدف او بوده و یا تفاضل میانگین دو یا چند جامعه ($\mu_{۱} - \mu_{۲}$) را به‌عنوان هدف و غایت عملیات انتخاب می‌کند، هریک از این دو پارامتر که هدف او باشند معمولاً او باید مقدار متغیر مشابه آن یعنی ($\bar{X}$) و یا ($\bar{X}_{۱} - \bar{X}_{۲}$) را برای حداقل یک نمونه بدست آورد و با استفاده از توزیع نمونه‌گیری همان متغیر به‌صورتی که در فصول بعد خواهند آمد در مورد پارامتر هدف در جامعه استنتاج نماید. چنان که قبلاً" گفته‌شد تعداد توزیعهای احتمال نمونه‌گیری مربوط به میانگین که برای تعمیم نتایج نمونه به جامعه می‌توانند بما کمک کنند محدودبوده و از دوتوزیع فراتر نمی‌روند. در بسیاری از موارد توزیع نمونه‌گیری نرمال بوده و شرایط مسئله به ما اجازه می‌دهد تا با استاندارد کردن آن از توزیع (Z) به عنوان توزیع نمونه‌گیری استفاده کنیم. از آنجا که خواننده گرامی با توزیع (Z) آشنائی‌کامل‌دارد در اینجا از توضیح مجدد پیرامون این توزیع احتمال خودداری نموده و صرفاً" به قضایایی که به‌نحوی با استفاده از توزیع (Z) بین جامعه و نمونه ارتباط برقرار می‌کنند اکتفا می‌شود. در مواردی که به‌دلایلی استفاده از توزیع (Z) به‌عنوان توزیع احتمال رابط بین نمونه و جامعه امکان‌پذیر نیست، از توزیع t دانشجو استفاده می‌شود که ابتدا به‌طور خلاصه این توزیع احتمال توضیح داده شده و سپس چگونگی ارتباط بین نمونه و جامعه از طریق‌این‌توزیع احتمال در قالب قضایایی توضیح داده می‌شود. بنابراین بحث بند الف در ۳ عنوان کلی زیر ارائه خواهد شد.

الف - ۱ - قضایایی که با استفاده از توزیع Z بین (μ) و متغیر تصادفی ($\bar{X}$) رابطه بر قرار می‌کنند.

الف - ۲ - قضایایی که با استفاده از توزیع Z بین ($\mu_{۱} - \mu_{۲}$) و متغیر تصادفی ($\bar{X}_{۱} - \bar{X}_{۲}$) رابطه برقرار می‌کنند.

الف - ۳ - توضیحات کلی پیرامون توزیع t و ارتباط بین میانگین نمونه و جامعه از طریق این توزیع.

الف - ۱ - رابطه بین μ و $\bar{x}$ با استفاده از توزیع Z

در فصل قبل دیدیم که هرگاه مقدار میانگین متغیری مثل (X) را که دارای توزیع نرمال‌است از آن کم نموده و به‌انحراف معیار آن تقسیم کنیم، آن متغیر به‌متغیر نرمال استاندارد (Z) تبدیل می‌گردد. اساس استدلال در ارتباط موضوع این بند که در فصول

آینده مورد بهره‌برداری قرار می‌گیرد نیز همین مسئله است . توضیح آنکه همان طور که در مقدمه همین فصل گفته شد میانگین نمونه ($\bar{X}$) یک متغیر تصادفی است و مقدار آن از یک نمونه به‌نمونه دیگر متفاوت است . این میانگین به‌نوبه خود از یک شکل توزیع برخورداربوده و دارای یک میانگین و یک انحراف معیار می‌باشد . واضح است که اگر مقدار میانگین ($\bar{X}$) از میانگین میانگین $E(\bar{X})$ کم شده و حاصل تفریق بر انحراف معیار آن ($\sigma_{\bar{X}}$) تقسیم گردد ، متغیر حاصله توزیع نرمال استاندارد (Z) را خواهد داشت .

$$Z = \frac{\bar{x} - E(\bar{X})}{\sigma_{\bar{X}}}$$

همچنان که بعدا" خواهیم گفت $E(\bar{X}) = \mu$ می‌باشد یعنی میانگین میانگینهای نمونه‌های مأخوذه از یک جامعه با میانگین جامعه مساوی می‌باشد و لذا رابطه بالا به صورت زیر نیز نوشته می‌شود .

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{X}}} \qquad (۴ - ۱)$$

نتیجه‌ای که از بحث فوق گرفته می‌شود این است که اگر بدانیم میانگین نمونه ($\bar{X}$) توزیع نرمال داشته و انحراف معیار آن ($\sigma_{\bar{X}}$) نیز برای ما معلوم باشد می‌توان توزیع ($\bar{X}$) را به‌توزیع Z مبدل نموده و چون مقدار ($\bar{X}$) نیز معمولا" برای ما روشن است و مقادیر مختلف Z نیز در جداول مختلف محاسبه گردیده است ،مقدارتمامی عبارات رابطه (Z)فوق بجز (μ)مشخص شده و می‌توان با استفاده از این رابطه مقدار μ را بدست آورد . تعیین مقدار μ ازمباحث فصول آینده است و در اینجا فقط در مورد شکل توزیع $\bar{X}$ ، میانگین و انحراف معیار آن در موارد مختلف بحث می‌شود .

به‌طور کلی میانگین ($\bar{X}$) وضعیتی مشخص تر یا ثابت تر از شکل توزیع وانحراف معیار آن دارد و می‌توان در مورد آن یک حکم قطعی و عام صادر نمود . اما شکل توزیع و انحراف معیار($\bar{X}$)بستگی به‌عواملی همچون حجم جامعه ، حجم نمونه و شکل توزیع (X) در جامعه دارد که ذیلا" احکام مربوط به‌هریک از این سه مفهوم (میانگین ، انحراف معیار و شکل توزیع $\bar{X}$) در قالب قضایایی بیان می‌گردند .

قضیه (۱-۴) .هرگاه از یک جامعه به حجم N با میانگین (μ) تمام نمونه‌های n تایی ممکن اخذ شوند و میانگین هریک از این نمونه‌های n تایی با($\bar{x}_i$)نشان داده شوند میانگین

$\bar{x}$ ها با میانگین جامعه (μ) برابر خواهد بود .

$$E(\bar{X}) = \frac{\sum_{i=1}^{m} \bar{x}_i}{m} = \mu \qquad (۴-۲)$$

دراین رابطه (n)حجم هر نمونه و (m)تعداد نمونه‌های n تایی مأخوذه را نشان‌می‌دهد . واضح است که اگر نمونه‌گیری با روش بدون بازگردانی انجام گیرد ($m = C_N^n$) خواهدشد.

مثال ۱ ـ ۴ : از جامعه‌ای مشتمل بر ۴ عدد (۰، ۱، ۲، ۳) همه نمونه‌های ۲ تایی ممکن را بدون بازگردانی اخذ نموده و ثابت کنید که رابطه (۱ـ۴) در مورد میانگین نمونه‌ها صادق است .

پاسخ : تعداد نمونه‌هائی که ممکن است گرفته شوند ($m_1 = C_4^2 = ۶$) می‌باشد که عناصرهر نمونه و نیز میانگین آنرا در جدول زیر می‌نویسیم .

x_i	(۳و۲)و(۳و۱)و(۲و۱)و(۳و۰)و(۲و۰)و(۱و۰)
$\bar{x}_i$	۲/۵ و ۲ و ۱/۵ و ۱/۵ و ۱ و ۰/۵

جدول ۲-۴

$$E(\bar{X}) = \frac{\sum \bar{x}_i}{m} = \frac{۰/۵ + ۱ + ۱/۵ + ۱/۵ + ۲ + ۲/۵}{۶} = \frac{۹}{۶} = ۱/۵$$

حال میانگین جامعه (μ) را مستقیما" حساب می‌کنیم .

$$\mu = \frac{\sum x_i}{N} = \frac{۰+۱+۲+۳}{۴} = \frac{۶}{۴} = ۱/۵$$

قضیه (۱ـ۴) در مورد تمام جوامع با هرگونه توزیع احتمال صادق است .

حال پس‌از مشخص شدن وضعیت میانگین ، به‌بررسی شکل توزیع $\bar{X}$ می‌پردازیم .همان طور که قبلا" گفته شد شکل توزیع ($\bar{X}$) به‌سه عامل حجم نمونه (n) و چگونگی توزیع (X) در جامعه و نیز حجم جامعه (N) بستگی دارد .

قضیه ۲ـ۴ : هرگاه تمام نمونه‌های n تایی از جامعه‌ای که صفت (X) درآن دارای توزیع‌نرمال (یا تقریبا" نرمال)است اخذ شوند درآن‌صورت میانگین نمونه‌ها ($\bar{x}_i$)دارای توزیع نرمال یا تقریبا" نرمال خواهد بود .

همانطور که در قضیه (۲ـ۴) به‌خوبی مشهود است اگر توزیع (X) در جامعه نرمال باشد بدون توجه به‌حجم نمونه و حجم جامعه ($\bar{X}$) توزیع نرمال خواهد داشت .

مثال ۲ـ۴ : مقادیر صفت (X) در جامعه‌ای که ده عضو دارد مطابق جدول زیر دارای‌توزیع احتمال نرمال می‌باشند . اگر با روش بدون جایگذاری همه نمونه‌های ۹ تایی ممکن را از این جامعه اخذ کنیم ثابت کنید که قضیه (۲ـ۴) در مورد $\bar{x}$ ها صادق است .

x_i	۱	۲	۳	۴	۵
F_i	۱	۲	۴	۲	۱
$P_i=\frac{f_i}{N}$	۰/۱	۰/۲	۰/۴	۰/۲	۰/۱

جدول ۳-۴

پاسخ : تعداد نمونه‌هایی که ممکن است گرفته شوند برابر ($m = C_{10}^{9}=10$) می‌باشد . که به‌ترتیب هریک از نمونه‌ها و میانگین مربوطه در جدول (۴ـ۴) نوشته می‌شوند .

x_i	$\bar{x}_i$
۱،۲،۲،۳،۳،۳،۳،۴،۴	۲/۸
۱،۲،۲،۳،۳،۳،۳،۴،۵	۲/۹
۱،۲،۲،۳،۳،۳،۳،۴،۵	۲/۹
۱،۲،۲،۳،۳،۳،۴،۴،۵	۳
۱،۲،۲،۳،۳،۳،۴،۴،۵	۳
۱،۲،۲،۳،۳،۳،۴،۴،۵	۳
۱،۲،۲،۳،۳،۳،۴،۴،۵	۳
۱،۲،۳،۳،۳،۳،۴،۴،۵	۳/۱
۱،۲،۳،۳،۳،۳،۴،۴،۵	۳/۱
۲،۲،۳،۳،۳،۳،۴،۴،۵	۳/۲

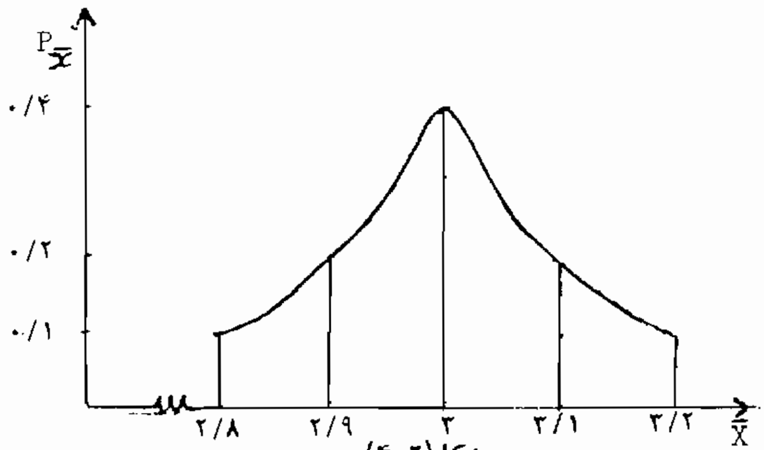
جدول (۴-۴)

حال جدول توزیع فراوانی و احتمال $\bar{x}$ ها را به‌صورت جدول (۵ ـ ۴) رسم می‌کنیم

$\bar{x}_i$	۲/۸	۲/۹	۳	۳/۱	۳/۲
F_i	۱	۲	۴	۲	۱
P_i	۰/۱	۰/۲	۰/۴	۰/۲	۰/۱

جـــدول (۵-۴)

منحنی توزیع احتمال $\bar{X}$ بر اساس جدول فوق به‌صورت زیر است که به‌وضوح یک‌منحنی توزیع احتمال نرمال بوده و صدق قضیه (۴/۲) ثابت می‌شود .



شکل(۲-۴)

معمولاً" از وضعیت توزیع احتمال صفت مورد نظر در جامعه اطلاعی در اختیار نیست‌و به‌همین دلیل از قضیه (۲-۴) نمی‌توان در اکثر موارد استفاده نمود . در اکثر قریب‌به‌اتفاق چنین مواردی قضیه زیر در مورد شکل توزیع می‌تواند مورد استفاده قرار گیرد .

قضیه ۳-۴ : هرگاه جامعه مورد مطالعه ما یکی از سه خصوصیت زیر را دارا باشد .

الف ــ حجم جامعه (N) بسیار بزرگ بوده و به‌تعبیر آمارگران جامعه‌ای نامتناهی و نا محدود ($n \geqslant ۳۰$) باشد .

ب ــ حجم جامعه (N) زیاد نباشد اما نمونه‌گیری با روش بازگردانی انجام گیرد .

ج ــ حجم جامعه کوچک بوده و نمونه‌گیری نیز با روش بدون بازگردانی انجام گیرد اما حجم جامعه حداقل به‌دو برابر حجم نمونه ($N \geqslant ۲n$) باشد .

آری اگر جامعه مورد مطالعه در یکی از سه وضعیت ذکر شده باشد بدون توجه به‌شکل توزیع (X) در جامعه هرگاه از این جامعه همه نمونه‌های n تایی ممکن اخذ شوند توزیع میانگین نمونه‌ها ($\bar{x}$) توزیعی تقریبا" نرمال خواهد بود .

مثال ۳ـ۴ : جامعه‌ای از ۴ عدد (۳،۲،۱،۰) تشکیل‌شده است . تمام نمونه‌های دوتایی‌ممکن را با روش بازگردانی از این جامعه انتخاب نموده و ثابت کنید که قضیه (۱ـ۴) و (۳ ـ ۴) در مورد آن صادق است .

پاسخ : از این جامعه ۱۶ نمونه دوتایی با روش بازگردانی می‌توان انتخاب نمود که ذیلا" این نمونه‌ها و میانگین هریک از آنها در جدول (۶ـ۴) نشان داده می‌شود .

ردیف	عناصرنمونه	$\bar{x}$
۱	۰،۰	۰
۲	۱،۰	۰/۵
۳	۲،۰	۱
۴	۳،۰	۱/۵
۵	۰،۱	۰/۵
۶	۱،۱	۱
۷	۲،۱	۱/۵
۸	۳،۱	۲

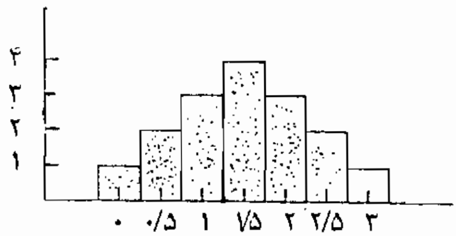
ردیف	عناصرنمونه	$\bar{x}$
۹	۰،۲	۱
۱۰	۱،۲	۱/۵
۱۱	۲،۲	۲
۱۲	۳،۲	۲/۵
۱۳	۰،۳	۱/۵
۱۴	۱،۳	۲
۱۵	۲،۳	۲/۵
۱۶	۳،۳	۳

جدول (۶-۴)

جدول توزیع فراوانی و هیستوگرام $\bar{x}$ به‌صورت زیر می‌باشد .

f	$\bar{x}$
۱	۰
۲	۰/۵
۳	۱/۰
۴	۱/۵
۳	۲/۰
۲	۲/۵
۱	۳/۰

جدول (۷-۴)



شکل (۳-۴)

همان‌طور که در منحنی (۳-۴) به‌خوبی مشخص است توزیع نمونه‌گیری ($\bar{X}$) یک توزیع نرمال می‌باشد و لذا صدق قضیه (۳-۴) تأیید می‌گردد. برای تحقیق در مورد صدق قضیه (۱-۴) باید میانگین جامعه و نیز میانگین ($\bar{X}$) را به‌دست آورد زیرا قضیه (۱-۴) می‌گوید که این دو میانگین مساوی یکدیگر هستند.

$$E(\bar{X}) = \frac{\Sigma \bar{x}_i F_i}{\Sigma F_i} = \mu$$

ابتدا میانگین جامعه را بدست می‌آوریم.

$$\mu = \frac{۰+۱+۲+۳}{۴} = ۱/۵$$

حال بر اساس اطلاعات جدول (۷-۴) میانگین ($\bar{x}$) را بدست می‌آوریم.

$$E(\bar{X}) = \frac{\Sigma \bar{x}_i F_i}{\Sigma F_i} = \frac{(۰\times۱)+(۰/۵\times۲)+\cdots+(۳\times۱)}{۱۶} = \frac{۲۴}{۱۶} = ۱/۵$$

بدین ترتیب صدق قضیه (۱-۴) نیز در مورد مثال فوق تأیید می‌شود.

دو قضیه فوق شکل توزیع نمونه‌گیری ($\bar{X}$) را در مواردی که شکل توزیع جامعه نرمال بوده و یا در صورت نرمال نبودن شکل توزیع (X) در جامعه، حجم نمونه‌ها ($n \geqslant ۳۰$) باشد مشخص نمود. اگر توزیع جامعه نرمال نباشد و نمونه مأخوذه نیز حجمی به‌اندازهٔ ($n < ۳۰$) داشته باشد نمی‌توان در مورد شکل توزیع نمونه‌گیری ($\bar{X}$) قضاوت نمود.

اینک وضعیت میانگین ($\bar{X}$) و نیز کیفیت شکل توزیع آن مشخص شده است. حال باید در مورد انحراف معیار میانگین نمونه ($\sigma_{\bar{X}}$) صحبت نمود. مطالب مربوط به انحراف معیار ($\bar{X}$) را نیز می‌توان در قالب قضایایی به‌شرح زیر خلاصه کرد.

قضیه ۴ـ۴: هرگاه

الف ـ از یک جامعه بزرگ و نامتناهی با هر روش دلخواه (بازگردانی یا بدون بازـ گردانی) با انحراف معیار (σ).

ب ـ از یک جامعه کوچک با روش بازگردانی با انحراف معیار (σ)

همه نمونه‌های n تایی ممکن اخذ شوند (مقدار n اهمیتی ندارد) ، میانگین نمونه‌ها ($\bar{X}$) دارای انحراف معیاری با مقدار تقریبی زیر خواهد بود .

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} \qquad (۴-۳)$$

در رابطه (۳ـ۴) σ انحراف معیار جامعه و n حجم نمونه را نشان می‌دهند .

مثال ۴ ـ ۴: در مثال (۳ـ۴) ثابت کنید که قضیه (۴ـ۴) صادق است .

پاسخ: واریانس ($\bar{X}$) را یک بار با استفاده مستقیم از جدول (۷ ـ ۴) و بسار دیگر با استفاده از رابطه (۳ـ۴) بدست می‌آوریم .اگر نتایج به‌طورتقریبی با یکدیگر مساوی باشند مطلوب مسئله اثبات شده است .

$$\sigma^۲_{\bar{X}} = \frac{\Sigma F_i(\bar{x}_i-\mu)^۲}{\Sigma F_i} = \frac{(۰-۱/۵)^۲+۲(۱-۱/۵)^۲+\cdots+۱(۳-۱/۵)^۲}{۱۶} = \frac{۵}{۸}$$

حال واریانس جامعه ($\sigma^۲$) را محاسبه می‌کنیم .

$$\sigma^۲_X = \frac{\Sigma(x_i-\mu)^۲}{N} = \frac{(۰-۱/۵)^۲+\cdots+(۳-۱/۵)^۲}{۴} = \frac{۵}{۴}$$

واریانس جامعه($\sigma^۲$)را براساس رابطه (۳ـ۴) بر حجم نمونه ($n=۲$) تقسیم می‌کنیم تا ($\sigma_{\bar{X}}$) از طریق این رابطه بدست آید .

$$\sigma^۲_{\bar{X}} = \frac{\sigma^۲}{n} = \frac{\frac{۵}{۴}}{۲} = \frac{۵}{۸}$$

ملاحظه می‌گردد که نتیجه از هر دو راه کاملا" مساوی یکدیگر بوده و قضیه (۴ـ۴) در این مورد صادق است .

قضیه ۵ ـ ۴: هرگاه از یک جامعه محدود با انحراف معیار (σ) همه نمونه‌های n تایی ممکن با روش بدون جایگذاری انتخاب شوند ، میانگین نمونه‌ها($\bar{X}$) انحراف معیاری معادل

مقدار زیر خواهد داشت .

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \qquad (۴-۴)$$

واضح است که در رابطه فوق (σ) انحراف معیار جامعه ، (N) حجم جامعه و (n) حجم نمونه می‌باشد .

مثال ۵ ـ ۴ : ثابت کنید قضیه (۵ ـ ۴) در مورد انحراف معیار ($\bar{X}$) در مثال (۲ـ۴)صادق است .

پاسخ : همانند مثال قبل واریانس ($\bar{X}$) را یک بار با روش مستقیم و بار دیگر با استفاده از قضیه (۵ ـ ۴) محاسبه نموده و نتایج را باهم مقایسه می‌کنیم .

$$\sigma^{۲}_{\bar{X}} = \frac{\Sigma F_i\left[\bar{x}_i - E(\bar{X})\right]^{۲}}{\Sigma F_i} = \frac{\Sigma F_i(\bar{x}_i - \mu)^{۲}}{\Sigma F_i}$$

ابتدا باید (μ) را حساب کنیم تا بتوانیم از رابطه فوق $\sigma^{۲}_{\bar{X}}$ را به‌دست آوریم .

$$\mu = \frac{\Sigma F_i x_i}{\Sigma F_i} = \frac{۱(۱) + ۲(۲) + ۳(۴) + ۴(۲) + ۵(۱)}{۱۰} = \frac{۳۰}{۱۰} = ۳$$

$$\sigma^{۲}_{\bar{X}} = \frac{۱(۲/۸-۳)^{۲} + ۲(۲/۹-۳)^{۲} + \cdots + ۱(۳/۲-۳)^{۲}}{۱۰} = \frac{۰/۱۲}{۱۰} = ۰/۰۱۲$$

حال ($\sigma^{۲}_{\bar{X}}$) را با استفاده از رابطه (۴ـ۴) محاسبه می‌کنیم .

$$\sigma^{۲}_{\bar{X}} = \frac{\sigma^{۲}}{n} \times \frac{N-n}{N-1}$$

انحراف معیار جامعه ($\sigma^{۲}$)را در اختیار نداریم و لذا باید ابتدا آن را حساب کنیم .

$$\sigma_X^{۲} = \frac{\Sigma F_i(x_i - \mu)^{۲}}{\Sigma F_i} = \frac{۱(۱-۳)^{۲} + ۲(۲-۳)^{۲} + \cdots + ۱(۵-۳)^{۲}}{۱۰} = ۱/۲$$

$$\sigma^{۲}_{\bar{X}} = \left(\frac{۱/۲}{۹}\right)\left(\frac{۱۰-۹}{۱۰-۱}\right) = \frac{۱/۲}{۹} \times \frac{۱}{۹} = ۰/۰۱۵$$

ملاحظه می‌گردد که مقدار انحراف معیار($\bar{X}$)از هردو طریق مقادیری نزدیک به‌یکدیگر داشته

و لذا می‌توان قضیه (۵ــ۴) را صادق دانست .

واضح است که در تمامی موارد که توزیع نمونه‌گیری میانگین نمونه‌ها($\bar{X}$) نرمال است می‌توان با استفاده از قاعده کلی تبدیل توزیع نرمال به‌نرمال استاندارد ، توزیع ($\bar{X}$) را نیز به‌نرمال استاندارد(Z) بدل نمود که قاعده کلی انجام این امر بر اساس رابطه (۵ــ۴) می‌باشد .

$$Z = \frac{\bar{x} - E(\bar{X})}{\sigma_{\bar{X}}} = \frac{\bar{x} - \mu}{\sigma_{\bar{X}}} \qquad (۴-۵)$$

رابطه (۵ ــ ۴) یکی از پایه‌های اصلی کار در تئوریهای تخمین ، آزمون فرضیه و غیره می‌باشد که به‌خواست خدای بزرگ در فصول آینده مورد بحث قرار می‌گیرد و طی آنها در بسیاری‌از موارد بااستفاده ازهمین رابطه و قضایای پنج‌گانه فوق‌مقدار میانگین جامعه تخمین زده می‌شود . اما فعلاً جهت ترسیم یک تصویر کلی از آن مباحث یک مثال ساده از آنها در اینجا ذکر می‌شود .

مثال ۶ ــ ۴ : از جامعه‌ای به‌حجم $N=4$ یک نمونه دوتایی اخذ شده و میانگین آن نمونه معادل ($\bar{x}=3$) محاسبه گردیده است . اگر انحراف معیار جامعه ($\sigma=4$) بوده و بدانیم‌که در صورت نرمال بودن توزیع $\bar{X}$ نقطه متناظر با نقطه ($\bar{x}=3$) در توزیع نرمال استاندارد ، نقطه ($Z=-0/25$) است ، میانگین جامعه μ را محاسبه کنید .

پاسخ : براساس قضیه (۳ــ۴) اگر تمام نمونه‌های n تایی ممکن از یک جامعه کوچک با روش بازگردانی انتخاب شوند درآن صورت میانگین نمونه‌ها دارای توزیع نرمال خواهد بود . این قضیه را به‌مثال فوق تعمیم داده و می‌گوییم اگر تمام نمونه‌های دوتایی از جامعه فوق با روش بازگردانی اخذ می‌شدند توزیع $\bar{X}$ یک توزیع نرمال بوده حال که به‌جای تمام‌نمونه‌های دوتایی فقط یک نمونه اخذ شده و میانگین آن ($\bar{x}=3$) بدست آمده است می‌توان گفت که نقطه ($\bar{x}=3$) یکی از نقاط یک توزیع نرمال است و اگر این نقطه را از میانگین μ کم نموده و حاصل‌رابرانحراف معیارتقسیم کنیم‌نقطه متناظر این نقطه در توزیع (Z) بدست‌می‌آید .

$$Z = \frac{\bar{x} - E(\bar{X})}{\sigma_{\bar{x}}}$$

از قضیه (۱ــ۴) داریم .

$$E(\bar{X}) = \mu$$

و از قضیه (۳ـ۴) داریم :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{۴}{\sqrt{۲}} = ۲/۸$$

بنابراین رابطه Z به‌صورت زیر در می‌آید .

$$Z = \frac{\bar{x} - \mu}{۲/۸}$$

اما مقدار Z متناظر با نقطه ($\bar{x}=۳$) در صورت مسئله داده شده است . با قراردادن مقدار Z و $\bar{x}$ در رابطه فوق مقدار μ محاسبه می‌شود .

$$-۰/۲۵ = \frac{۳ - \mu}{۲/۸} \longrightarrow -۰/۷ = ۳ - \mu \longrightarrow \mu = ۳/۷$$

در خاتمه این مبحث ذکر قضیه زیر که با قضایای قبلی سنخیت و تجانس نسبی داشته و در مباحث بعدی مورد استفاده قرار می‌گیرد مفید به‌نظر می‌رسد .

قضیه ۶ـ۴ : اگر تعدادی نمونه تصادفی n تایی از جامعه‌ای با توزیع دو جمله‌ای با میانگین $\mu = nP$ و $\sigma^2 = nPq$ استخراج شوند در این صورت توزیع نمونه‌گیری متغیر $\hat{p}_i = \frac{x_i}{n}$ توزیعی تقریباً نرمال با میانگین $\mu_{\hat{p}} = P$ و $\sigma_{\hat{p}} = \sqrt{\frac{Pq}{n}}$ می‌باشد و بنابراین می‌توان نوشت .

$$Z = \frac{\hat{P} - P}{\sqrt{\frac{Pq}{n}}} \qquad (۴-۶)$$

در قضیه فوق x_i تعداد موفقیتها را در هر آزمایش نشان می‌دهد .

تفاوت قضیه (۶ـ۴) با قضایای قبلی این است که در قضایای قبلی تمام نمونه‌های n تایی ممکن از جامعه باید اخذ می‌شدند در حالی که در قضیه فوق فقط تعدادی از نمونه‌ها کفایت می‌کند . البته هرچه تعداد نمونه‌ها بیشتر باشد توزیع $\hat{p}$ به‌توزیع نرمال نزدیکتر می‌شود .

مثال ۷ـ۴ : شیئی ۴ وجهی در اختیار داریم که یکی از وجوه آن قرمز رنگ بوده و مشاهده همین وجه قرمز رنگ پس از پرتاب این ۴ وجهی به‌هوا برای ما به‌منزله موفقیت تلقی می‌شود. اگر در یک نمونه ،آزمایش پرتاب را ($n=۳۰$) بار انجام‌دهیم در صورتی که احتمال موفقیت

($P = ۰/۶$) و نقطه متناظر با P مربوط به این نمونه ($Z = ۲$) باشد، تعداد موفقیتها را در این نمونه حساب کنید.

پاسخ: اگر تعدادی نمونه ۳۰ تایی اخذ می‌شد متغیر $\hat{P}_i = \frac{x_i}{n}$ توزیع نرمال داشت. حال که فقط یک نمونه از این جامعه اخذ شده است $\hat{P}$ مربوط به این نمونه یکی از نقاط یک توزیع نرمال بوده و بنابراین با استفاده از قضیه (۶-۴) می‌توان آن را تبدیل به استاندارد نموده و مطلوب مسئله را محاسبه نمود.

$$Z = \frac{\hat{P} - P}{\sqrt{\frac{Pq}{n}}} \longrightarrow ۲ = \frac{\hat{P} - ۰/۶}{\sqrt{\frac{(۰/۶)(۰/۴)}{۳}}}$$

$$\hat{P} = ۰/۷۸$$

$$\hat{P} = \frac{x}{n} \rightarrow x = n\hat{P} = ۳۰(۰/۷۸) \simeq ۲۳$$

الف ـ ۲ـ رابطه بین ($\mu_۱ - \mu_۲$) با متغیر ($\bar{X}_۱ - \bar{X}_۲$) با استفاده از توزیع Z

همان طور که در مقدمه کتاب گفته شد در بسیاری از موارد هدف آمارگر مقایسه بین میانگین دو جامعه می‌باشد یعنی می‌خواهد میانگین جامعه اول ($\mu_۱$) را با میانگین جامعه دوم ($\mu_۲$) مقایسه کند بدون آنکه به مقادیر تک تک این دو پارامتر توجهی داشته باشد. در چنین مواردی معمولا" چنین عمل می‌کنند که هدف فوق را از طریق بررسی مقدار یک پارامتر جدید یعنی ($\mu_۱ - \mu_۲$) محقق می‌سازند. شیوه عمل براساس این استدلال شکل می‌گیرد که اگر میانگین نمونه‌های مربوط به یک جامعه ($\bar{X}_۱$) به تنهایی توزیع نمونه‌گیری نرمال داشته و به همین ترتیب ($\bar{X}_۲$) نیز توزیع نمونه‌گیری نرمال دارد بنابراین تفاضل این دو یعنی ($\bar{X}_۱ - \bar{X}_۲$) نیز متغیر تصادفی جدیدی خواهد بود که توزیع نمونه‌گیری نرمال داشته و لذا کافی است میانگین و انحراف معیار این متغیر جدید را پیدا کرده و آن را به نرمال استاندارد تبدیل نمود. از آنجا که به سادگی ثابت می‌شود که $E(\bar{X}_۱ - \bar{X}_۲) = (\mu_۱ - \mu_۲)$ می‌باشد رابطه جدیدی که برای Z بدست می‌آید کار تأمین هدف فوق یعنی محاسبه تقریبی ($\mu_۱ - \mu_۲$) را عملی می‌سازد. برای بدست آوردن رابطه Z براساس متغیر ($\bar{X}_۱ - \bar{X}_۲$) کافی است میانگین و انحراف معیار آن را محاسبه نمود.

براساس قواعد امید ریاضی محاسبه میانگین متغیر ($\bar{X}_۱ - \bar{X}_۲$) یعنی اثبات رابطه‌ای که چند سطر پیش گفته شد بسیار ساده است و به صورت زیر انجام می‌گیرد.

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

به‌همین ترتیب می‌توان انحراف معیار این متغیررا نیز به‌راحتی محاسبه نمود .

$$\sigma^2_{(\bar{x}_1 - \bar{x}_2)} = \sigma^2_{\bar{x}_1} + \sigma^2_{(-\bar{x}_2)} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

مجموع مطالب فوق را می‌توان در قالب قضیه‌ای به‌صورت زیر بیان نمود .

قضیه ۶ ــ ۴ : هرگاه توزیع نمونه‌گیری میانگین ($\bar{X}_1$) یک جامعه با میانگین (μ_1)و انحراف معیار (σ_1) تقریبا" نرمال بوده و به‌همین ترتیب توزیع نمونه‌گیری میانگین ($\bar{X}_2$) جامعه دیگری به‌میانگین (μ_2)و انحراف معیار (σ_2) تقریبا" نرمال باشد در آن صورت تفاضل میانگین نمونه‌های این دو جامعه یعنی($\bar{X}_1 - \bar{X}_2$)نیز متغیری تصادفی با توزیع نمونه‌گیری تقریبا" نرمال خواهد بود . میانگین این متغیر جدید ($\mu_1 - \mu_2$)وانحراف‌معیار آن $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ می‌باشد .

واضح است که وقتی متغیری مثل ($\bar{X}_1 - \bar{X}_2$) دارای توزیع نمونه‌گیری نرمال بوده و میانگین و انحراف معیار آن نیز در اختیار باشد می‌توان آن را به‌توزیع نرمال استاندارد (Z) تبدیل نموده و از جداول توزیع (Z) در مورد مسائل این متغیر استفاده کرد . این تبدیل بر اساس رابطه زیر انجام می‌گیرد .

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad (۴-۷)$$

مثال ۸ ــ ۴ : یک جامعه ۳ عضوی با اعضای (۳و۴و ۵) و یک جامعه دو عضوی با اعضای (٥ و ۳) دراختیار است .از جامعه‌اول همه نمونه‌های دوتایی ممکن و ازجامعه دوم همه‌نمونه‌های

سه‌تایی ممکن را با جایگذاری انتخاب می‌کنیم ، مطلوبست تحقیق موارد زیر:

الف ـ توزیع نمونه‌گیری $(\bar{X}_1 - \bar{X}_2)$ نرمال است .

ب ـ $\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_1 - \mu_2$

ج ـ $\sigma^2_{(\bar{x}_1 - \bar{x}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

پاسخ: چون نمونه‌گیری‌ها با جایگذاری انجام شده‌اند توزیع نمونه‌گیری $\bar{x}$ ها تقریبا" نرمال خواهد بود . تعداد نمونه‌های ممکن دو جامعه به‌همراه میانگین‌های مربوطه در جدول (۸ـ۴) منعکس شده است و صحت ادعای فوق توسط این جدول تأیید می‌شود .

جامعه۱			جامعه ۲		
شماره	نمونه	$\bar{x}_i$	شماره	نمونه	$\bar{x}_i$
۱	۳،۳	۳/۰	۱	۰،۰،۰	۰
۲	۳،۴	۳/۵	۲	۰،۰،۳	۱
۳	۳،۵	۴/۰	۳	۰،۳،۰	۱
۴	۴،۳	۳/۵	۴	۳،۰،۰	۱
۵	۴،۴	۴/۰	۵	۰،۳،۳	۲
۶	۴،۵	۴/۵	۶	۳،۰،۳	۲
۷	۵،۳	۴/۰	۷	۳،۳،۰	۲
۸	۵،۴	۴/۵	۸	۳،۳،۳	۳
۹	۵،۵	۵/۰			

جـدول (۸-۴)

الف ـ براساس جدول (۸ـ۴) متغیر $(\bar{X}_1 - \bar{X}_2)$ ۷۲ مقدار به‌شرح جدول (۹ـ۴) خواهد داشت.

$\bar{x}_{2i}$	$\bar{x}_{1i}$								
	۳/۰	۳/۵	۴/۰	۳/۵	۴/۰	۴/۵	۴/۰	۴/۵	۵/۰
۰	۳/۰	۳/۵	۴/۰	۳/۵	۴/۰	۴/۵	۴/۰	۴/۵	۵/۰
۱	۲/۰	۲/۵	۳/۰	۲/۵	۳/۰	۳/۵	۳/۰	۳/۵	۴/۰
۱	۲/۰	۲/۵	۳/۰	۲/۵	۳/۰	۳/۵	۳/۰	۳/۵	۴/۰
۱	۲/۰	۲/۵	۲/۰	۲/۵	۳/۰	۳/۵	۳/۰	۳/۵	۴/۰
۲	۱/۰	۱/۵	۲/۰	۱/۵	۲/۰	۲/۵	۲/۰	۲/۵	۳/۰
۲	۱/۰	۱/۵	۲/۰	۱/۵	۲/۰	۲/۵	۲/۰	۲/۵	۳/۰
۲	۱/۰	۱/۵	۲/۰	۱/۵	۲/۰	۲/۵	۲/۰	۲/۵	۳/۰
۳	۰	۰/۵	۱/۰	۰/۵	۱/۰	۱/۵	۱/۰	۱/۵	۲/۰

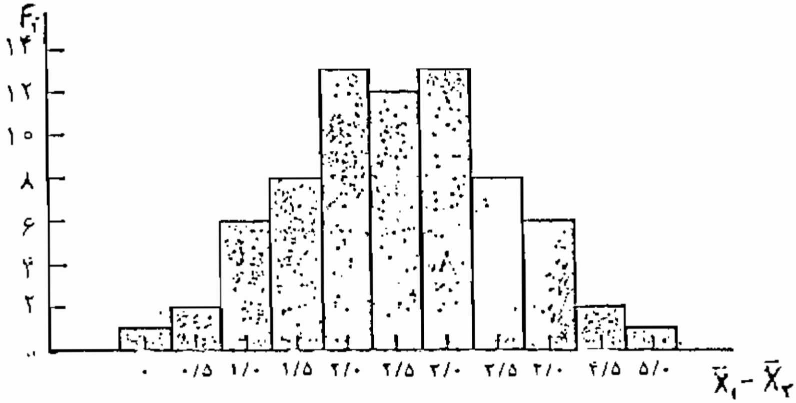
جـــــدول (۹-۴)

جدول توزیع فراوانی متغیر ($\bar{X}_1 - \bar{X}_2$) براساس جدول (۹-۴) به‌صورت جـــدول (۱۰-۴) خواهد بود.

$(\bar{x}_1 - \bar{x}_2)_i$	F_i
۰	۱
۰/۵	۲
۱/۰	۶
۱/۵	۸
۲/۰	۱۳
۲/۵	۱۲
۳/۰	۱۳
۳/۵	۸
۴/۰	۶
۴/۵	۲
۵/۰	۱

جدول (۱۰-۴)

جدول (۱۰–۴) به‌خوبی توزیع نمونه‌گیری تقریباً" نرمال متغیر $(\bar{X}_1 - \bar{X}_2)$ را نشان می‌دهد با این وجود برای نمایش بهتر این مطلب هیستوگرام این توزیع درشکل (۴–۴) رسم شده است .



شکل (۴ – ۴)

ب – برای اثبات بند ب یکبار $(\mu_1 - \mu_2)$ و بار دیگر $\mu_{(\bar{X}_1 - \bar{X}_2)}$ را محاسبه می‌کنیم. داریم .

$$\mu_1 = \frac{۳ + ۴ + ۵}{۳} = ۴$$

$$\mu_2 = \frac{۰ + ۳}{۲} = ۱/۵$$

$$\mu_1 - \mu_2 = ۴ - ۱/۵ = ۲/۵$$

براساس جدول (۱۰–۴) داریم

$$\mu_{(\bar{x}_1 - \bar{x}_2)} = \frac{\Sigma F_i (\bar{x}_1 - \bar{x}_2)_i}{\Sigma F_i} = \frac{(۰)(۱) + (۰/۵)(۲) + \cdots + (۵)(۱)}{۷۲} = \frac{۱۸۰}{۷۲} = ۲/۵$$

و بدین ترتیب مطلوب بند ب ثابت می‌شود .

ج – برای اثبات این بند نیز مانند بند ب دو طرف تساوی را جداگانه حساب می‌کنیم .

$$\sigma^{۲}_{\bar{x}_۱-\bar{x}_۲}=\frac{\Sigma F_i[(\bar{x}_۱-\bar{x}_۲)-(\mu_۱-\mu_۲)]^۲}{\Sigma F_i}=\frac{(۰-۰/۲۵)^۲+۲(۰/۵-۲/۵)^۲+\cdots+(۵-۲/۵)^۲}{۷۲}$$

$$\sigma^{۲}_{\bar{x}_۱-\bar{x}_۲}=\frac{۷۸}{۷۲}=۱/۰۸$$

برای محاسبه طرف دیگر تساوی باید قبلا" واریانس دو جامعه را محاسبه کنیم .

$$\sigma^۲_۱=\frac{(x_{۱i}-\mu_۱)^۲}{N_۱}=\frac{(۳-۴)^۲+(۴-۴)^۲+(۵-۴)^۲}{۳}=\frac{۲}{۳}$$

$$\sigma^۲_۲=\frac{(۰-۱/۵)^۲+(۳-۱/۵)^۲}{۲}=\frac{۹}{۴}$$

حال طرف دوم تساوی را محاسبه می‌کنیم .

$$\frac{\sigma^۲_۱}{n_۱}+\frac{\sigma^۲_۲}{n_۱}=\frac{\frac{۲}{۳}}{۲}+\frac{\frac{۹}{۴}}{۳}=\frac{۱۳}{۱۲}=۱/۰۸$$

بدین ترتیب قضیه (۶ـ۴) بطور کامل توسط مثال (۸ـ۴) تأیید می‌گردد . واضح است که در قضیه (۶ـ۴) شکل توزیع متغیرها در دو جامعه اهمیتی نداشته و هرگاه توزیع نمونه‌گیری ($\bar{X}$)ها نرمال و یا نزدیک به‌آن باشد توزیع نمونه‌گیری ($\bar{X}_۱-\bar{X}_۲$) نرمال خواهدبود خواه توزیع جوامع اصلی نرمال ، دوجمله‌ای و یا هر توزیع دیگری باشد . به‌همین ترتیب قضیه (۶ ـ ۴)درمورد متغیر ($\hat{P}_۱-\hat{P}_۲$) در صورتی که هریک از دو متغیر $\hat{P}_۱$ و $\hat{P}_۲$ توزیع نرمال داشته باشند نیز عینا" صادق است .

الف ـ ۳ ـ توزیع t و ربط بین میانگین نمونه و جامعه از طریق آن

تا کنون برای حل مسائل مربوط به متغیرهایی از قبیل X و $\bar{X}$ که دارای توزیع نرمال بودند ابتدا این متغیرها را به‌متغیر Z تبدیل می‌کرده و با استفاده از جداول مربوط به‌توزیع

نرمال استاندارد Z ، کار حل مسئله را به‌مقدار زیادی تسهیل می‌نمودیم . شیوه تبدیل هر متغیر تصادفی به‌متغیر نرمال استاندارد Z نیز بدین صورت بود که مقدار آن متغیررا از میانگین آن کم نموده و تفاضل را بر انحراف معیار متغیر تقسیم می‌نمودیم . واضح است که وقتی می‌توان‌این عمل تبدیل را انجام داد که مقدار انحراف معیار متغیرموردنظر رابدانیم. اگر متغیر اصلی ما X باشد تبدیل آن به Z منوط به‌دانستن مقدار (σ_X) جامعه است‌واگراین متغیر $\bar{X}$ باشد بازهم محتاج دانستن σ جامعه می‌باشیم ، زیرا همان طور که درقضایای قبلی ملاحظه گردید برای‌بدست آوردن $\sigma_{\bar{X}}$ از انحراف‌معیار جامعه (σ)استفاده می‌شود . به‌دلیل همین دخیل بودن (σ_X) در عملیات تبدیل متغیر نرمال به‌متغیر (Z) ، در اکثر عملیات آماری استفاده از متغیر (Z) به‌روش فوق ممکن نیست زیرا معمولاً" مقدار (σ) نه تنها بر محقق مجهول است بلکه در بیشتر اوقات پیدا کردن مقداری برای (σ) یکی از اهداف عملیات آماری می‌باشد .

چاره چیست ؟ آیا باید در برخورد بااین مشکل به‌کلی ازعملیات آماری‌صرفنظرنمود ؟ یا آن که بدون استفاده از متغیری مانند (Z) از روشهای پرپیچ و خم و طولانی دیگری‌برای تحقق هدف آماری استفاده نمود ؟ که البته در بسیاری موارد چنین کاری هم ممکن نیست . آمارگران هیچیک از دو طریق نامعقول فوق را اختیار نکرده‌اند . روش متخذه در برخورد با این مشکل از سوی‌آنان این بوده است که برای (σ) یک مشخصه جایگزین انتخاب نموده و با استفاده از این مشخصه راه حل مشکل را بدست آورده‌اند . بدون شک اکنون سئوال‌این است که این مشخصه چیست ؟ در پاسخ گفته می‌شود که مناسب‌ترین مشخصه‌ای که می‌تواند به‌جای انحراف معیار جامعه (σ) مورد استفاده قرار گیرد ، انحراف معیار نمونه می‌باشد . بدین ترتیب آنان در مواردی که مقدار (σ) مجهول است در هنگام استانداردکردن متغیر مورد مطالعه از انحراف معیار نمونه (S) استفاده می‌نمایند . بدون تردید اینک خواهید پرسید که محصول و فرایند چنین تبدیلی بازهم متغیر نرمال استاندارد (Z) خواهد بود ؟ پاسخ منفی است و متغیر حاصله متغیر نرمال دیگری به‌نام متغیر(t)می‌باشد . رابطه این تبدیل همان طور که گفته شد همان رابطه مربوط به (Z) است . با این تفاوت که در مخرج کسر به‌جای (σ) انحراف معیار نمونه یعنی (S) به‌کار رفته است .

$$t = \frac{x-\mu}{S} \qquad (۴-۸)$$

مثال ۹ ــ ۴ : متغیر (X) در جامعه‌ای دارای توزیع نرمال با میانگین ($\mu =$ ۱۰) و انحراف معیار مجهول (σ) می‌باشد . در صورتی که ($x_i =$ ۴) یکی از مقادیر این متغیر بوده

و انحراف معیار یک نمونه ۵ تائی از این جامعه ($S=۲$) باشد ، مقدار (t) متناظر با نقطه (x_1) را محاسبه کنید .

پاسخ : از رابطه (۸ــ۴) استفاده می‌کنیم .

$$t = \frac{x - \mu}{S_x} = \frac{۴-۱۰}{۲} = \frac{-۶}{۲} = -۳$$

متغیر t برای اولین بار در سال ۱۹۰۸ در یک مقاله به‌قلم فردی به‌نام و. س. گوست[1] مورد بررسی قرار گرفت . از آنجا که در آن زمان این شخص در استخدام یک شرکت ایرلندی بود و این شرکت اجازه انتشار مقالات را با نام خود گوست به‌وی نمی‌داد ، او مقالات خود را با نام مستعار Student یا دانشجو منتشر می‌نمود و به‌همین دلیل این متغیر تدریجا" به t استیودنت یا t دانشجو معروف گردید . توزیع احتمال این متغیر بعدها توسط شخص دیگری به‌نام ر. ا. فیشر[2] مورد بررسی قرار گرفته و مشخص گردید .

همان طور که گفته شد متغیر (t) نیز مانند (Z) دارای توزیع احتمال نرمال می‌باشد تنها تفاوت اصلی بین این دو متغیر این است که در رابطه محاسباتی مربوط به (t) از انحراف معیار نمونه (S) استفاده شده است در حالی که در رابطه (Z) انحراف معیار جامعه (σ) مورد استفاده قرار می‌گیرد . همان طور که می‌دانیم انحراف معیار یک جامعه (σ) مقدار ثابتی داشته و می‌توان آن‌را به‌عنوان یک پارامتر تلقی نمود در حالی‌که انحراف معیار نمونه (S) مشخصه‌ای متغیر است که مقدار آن به‌چگونگی نمونه‌گیری و بخصوص حجم نمونه ($n-۱$) کاملا" وابسته است . طبیعی ترین اثر این مسئله این است که مقدار متغیر نرمال استاندارد (Z) فقط به‌یک متغیر یعنی (X)[3] بستگی داشته باشد در حالی که مقدار متغیر نرمال (t) به‌دو متغیر (X) و S_x وابسته است[4] . از آنجا که مقدار S_x نیز چنان که می‌دانیم به‌حجم نمونه ($n-۱$) بستگی دارد می‌توان گفت مقدار متغیر (t) تابع مقادیر دو متغیر

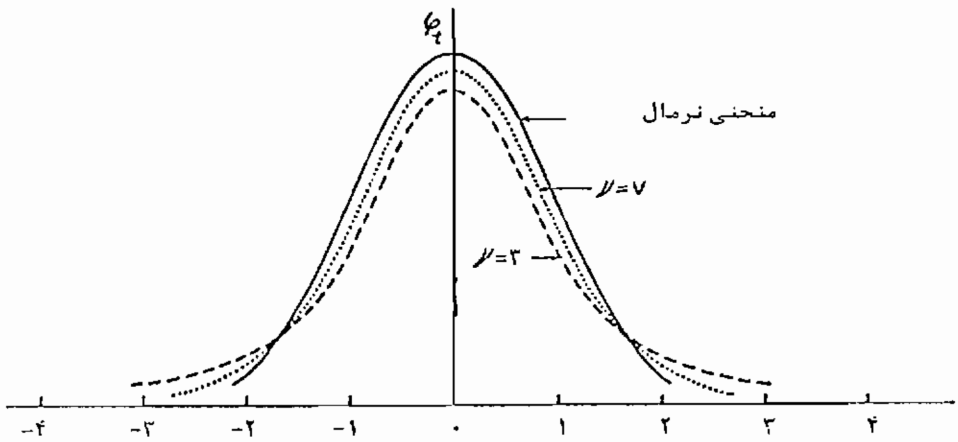
---

1 - W.S.Gosset 2 - R.A.Fisher

۳ ــ در صورتی که برای محاسبه (Z) متغیرهائی از قبیل $\bar{X}$ یا $(\bar{X}_1 - \bar{X}_2)$ و غیره به‌کار رفته باشد مقدار Z فقط به‌مقدار این متغیرها وابسته است و $\sigma_{\bar{X}}$ و یا $\sigma_{(\bar{X}_1-\bar{X}_2)}$ نیز همانند σ_x مقدارشان ثابت است .

۴ ــ در توزیع t نیز می‌توان به‌جای متغیر (X) از متغیرهای دیگری نظیر ($\bar{X}$) یا $(\bar{X}_1 - \bar{X}_2)$ استفاده نمود . در این صورت انحراف معیار این متغیرها نیز خود یک متغیر است و باز هم مقدار (t) به‌مقادیر دو متغیر بستگی پیدا می‌کند .

(X) و (۱ـ n) می‌باشد . به متغیر (۱ـ n) اصطلاحا" درجه آزادی گفته می‌شود و با حرف یونانی (ν)نشان داده می‌شود .وابستگی مقدارمتغیر (t) به‌دو متغیر (X)و (۱ـ n)طی جریاناتی که از طریق تابع چگالی احتمال توزیع نرمال (t) استیودنت عمل می‌کند سبب می‌شود که برخلاف منحنی نرمال استاندارد (Z) ، متغیر (t) برحسب مقادیر مختلف (X) و (S) منحنی‌های توزیع احتمال متعددی داشته باشد که البته همه این منحنی‌هااز منحنی (Z) پهن‌تر بوده و همان‌طور که در شکل (۵ ــ ۴) ملاحظه می‌شود هرچقدر حجم نمونه (n) و به‌تبع آن درجه آزادی ($\nu = n-1$) بزرگتر می‌شود منحنی توزیع احتمال (t) به‌منحنی (Z) نزدیکتر می‌گردد به‌طوری که در حجم های نمونه بالا ($n \geqslant 30$) این دو منحنی به‌قدری به‌یکدیگر شبیه هستند که آمارگران در عمل ترجیح می‌دهند علیرغم استفاده از انحراف معیار نمونه (S) ، از جداول مربوط به (Z) استفاده کنند . معنای این حرف این است‌که وقتی ($n \geqslant 30$)باشد انحراف معیارنمونه (S) تخمین خوبی از انحراف معیار جامعه (σ) بوده و می‌توان آن را به‌جای (σ) مورد استفاده قرار داد . بدین‌ترتیب استفاده ازمتغیر (t) به‌مواردی محدودمی‌شود که اولا" مقدار انحراف معیار جامعه مجهول و نامشخص باشد و ثانیا" ($n < 30$) بوده و به‌عبارت بهتر حجم نمونه نیز کوچک باشد ، تا دراثر آن ، فرض ($\sigma \simeq S$) غیر ممکن گردد نتیجتا" مقدار (σ) همچنان مجهول باقی‌بماند چه در غیر این صورت با استفاده از σ توزیع نرمال را استاندارد نموده و از جداول(Z) که کار با آنها راحت تر است استفاده می‌کنند .



شکل (۵ ــ ۴)

خوشبختانه مقادیراحتمال متغیر (t)برحسب مقادیرمختلف این متغیر و نیز برحسب مقادیر مختلف درجه آزادی (ν) قبلاً" محاسبه شده و به‌صورتی متناسب با نیازهای آماری طی جداولی که بعضی از آنها به‌همین کتاب نیز ضمیمه شده‌اند در اختیار آمارگران قرار دارند و درواقع به‌خاطر رسیدن به‌چنین جداولی بودکه آمارگران وقت خودرا صرف بررسی متغیر t نمودند تا در مواردی که محاسبات مربوط به‌متغیرهای نرمال از طریق توزیع (Z) ممکن نیست بتوانند بااستفاده ازاین جداول خودرا از دام محاسبات پیچیده توابع چگالی و غیره رهایی بخشند . معمولاً" در منحنی توزیع (t) رسم براین است که سطح زیر منحنی در طرف راست هریک از مقادیر (t) را با حرف یونانی (α) نشان داده و آن مقدار از (t) را که مساحتی معادل (α)را در سمت راست خود جدامی‌کند با t_α نمایش‌می‌دهند . درجدول ضمیمه کتاب مقادیر (t_α) برای ۵ مقدار مختلف (α) که بیشتر موردنیاز استبر حسب درجات آزادی مورد لزوم $(29 > n-1 = \nu)$ داده شده‌اند . به‌عنوان مثال مقدار متغیر t با درجه آزادی ($\nu = 6$) در سطح ($\alpha = 0/05$) برابر ($t = 1/943$) می‌باشد . چنانکه قبلاً" گفته شد برای درجات آزادی بالاتر می‌توان از توزیع نرمال استاندارد (Z) به‌جای (t) استفاده نمود .

با وجود جداول فوق معمولاً" محققین در عمل نیازی به‌تابع چگالی احتمال (φ_t) و نیز میانگین (μ_t) و انحراف معیار (σ_t) پیدا نمی‌کنند اما جهت مزید اطلاع خواننده گرامی خاطرنشان می‌گردد که میانگین (t) همانند میانگین (Z) مساوی صفر بوده ($\mu_t = \mu_z = 0$) و انحراف‌معیارآن در درجات آزادی ($\nu < 29$) همواره از ۱ بزرگتر است ($\sigma_t > 1$) زیرا منحنی‌های (t) پهن‌تر از منحنی (Z) بوده‌وواضح است که هرچه‌یک منحنی با شکل‌کلی ثابت مثل منحنی نرمال پهن تر می‌شودانحراف معیارآن بزرگترمی‌گردد و چون ($\sigma_z = 1$)می‌باشد قطعاً" همه‌منحنی‌های (t)مزبور انحراف معیاری بزرگتر از ۱ ($\sigma_t \gtrsim 1$) خواهند داشت .

مثال ۱۰ ــ ۴ : متغیر (X) درجامعه‌ای دارای توزیع‌نرمال با میانگین ($\mu = 20$) و انحراف معیار مجهول (σ)می‌باشد . اگر بدانیم‌نقاط متناظر با یکی‌ازمقادیراین متغیر در توزیع (z ، t) به‌ترتیب نقطه ($Z = 2$ و $t = 3$) می‌باشد مطلوب است :

الف ــ مقدار (X) در نقطه مذکور در صورتی که بدانیم انحراف معیار یک نمونه ۳۵ تایی‌از این جامعه $S = 3$ می‌باشد .

ب ــ انحراف‌معیار نمونه‌ای که براساس‌آن مقدار (t)متناظر با نقطه فوق‌محاسبه‌شده‌است .

ج ــ اگر بدانیم متغیر (t) مزبور مربوط به نقطه ($\alpha = 0/01$) می‌باشد مقدار درجه آزادی (ν) و حجم نمونه (n) را محاسبه کنید .

پاسخ: الف ـ چون $(n = ۳۵)$ است داریم :

$$Z = \frac{x - \mu}{S_X \simeq \sigma_X} \longrightarrow x - \mu = (Z)(\sigma) \longrightarrow x = Z\sigma + \mu = (۲\times۳) + ۲۰ = ۲۶$$

ب ـ براساس رابطه (۷ـ۴) داریم :

$$t = \frac{x - \mu}{S} \longrightarrow S = \frac{x - \mu}{t}$$

در بند الف مقدار $(x = ۲۶)$ محاسبه گردید که آنرا در رابطه فوق قرار داده و مقدار S را محاسبه می‌کنیم .

$$S = \frac{۲۶ - ۲۰}{۳} = ۲$$

ج ـ با مراجعه به‌جدول در می‌یابیم که در نقطه $(\alpha = ۰/۰۱)$ مقدار t با درجه آزادی $(\nu = ۷)$ مقدار تقریبی (۳) دارد .

$$t_{(۰/۰۱ \text{ و } ۷)} \simeq ۳ \longrightarrow \nu = ۷$$

داریم :

$$\nu = n - ۱ \rightarrow n = \nu + ۱ = ۷ + ۱ = ۸$$

همان‌طورکه قبلاً گفته شداگر مقدار یک متغیررا از میانگین آن‌کم نموده و بر انحراف معیار نمونه مربوط به‌همان متغیر تقسیم کنیم متغیر حاصله (t) استیودنت خواهد بود . ممکن است این متغیر (X) یا $(\bar{X})$ یا $(\bar{X}_۱ - \bar{X}_۲)$ و یا هر متغیر با هرشکل دیگری باشد . واضح است‌که وقتی متغیر $(\bar{X})$ یا $(\bar{X}_۱ - \bar{X}_۲)$متغیر اصلی ما باشدباید انحراف‌معیار همین متغیر را در مخرج‌کسر رابطه (۸ـ۴) مورد استفاده قرار داد که در این صورت رابطه به‌صور زیر تبدیل می‌گردد .

$$t = \frac{\bar{x} - \mu}{S_{\bar{X}}}$$

یا

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_{(\bar{x}_1-\bar{x}_2)}}$$

سئوالی که اکنون مطرح می‌شود این است که مقادیر $S_{\bar{X}}$ و $S_{(\bar{X}_1-\bar{X}_2)}$ چگونه‌تعیین می‌گردد ؟ در مورد ($S_{\bar{X}}$) مشکل چندانی وجود ندارد زیرا طرز محاسبهٔ آن درست مشابه $\sigma_{\bar{X}}$ می‌باشد یعنی داریم :

$$S_{\bar{X}} = \frac{S_X}{\sqrt{n}} \qquad (۴-۹)$$

با قرار دادن این مقدار در رابطه فوق ، این رابطه برحسب متغیر ($\bar{X}$) به‌شکل‌زیر کامل می‌شود .

$$t = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} \qquad (۴-۱۰)$$

اما اگر متغیر اصلی ما ($\bar{X}_1 - \bar{X}_2$) باشد محاسبه انحراف معیار آن کمی متفاوت باروشهای قبلی خواهد بود . در بررسی متغیر نرمال استاندارد (Z) برحسب ($\bar{X}_1 - \bar{X}_2$) ملاحظه کردیم که انحراف معیار این متغیر برابر با $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ می‌باشد . حال اگر مقادیر انحراف معیار از دو نمونه مربوط به دو جامعه با حجم‌های ($n_1 \geq ۳۰$ و $n_2 \geq ۳۰$) بدست آمده باشند ، چنان که قبلا" گفته شد مقادیر S_1^2 و S_2^2 حاصله تخمین‌های نسبتا" خوبی از σ_1^2 و σ_2^2 بوده و بازهم می‌توان با استفاده از این دو انحراف معیار نمونه در رابطه فوق، انحراف معیار جامعه مربوط به‌متغیر ($\bar{X}_1 - \bar{X}_2$) را بطور تقریبی بدست آورد . مشکل اصلی وقتی بروزمی‌کند که حجم یکی از دو نمونه یا هردوی آنها از ۳۰ کوچکتر باشد دراین‌صورت جمله $\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$ تخمین خوبی از انحراف معیار جامعه $\sigma_{(\bar{x}_1-\bar{x}_2)}$ نخواهد بودو اگر از این عبارت به‌جای انحراف معیار جامعه استفاده شود متغیر حاصله نه تنها توزیع نرمال

استاندارد (Z) نخواهد داشت بلکه توزیع احتمال آن با توزیع (t) نیز بسیار متفاوت می‌باشد. نتیجه تلاشهایی‌که برای حل معضل‌فوق صورت گرفته‌است بدین صورت قابل بیان است که اگر متغیر اصلی ما متغیر $(\bar{X}_1 - \bar{X}_2)$ با توزیع نرمال باشد و انحراف معیار جامعه نیز مجهول بوده و حجم حداقل یکی از دو نمونه نیز کمتر از ۳۰ باشد فقط‌وقتی می‌توان این متغیر را به‌متغیر (t) تبدیل نمود که بدانیم انحراف معیار دو جامعه با یکدیگر برابر است یعنی($\sigma_1^2 = \sigma_2^2$)می‌باشد. معمولاً" در آمار چنین متداول است که اگر در جایی از وجود این شرط مطمئن نباشند ابتدا به‌نحوی وجود آنرا ثابت می‌کنند و بعد متغیر فوق را به‌متغیر (t) تبدیل می‌کنند. به‌هرحال اگر به‌نحوی از وجود شرط فوق‌اطمینان‌حاصل شود در آن صورت می‌توان متغیر نرمال $(\bar{X}_1 - \bar{X}_2)$ را براساس رابطه (۱۱ــ۴) به‌متغیر نرمال t تبدیل کرد.

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \qquad (۴-۱۱)$$

متغیر (t) حاصل از رابطه (۱۱ــ۴) دارای توزیع احتمال (t) با درجه آزادی ($\nu = n_1 + n_2 - 2$)می‌باشد. مشخصه(S_p)که در مخرج کسر فوق به‌کار برده شده است از طریق رابطه (۱۲ــ۴) قابل محاسبه است.

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

مثال ۱۱ ــ ۴: متغیر (X) در جامعه‌ای دارای توزیع نرمال با میانگین ($\mu = 30$)و انحراف معیار مجهول (σ) می‌باشد.

الف ــ از این جامعه نمونه‌ای به‌حجم ($n=16$) اخذ شده و میانگین و انحراف معیار آن به‌ترتیب معادل ($\bar{x}=32$ و $S=3$) محاسبه شده‌است. مقدار(t) متناظر با میانگین این نمونه را محاسبه کنید.

ب ــ از جامعه دیگری با متغیر نرمال (X) به میانگین ($\mu_2=32$) و انحراف معیار مجهول (σ_2) یک نمونه ۹ تایی اخذ شده و مشخصه‌های آن به‌ترتیب ($\bar{x}_2=30$ و $S_2=2$)محاسبه شده‌است مقدار t متناظر با مقدار ($\bar{x}_1 - \bar{x}_2$)را محاسبه کنید، در صورتی که‌واریانسهای

دو جامعه مساوی باشند .

پاسخ : الف ــ از رابطه (۹ــ۴) استفاده می‌کنیم .

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} = \frac{۳۲-۳۰}{\frac{۲}{\sqrt{۱۶}}} = ۴$$

ب ــ ابتدا از رابطه (۱۱ــ۴) مقدار S_p را محاسبه نموده و سپس از طریق رابطه (۱۰ ــ ۴) مقدار t را به‌دست می‌آوریم .

$$S_p^۲ = \frac{(n_۱-۱)S_۱^۲ + (n_۲-۱)S_۲^۲}{n_۱+n_۲-۲} = \frac{(۱۵\times۹) + (۸\times۴)}{۱۶+۹-۲} = \frac{۱۶۷}{۲۳} = ۷/۲$$

$S_p = ۲/۷$

$$t = \frac{(\bar{x}_۱-\bar{x}_۲)-(\mu_۱-\mu_۲)}{S_p\sqrt{\frac{1}{n_۱}+\frac{1}{n_۲}}} = \frac{(۳۲-۳۰) - (۳۰-۳۲)}{۲/۷ \sqrt{\frac{۱}{۱۶}+\frac{۱}{۹}}} = \frac{۳ + ۲}{(۲/۷)(\frac{۱}{۵})} = ۴/۲۴$$

مثال ۱۲ ــ ۴ : یک کارخانه کمپوت وزن خالص قوطی‌های کمپوت پر شده خود را $۴۰۰\ gr$ اعلام نموده‌است . برای آزمون این مدعا یک نمونه ۲۱ تایی از این محصولات موردآزمایش قرار گرفته و مشخصه‌های این نمونه به‌ترتیب ($\bar{x} = ۴۰۵\ gr$ و $S = ۱۰$) محاسبه شده‌اند .

الف ــ اگر ملاک صحت ادعا این باشد که t محاسبه شده در فاصله ($t_{۰/۰۱}$ و $-t_{۰/۰۱}$) قرار گیرد آیا می‌توان ادعای فوق را رد کرد ؟

ب ــ اگر یک‌نمونه ۴ تایی‌ازمحصولات یک‌کارخانه‌دیگربه‌ترتیب‌دارای‌مشخصه‌های ($\bar{x}_۲ = ۴۱۰\ gr$ و $S_۲ = ۴$) بوده و صاحب‌کارخانه اول مدعی‌باشد که‌وزن قوطی‌های کارخانه‌دوم نیز همان $۴۰۰\ gr$ است آیا می‌توان این ادعا را باطل دانست در صورتی که ملاک قضاوت همان‌ملاک بند الف بوده و واریانس محصولات دو کارخانه مساوی هم باشند .

پاسخ : الف ــ برای آزمون صحت ادعای فوق کافی است مقدار ($t_{۰/۰۱}$ و $-t_{۰/۰۱}$) را با درجه آزادی ($\nu = ۲۱-۱ = ۲۰$) از جدول بدست‌آوریم و سپس مقدار t مربوط به بند الف را با استفاده از اطلاعات صورت مسئله حساب کنیم . اگر مقدار t محاسبه شـده در داخـل فاصله بدست آمده از جدول باشد ادعا قابل رد نیست و در غیر این صورت رد می‌شود .

$$t_{(۰/۰۱\ ,\ ۲۰)} = ۲/۵۲۸ \qquad ، \qquad -t_{(۰/۰۱\ ,\ ۲۰)} = -۲/۵۲۸$$

ناحیه بحرانی یا ناحیه رد ادعا (۲/۵۲۸- < t محاسبه شده < ۲/۵۲۸-) می‌باشد حال t را محاسبه می‌کنیم . ادعای کارخانه به‌منزله ۴۰۰=μ تلقی می‌گردد .

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} = \frac{۴۰۵ - ۴۰۰}{\frac{۱۰}{\sqrt{۲۱}}} = ۲/۲۹$$

t محاسبه شده در ناحیه بحرانی نیست و نمی‌توان ادعای کارخانه مزبور را رد نمود .

ب ـ نظیر عملیات بالا را برای متغیر $(\bar{X}_1 - \bar{X}_2)$ تکرار می‌کنیم . ناحیه بحرانی همان ناحیه بند الف می‌باشد . ابتدا S_p را محاسبه می‌کنیم .

$$S_p = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2} = \frac{(۲۰\times۱۰۰)+(۳\times۱۶)}{۲۱+۴-۲} = \frac{۲۰۴۸}{۲۳} = ۸۹$$

$$S_p = ۹/۴$$

حال مقدار (t) مربوط به متغیر فوق را حساب می‌کنیم .

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(۴۰۵-۴۱۰)-(۴۰۰-۴۰۰)}{۹/۴\sqrt{\frac{۱}{۲۱}+\frac{۱}{۴}}} = \frac{-۵}{۵/۱۲۸} = -۰/۹۷۵$$

t محاسبه شده بزرگتر از (۲/۵۲۸-) بوده و در ناحیه بحرانی قرار ندارد . لذا ادعای دوم کارخانه کمپوت سازی صورت مسئله قابل قبول است .

ب ـ توزیعهای احتمال رابط بین انحراف معیار نمونه (S) و انحراف معیار جامعه (σ)

تا کنون در مباحث فصل چهارم در مورد توزیعهای احتمالی صحبت شد که به نحوی بین میانگین نمونه ($\bar{x}$) با میانگین جامعه (μ) ارتباط برقرار می‌نمودند . در فصول بعدی از این توزیعهای رابط برای بدست آوردن مقدار میانگین جامعه و یا مقایسه بین میانگین‌های جوامع ،انشاءا . . .استفاده خواهد شد .نظیر جریانات فوق برای انحراف معیار جامعه (σ) نیز می‌تواند مطرح شود ،بدین ترتیب که در بسیاری از عملیات آماری محقق با در اختیار داشتن مشخصه‌های یک یا چند نمونه ، می‌خواهد نتایجی را در مورد انحراف معیار یک یا چند جامعه بدست آورد . همانند مورد میانگین بدون وجود یک حلقه رابط بین نمونه و جامعه

نمی‌توان صرفا" با داشتن مشخصه‌های نمونه از قبیل ($\bar{x}$ و S) به‌هیچ نتیجه‌ای در مورد انحراف معیار جامعه (σ) دست یافت در مورد میانگین این حلقه‌های رابط و این پلهای ارتباطی دو متغیر (Z و t) و جداول توزیع احتمال آنها بودند متغیرهای شناخته شده‌ای نیز وجود دارند که از طریق آنها انحراف معیار نمونه (S) با همین مشخصه در جامعه یعنی (σ) مرتبط می‌شود و اگر آمارگری بخواهد با داشتن مقدار انحراف معیار یک یا چند نمونه از یک یا چند جامعه ، نتایجی را در مورد انحراف معیارهای آن جامعه یا جوامع بدست آورد می‌تواند از این متغیرها استفاده نموده و به‌کمک جداول توزیع احتمال از پیش محاسبه شده آنها نتایج دلخواه خود را بدست آورد . یکی از این متغیرها ، متغیر تصادفی چی دو (χ^2) می‌باشد . این متغیر نتیجه وجود رابطه واقعی است که بین انحراف معیار یک نمونه (S) با انحراف معیار جامعه مربوط به‌همان نمونه (σ) وجود دارد و از آنجا که جدول توزیع احتمال آن از قبل مشخص شده است ، در مواردی که محقق (S) را در اختیار داشته و مقدار (σ) را می‌خواهد می‌تواند مورد استفاده وی قرار گیرد . دومین متغیر رابط بین انحراف معیار نمونه و جامعه ، متغیر (F) است که این یکی نیز محصول ارتباط ذاتی بین (S) های دو نمونه از دو جامعه متفاوت با (σ) های مربوطه می‌باشد و توزیع احتمال شناخته شده‌ای دارد که در جداولی در اختیار محققین می‌باشد . واضح است که متغیر (F) وقتی مورد استفاده قرار می‌گیرد که محقق انحراف معیارهای دو نمونه را در اختیار داشته و تصمیم داشته باشد بین انحراف معیارهای دو جامعه مربوط به‌آن دو نمونه مقایسه به‌عمل آورد . ذیلا" هریک از این دو توزیع بطور مختصر و در حد ضرورت مورد بحث قرار می‌گیرند .

ب ـ ۱ متغیر (χ^2) و ربط بین (S و σ) از طریق آن :

در بحث میانگین مشاهده گردید که میانگین نمونه‌های مأخوذه از یک جامعه متغیری تصادفی است که مقدار آن از یک نمونه به نمونه دیگر تغییر می‌کند . واریانس یک نمونه‌نیز چنین حالتی دارد زیرا بدون شک مقادیر واریانس نمونه‌های n تایی که از یک جامعه‌ماخذ می‌شوند با یکدیگر مساوی نبوده و هر نمونه واریانسی مخصوص به‌خود دارد و بدین لحاظ واریانس نمونه یک متغیر است و از آنجا که‌نمونه‌های فوق به‌طورتصادفی‌از جامعه‌اخذ شده‌اند واریانس نمونه یک متغیر تصادفی بوده و بنابراین دارای یک توزیع احتمال خواهد بود .

مثال ۳ ـ ۴ : تمامی ده نمونه سه‌تایی بدون جایگذاری ممکن از یک جامعه ۵عضوی اخذ شده و واریانس نمونه‌ها به‌ترتیب مقادیر ۲ ، ۳ ، ۴ ، ۳ ، ۴ ، ۳ ، ۳ ، ۴ ، ۵ ، ۴ می‌باشند . مطلوبست جدول توزیع احتمال واریانسها .

پاسخ : چون نمونه‌گیری انجام شده است بنابراین احتمال تجربی یا فراوانی نسبی مطلوب

مسئله می‌باشد .

S_i^2	F_i
۲	۱
۳	۵
۴	۳
۵	۱

m=۱۰

جدول (۱۱ ــ ۴)

در جدول فوق m نشانگر تعداد نمونه‌ها می‌باشد .

بررسی‌های انجام شده نشان می‌دهد که اگر برروی متغیر تصادفی (S^2)که نمونه آن از یک جامعه نرمال اخذ شده باشد عملیاتی به‌صورت رابطه (۱۳ــ۴) انجام گیرد این متغیر تصادفی ، به متغیر تصادفی جدیدی بنام چی دو تبدیل می‌شود که این اسم تلفظ حرف یونانی (χ^2) می‌باشد و عده‌ای آنرا خی دو نیز تلفظ می‌کنند . از آنجا که این متغیر مربع متغیر چی (χ) می‌باشد چی دو یا خی دو خوانده می‌شود و عده‌ای نیز از آن با عنوان مربع چی و یا مربع خی یاد می‌کنند و بالاخره در نزد انگلیسی‌زبانان این متغیر بنام کای اسکویر معروف است .

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \qquad (۴-۱۳)$$

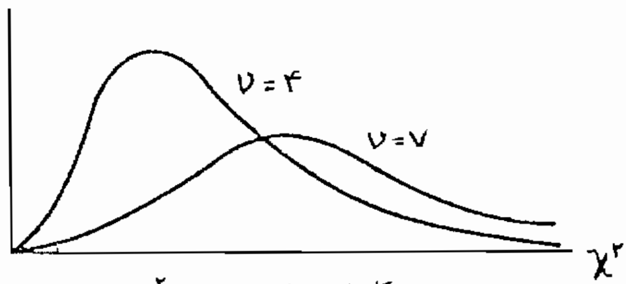
مثال ۱۴ ــ ۴ : انحراف معیار جامعه‌ای ($\sigma=3$) و انحراف معیار یک نمونه ده‌تایی از همین جامعه ($S=2$) می‌باشد . مقدار متغیر χ^2 متناظر با این نمونه را پیدا کنید .

پاسخ :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{(10-1)2^2}{3^2} = 4$$

چنان که در رابطه (۱۰ــ۴) به‌خوبی مشهود است مقدار متغیر (χ^2) تابعی از دو متغیر S^2 و ($n-1$) می‌باشد و از آنجا که مقدار S^2 نیز خود تابعی از ($n-1$) است می‌توان

گفت که مقدار (χ^2)عمدتا" از ($n-1$)تبعیت می‌کند که‌آن را درجه‌آزادی نام نهاده‌اند اگر تمام نمونه‌های n تایی ممکن از یک جامعه نرمال (مثلا" نمونه‌های ۴ تایی) اخذ شده و (S^2) هر نمونه محاسبه شود و سپس مقدار (χ^2) متناظر با هریک از S^2 ها محاسبه شده و جدول توزیع احتمال این (χ^2) ها مشخص گردد و سپس هیستوگرام فراوانی آن رسم شود می‌توان با تقریب این هیستوگرام ، منحنی توزیع احتمال (χ^2) را بدست آورد که بدون تردید شکل این منحنی برحسب درجات آزادی (ν) مختلف متفاوت خواهد بود . درشکل (۶ـ۴) دو منحنی χ^2 برای درجات آزادی ۴ و ۷ نشان داده شده است .



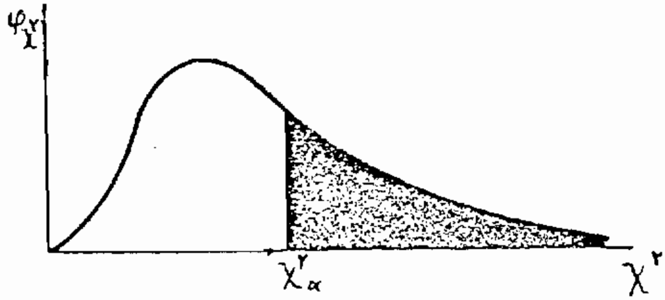
شکل (۶ـ۴) دو منحنی χ^2

متغیر (χ^2) دارای چند خاصیت مهم زیر می‌باشد . اولا" چنانکه از پاراگراف قبل به‌خوبی فهمیده می‌شود این متغیر ، متغیری ناپیوسته است که پس از رسم هیستوگرام توزیع احتمال آن ، این هیستوگرام به‌یک منحنی تقریب می‌شود. ثانیا" مقدار (χ^2) همواره‌مثبت می‌باشد زیرا هم در صورت وهم در مخرج رابطه (۱۳ـ۴) فقط‌اعداد مثبت وجود دارند و ثالثا"منحنی توزیع احتمال (χ^2)با افزایش حجم‌نمونه (n) به‌توزیع نرمال نزدیک‌می‌شود به‌طوریکه برای نمونه‌های ($n > ۱۰۰$) تایی می‌توان (χ^2) را با مختصر تغییراتی به منحنی (Z) تبدیل نمود . اما برخلاف توزیعهای (Z و t) تقارن نسبت به‌یک نقطه‌خاص ازقبیل ($\mu = ۰$ در Z و t) به‌هیچ وجه از خصوصیات شاخص یک منحنی (χ^2) به‌حساب نمی‌آید .

معمولا" در جریان یک عملیات آماری ، آمارگر به‌تابع چگالی احتمال χ^2 [$\varphi(\chi^2)$] و μ_{χ^2} و σ_{χ^2} محتاج نمی‌گردد زیرا مقادیر احتمال (χ^2) برحسب مقادیر مختلف این‌متغیر با درجات آزادی گوناگون محاسبه شده و در جداولی در اختیار محققان قرار دارد . شکل تنظیم این جداول درست مشابه جداول (t) می‌باشد یعنی در اولین ستون جدول‌درجات آزادی ، و دراولین سطر جدول مقادیر (α) مشخص شده است مقصود از (α)مساحتی‌است که هریک از مقادیر χ^2 در سمت راست منحنی خود جدا می‌کنند و به عبارت بهتر مقدار

($\chi^{۲}_{\alpha}$)گویای این مطلب‌است که مقدار متغیر ($\chi^{۲}$)سطحی معادل (α ـ۱) را در سمت‌چپ خود جدا می‌سازد . در شکل (۷ـ۴) سطح سیاه زیر منحنی معادل (α) و سطح سفید معادل (α ـ۱) می‌باشد . بدین ترتیب مقدار ($\chi^{۲}$) با درجه آزادی (۵= ν) و (α = ٪ ۵) که به‌طوراختصار به‌صورت ($\chi^{۲}_{۵\text{ و }۰/۰۵}$)نشان داده‌می‌شود ازجدول ضمیمه کتاب به‌صورت زیر بدست می‌آید .

$$\chi^{۲}_{۵\text{و}۰/۰۵}=۱۱/۰۷$$



شکل ۷ـ۴

وجود دو مشخصه $S^{۲}$ و $\sigma^{۲}$ در رابطه $\chi^{۲}$ به‌همراه مقادیر احتمال از پیش محاسبه شده و استاندارد این متغیر، به‌آمارگران این امکان را می‌دهد که با داشتن مقدار $S^{۲}$ ، نتیجه‌ای تقریبی را در مورد ($\sigma^{۲}$) به دست آورند .

مثال ۱۵ ـ ۴ : واریانس یک نمونه ده‌تایی از یک جامعه (۴= $S^{۲}$) می‌باشد . یک آمارگر می‌خواهد در مورد فرضیه (۵= $\sigma^{۲}$) تحقیق کند و ملاکهای قضاوت مورد نظر او بدین قرار می‌باشند :

الف ـ اگر ($\chi^{۲}$ محاسبه شده < $\chi^{۲}_{۹،۹۵\%}$) باشد از نظر او (۵< $\sigma^{۲}$) مردود است .

ب ـ اگر ($\chi^{۲}$ محاسبه شده > $\chi^{۲}_{۹،۵\%}$) باشد از نظر او (۵ > $\sigma^{۲}$) مردود است .

ج ـ اگر ($\chi^{۲}_{۹،۲/۵\%}$ > $\chi^{۲}$ محاسبه شده > $\chi^{۲}_{۹،۹۷/۵\%}$) باشد از نظر او (۵= $\sigma^{۲}$) قابل رد نیست . به نظر شما با توجه به‌این سه ملاک می‌توان در مورد فرضیه فوق قضاوت نمود ؟

پاسخ : ابتدا چهارمقدار $\chi^{۲}$ فوق‌را از جدول‌استخراج نموده و سپس مقدار $\chi^{۲}$مربوط به‌نمونه مأخوذه فوق را با فرض (۵= $\sigma^{۲}$) محاسبه می‌کنیم و نهایتا" با مقایسه $\chi^{۲}$محاسبه شده با $\chi^{۲}$ های جداول در مورد سه فرض فوق قضاوت می‌کنیم . براساس جدول داریم .

$$\chi^{2}_{9,95\%}=3/33 \,،\quad \chi^{2}_{9,5\%}=16/92 \,،\quad \chi^{2}_{9,2/5\%}=19/02 \,،\quad \chi^{2}_{9,97/5\%}=2/7$$

$$\chi^{2} = \frac{(n-1)S^{2}}{\sigma^{2}} = \frac{(10-1)4}{5} = 7/2$$

همان طور که ملاحظه می‌شود هیچیک از سه فرض فوق در این آزمون رد نمی‌شود . نتیجه‌ای که می‌توان گرفت این است که فرض ($\sigma^2 = 5$) برای منظور این محقق راهگشا نبوده و اوباید برای اخذ نتیجه فرض دیگری از قبیل $\sigma^2 = 2$ و غیره را آزمون کند .

در خاتمه‌مبحث توزیع (χ^2) یادآوری این نکته ضروریست که مواردا ستفاده این توزیع صرفا" در تخمین مقدار واریانس یک جامعه خلاصه نمی‌شوند . از این توزیع استفاده‌های دیگری نیز می‌شود که ذیلا" به دو مورد از مهمترین آنها اشاره می‌شود . یکی از این دومورد ایجاد ارتباط بین مقادیر نظری و تجربی یک متغیر می‌باشد . به‌عنوان مثال ونمونه‌این مورد استفاده می‌توان از آزمون همتراز بودن یک تاس نام برد . تا سی در اختیار شماست و می‌خواهید بدانید که آیا این تاس همتراز است یا خیر؟ به‌عبارت دیگر می‌خواهید بدانید که آیا می‌توان از مقادیر احتمال نظری به‌جای مقادیر احتمال تجربی در مورد آن استفاده کرد بدون آنکه در نتیجه کار تغییر عمده‌ای پدیدار گردد ؟ به‌کمک توزیع χ^2 می‌توان این مسئله را آزمون نمود . البته نتیجه‌گیری در مورد واریانس جامعه نیز که به‌کمک (χ^2) انجام می‌گرفت شکل دیگری از همین مورد استفاده می‌باشد . دومین مورد استفاده که با موردقبلی متفاوت‌است‌آزمون همبستگی بین‌دوعامل‌می‌باشد.بدین معنی‌که با استفاده از χ^2 می‌توان تحقیق نمود که آیا دو عامل مستقل از یکدیگرند یا به‌یکدیگر وابسته بوده و تغییرات یکی بر دیگری اثر می‌گذارد و بالعکس؟ به‌عنوان مثال با استفاده از χ^2 می‌توان تحقیق نمود که آیا مسئله تا هل و ازدواج بر سطح فهم دانشجو اثر حتمی دارد یا خیر؟

بحث توزیع (χ^2) را در همین‌جا خاتمه داده و به‌بررسی‌توزیع دیگری که‌بین‌واریانس نمونه و جامعه رابطه برقرار می‌کند می‌پردازیم .

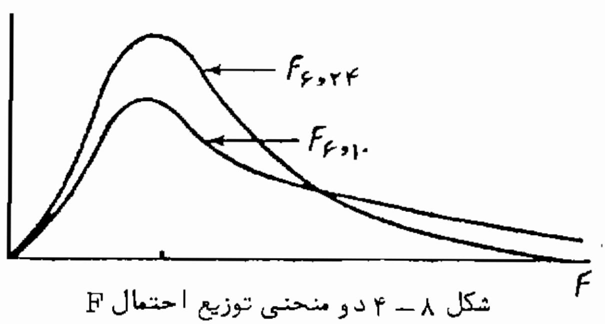
ب ــ ۲ ــ توزیع فیشر F

در بسیاری از عملیات آماری مثل تجزیه واریانس و غیره ، آمارگر نیازمند آن است که واریانس‌دو جامعه را با یکدیگر مقایسه کند . اگر این دو جامعه توزیع نرمال داشته باشند ، مقایسه مزبور به‌راحتی با استفاده از متغیری بنام فیشر (F) و توزیع احتمال مربوط به‌آن امکان‌پذیر است . رابطه مربوط به محاسبه این متغیر بصورت زیر است .

$$F = \frac{\dfrac{S_1^2}{\sigma_1^2}}{\dfrac{S_2^2}{\sigma_2^2}} = \frac{\dfrac{\chi_1^2}{\nu_1}}{\dfrac{\chi_2^2}{\nu_2}} = \frac{S_1^2}{S_2^2} \times \frac{\sigma_2^2}{\sigma_1^2} \qquad (۴-۱۴)$$

همان طورکه از رابطه (۱۴ـ۴) به‌خوبی‌دیده می‌شود مقدار متغیر F تحت تأثیرمقادیر دو متغیر S_1^2 و S_2^2 تعیین می‌گردد و از آنجا که مقادیر S_1^2 و S_2^2 نیز به‌نوبه خود از دومتغیر $(\nu_1 = n_1 - 1)$ و $(\nu_2 = n_2 - 1)$ متأثر می‌باشند می‌توان مقدار F را تحت تأثیر ν_1 و ν_2 دانست که ν_1 مربوط به (S^2) موجود در صورت کسر و ν_2 مربوط به S^2 موجود در مخرج کسر می‌باشد. به‌همین دلیل برای مشخص کردن دقیق یک مقدار F آن را با F_{ν_1, ν_2} نشان می‌دهند.

اگرتمامی نمونه‌های n_1 تایی‌از یک جامعه‌نرمال با واریانس σ_1^2 و همه‌نمونه‌های n_2تایی از جامعه نرمال دیگری با واریانس σ_2^2 را اخذ نموده و سپس تمام مقادیر F را که با ترکیبات مختلف S_1^2 و S_2^2 موجود می‌توانند ایجاد شوند، محاسبه نماییم و پس از آن منحنی توزیع احتمال مقادیر F حاصله را رسم نماییم منحنی حاصله را منحنی متغیر F با درجه آزادی $(n_1 - 1)$ و $(n_2 - 1)$ گویند. واضح است که اگر حجم هریک از نمونه‌ها تغییر کند، درجه آزادی F تغییر نموده و طبیعتاً" شکل منحنی توزیع F تغییر خواهد کرد. شکل (۸ ــ ۴) دو منحنی F را با درجات آزادی مختلف نشان می‌دهد.



شکل ۸ ــ ۴ دو منحنی توزیع احتمال F

تمامی مشخصه‌های مورد استفاده در صورت و مخرج رابطه (۱۴ـ۴) به صورت مجذور بوده و مقدارشان مثبت‌است و بنابراین همه مقادیر متغیر (F) نیز مثبت می‌باشند و ازآنجا

که در رابطه فوق از واریانسهای نمونه و جامعه برای محاسبه مقدار F استفاده شده است ، با توجه به‌اینکه توزیع متغیر F یک توزیع استاندارد بوده و مقادیر آن برای احتمالات مختلف محاسبه‌شده و در جداولی در اختیار پژوهشگران می‌باشد از این توزیع برای مقایسه‌واریانس دو جامعه و به‌عبارت‌دقیقتر برای تخمین و آزمون مقدار نسبت‌واریانس دو جامعه $\left(\frac{\sigma_1^2}{\sigma_2^2}\right)$ استفاده می‌شود .

مثال ۱۶ ــ ۴ : واریانس دو نمونه که از دو جامعه نرمال اخذ شده‌اند به‌ترتیب $S_1^2=8$ و $S_2^2=4$ می‌باشد . اگر مقدار (F) متناظر با این دو واریانس(F =۶) باشد مقدار نسبت $\frac{\sigma_1^2}{\sigma_2^2}$ را محاسبه کنید .

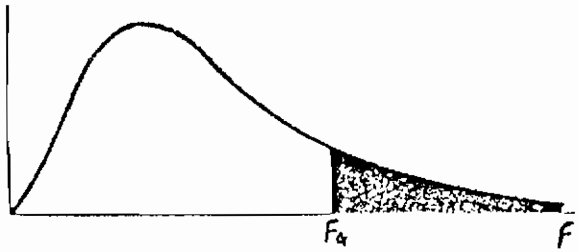
پاسخ : از رابطه (۱۴ــ۴) داریم :

$$F = \frac{S_1^2}{S_2^2} \times \frac{\sigma_2^2}{\sigma_1^2} \rightarrow 6 = \frac{8}{4} \times \frac{\sigma_2^2}{\sigma_1^2} \rightarrow \frac{\sigma_2^2}{\sigma_1^2} = \frac{6}{2} \rightarrow \frac{\sigma_1^2}{\sigma_2^2} = \frac{2}{6} = \frac{1}{3}$$

مقادیرمختلف متغیر (F)با درجات‌آزادی‌مختلف به‌ازاء دو مقدار $\alpha = 1\%$ و $\alpha = 5\%$ که در عملیات آماری بیشتر مورد احتیاج هستند محاسبه شده و در جدول ضمیمه کتاب در اختیاردانشجویان گرامی قراردارد . همانند متغیرهای t و χ^2 مقصود از F_α مقداری از متغیر F است که مطابق شکل (۹ــ۴) مساحتی معادل α را در سمت راست منحنی توزیع احتمال $F(\nu_1 \text{ و } \nu_2)$ جدا کند . معنای دقیقتر جمله فوق به‌صورت ریاضی زیر نوشته می‌شود .

$$P(F_{\nu_1,\nu_2} < F_{\alpha,\nu_1,\nu_2}) = 1 - \alpha$$

نمایش کامل یک مقدار F بصورت F_{α,ν_1,ν_2} می‌باشد . به‌عنوان مثال از ۶، ۵، $F\,1\%$ مقداری از متغیر F با درجه آزادی ($\nu_1=5$) و ($\nu_2=6$) می‌باشد که مساحتی‌معادل ۱% را در سمت راست خود جدا می‌کند و به‌عبارت بهتر۹۹% مقادیر متغیر F با همین درجه آزادی در سمت چپ این نقطه واقع می‌شوند .



شکل (۹ – ۴)

در جدول F ضمیمه کتاب مقادیر F فقط، برای سطوح اعتماد ($\alpha = 1\%$)و ($\alpha = 5\%$) داده‌شده است . عملاً" دربسیاری از مواردآمارگر محتاج دانستن مقادیر F برای ($\alpha = 99\%$) و ($\alpha = 95\%$) می‌باشد . اگر مقدار (ν_2 و ν_1 و F_α) در اختیار باشد از طریق رابطه زیر می‌توان مقدار [$F_{(1-\alpha),\nu_1,\nu_2}$] را بدست آورد .

$$F_{(1-\alpha),\nu_1,\nu_2} = \frac{1}{F_{\alpha,\nu_2,\nu_1}} \qquad (۴-۱۵)$$

مثال ۱۷ – ۴ : مقدار متغیر $F_{99\%,4,7}$ را تعیین کنید .

پاسخ : $F(\alpha = 0/99)$ در جدول وجود ندارد اما ($1-\alpha = 0/01$) می‌باشد و $F(1\%)$ در جدول موجود می‌باشد . براساس رابطه (۱۵–۴) اگر مقدار ($F\ 1\%,7,4$) را داشته باشیم می‌توانیم مقدار $F\ 99\%,4,7$ را بدست آوریم . از جدول داریم

$$F_{(1\%,7,4)} = 14/98$$

حال از رابطه (۱۵–۴) استفاده نموده و مقدار F مطلوب مسئله را تعیین می‌کنیم .

$$F_{(99\%,4,7)} = \frac{1}{F_{(1\%,7,4)}} = \frac{1}{14/98} \simeq 0/067$$

با این مبحث فصل چهارم به‌پایان می‌رسد و با پایان فصل چهارم جمیع مقدمات لازم برای تحقق اهداف عملیات آماری‌که در مقدمه کتاب گفته شد فراهم می‌گردد و انشاءالله درفصول

آینده با استفاده از مباحث این ۴ فصل چگونگی تحقیق این اهداف مورد بحث قـــرار خواهد گرفت .

مسائل فصل چهارم

۱ ــ جامعه‌ای متناهی از اعداد ۲ و ۴ تشکیل شده است.

الف ــ هیستوگرام فراوانی توزیع نمونه‌ای $\bar{X}$ را برای نمونه‌های ۴ تایی که با جایگذاری استخراج می‌شود رسم کنید.

ب ــ تحقیق کنید که $\mu_{\bar{X}} = \mu$, $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

ج ــ I را طوری تعیین کنید که ۶۸٪ میانگینهای نمونه در فاصله $\mu - I$, $\mu + I$ قرار گیرد.

۲ ــ جامعه‌ای از اعداد شامل ۵ عدد ۲، ۳، ۶، ۸ و ۱۱ می‌باشد. تمام نمونه‌های ممکن به‌حجم ۲ را یک بار با جایگذاری و بار دیگر بدون جایگذاری انتخاب می‌کنیم. برای هریک از دوحالت فوق مقادیر زیر را پیدا کنید.

الف ــ میانگین جامعه　　　　ب ــ انحراف معیار جامعه

ج ــ میانگین توزیع میانگینهای نمونه‌گیری

د ــ انحراف معیار توزیع میانگین‌های نمونه‌گیری یعنی اشتباه معیار میانگین ها.

۳ ــ اگر در مسأله ۱ یک نمونه۹ تایی با جایگذاری استخراج شود احتمال آنکه میانگین نمونه‌ای مزبور درفاصله ۲، ۴ قرارگیرد چیست؟ با فرض آنکه مقدارمیانگین‌ها به‌عدد صحیح تقریب شده باشند.

۴ ــ در امتحان درس بیولوژی که شامل ده سئوال بود نمرات ۱، ۲، ...، ۱۰ بر حسب تعداد پاسخهای صحیح به‌شاگردان داده شده است، میانگین نمرات ۶/۷ و انحراف معیار آن ۱/۲ می‌باشد. با فرض اینکه توزیع نمرات نرمال است تعیین کنید.

الف ــ درصد دانشجویانی که جمع نمرات آنها ۶ است.

ب ــ ماکزیمم نمره ده درصد دانشجویانی که کمترین نمره را گرفته‌اند.

ج ــ حداقل نمره ۱۰٪ دانشجویانی که بیشترین نمره را گرفته‌اند.

۵ ــ میانگین قیمت فروش یک نوع مشخص صابون در پانصد مغازه خواربارفروشی تهران ۵/۰۲ تومان و انحراف معیار آن ۰/۳ تومان می‌باشد. اگر با روش نمونه‌گیری تصادفی از پانصد مغازه فوق صدتا را انتخاب کرده و کالای مزبور را قیمت کنیم احتمال اینکه مجموع قیمتهای داده شده

الف ــ بین ۴۹۶ و ۵۰۰ تومان و

ب ـ بیش از ۵۱۰ تومان باشد چقدر است در صورتی که قیمت صابون‌ها توزیع نرمال داشته باشد .

۶ ـ میانگین عمر لامپهای ساخت تولید کننده A برابر ۱۴۰۰ ساعت و انحراف معیار آن ۲۰۰ ساعت است در حالی که این ارقام برای تولیدکننده B به‌ترتیب ۱۲۰۰ ساعت و ۱۰۰ ساعت می‌باشد . اگر توزیع عمر لامپ‌ها در دو جامعه نرمال و ۱۲۵ نمونه تصادفی از لامپهای هریک از تولیدکنندگان B و A آزمایش شوند احتمالات حالات ذیل چقدر است .

الف ـ میانگین عمر لامپهای ساخت A اقلاً ۱۶۰ ساعت بیش از لامپهای ساخت B باشد .

ب ـ میانگین عمر لامپهای ساخت A حداقل ۲۵۰ ساعت بیش از لامپهای ساخت B باشد .

۷ ـ وزن نوع خاصی از کاغذ هر برگ ۵/۰ گرم و انحراف معیار آن ۰/۰۲ گرم می‌باشد . دو گروه مستقل ۱۰۰۰ برگی را انتخاب می‌کنیم . احتمال اینکه وزن برگهای دو گروه بیش از ۲ گرم تفاوت داشته باشد چقدر است ؟

۸ ـ میانگین عمر نوع معینی لامپ ۱۵۰۰ ساعت و انحراف معیار آن ۱۵۰ ساعت است . سه لامپ طوری به‌هم مربوط شده‌اند که وقتی یکی بسوزد دیگری بلافاصله روشن می‌شود . با فرض اینکه توزیع عمر لامپها نرمال است احتمال وجود روشنایی را برای حالات ذیل پیدا کنید .

الف ـ حداقل ۵۰۰۰ ساعت　　　　ب ـ حداکثر ۴۲۰۰ ساعت

۹ ـ از جامعه‌ای نرمال با میانگین ۸۰ و انحراف معیار ۵ یک نمونه تصادفی ۲۵ تایی استخراج کرده‌ایم از یک جامعه‌نرمال دیگر با میانگین ۷۵ و انحراف معیار ۳ یک نمونه‌تصادفی ۳۶ تایی استخراج کرده‌ایم مطلوب است احتمال آنکه اختلاف میانگینهای این دو نمونه‌بین ۳/۵ و ۶/۵ قرار گیرد با فرض آنکه میانگینها را با ۰/۱ تقریب محاسبه کنیم .

۱۰ ـ از یک‌جامعه نرمال با میانگین ۴۰ یک نمونه ۱۶ تایی استخراج می‌کنیم مطلوبست احتمال آنکه‌میانگین نمونه $\bar{X}$ در فاصله $\mu_{\bar{X}}$ ۱/۴۷۵ و $\mu_{\bar{X}}$ - ۰/۶۷ قرار گیرد، با فرض آنکه میانگین را بتوان با هر دقت که بخواهیم محاسبه کرده و انحراف معیار نمونه ($S=2$) باشد

۱۱ ـ مسئله ۵ را با این فرض حل‌کنید که واریانس جامعه (σ^2) مجهول بوده و انحراف معیار یک نمونه ده تایی از آن $S=0/40$ باشد و ضمناً" مقادیر مطلوب به‌ترتیب به ۴۷/۶ ، ۴۸ برای بند الف ، و برای بند ب به ۵۱ تومان تغییر پیدا کنند .

۱۲ ـ با فرض ثبات سایر شرایط، در مسئله ۶ اگر واریانس جامعه محصولات تولیدکنندگان A و B (σ_1^2 و σ_2^2) مجهول بوده و مساوی یکدیگر فرض شوند و نمونه‌های مأخوذه از جامعه ۱۵ تایی باشند و بالاخره انحراف معیارهای این دو نمونه به‌ترتیب ($S_1=60$ و $S_2=70$) باشند ،بندهای الف و ب را حل‌کنید .در صورتی‌که‌توزیع عمر لامپهای دوکارخانه

نرمال باشد.

۱۳ ــ از دو جامعه نرمال با واریانسهای مجهول اما مساوی و میانگین‌های ۸۰ و ۷۵ دو نمونه ۷ و ۹ تایی استخراج می‌کنیم در صورتی که انحرافهای معیار این دو نمونه به ترتیب $S_1=5$ و $S_2=3$ باشند، مطلوب است احتمال آنکه اختلاف میانگین‌های دو نمونه بین ۱/۵ و ۸/۶ قرار گیرد مشروط بر آنکه میانگین‌ها با ۰/۱ تقریب محاسبه شوند.

۱۴ ــ از دو جامعه نرمال با میانگین‌های ($\mu_1=50$ و $\mu_2=80$) دو نمونه با حجم‌های ($n_1=16$ و $n_2=5$) اخذ شده است. اگر انحراف معیارهای دو نمونه به‌ترتیب ($S_1=7$ و $S_2=9$) بوده و واریانسهای دو جامعه نیز مساوی فرض شوند احتمالات زیر را پیداکنید.

الف ــ تفاوت میانگین دو نمونه حداقل (۳۸) و حداکثر (۴۵) باشد.

ب ــ میانگین‌های دو نمونه باهم مساوی باشند.

۱۵ ــ از دوجامعه نرمال بامیانگین و واریانس مجهول، دو نمونه‌به‌حجم های ($n_1=15$ و $n_2=10$) اخذ شده‌اند. اگر میانگین و انحراف‌معیارهای‌نمونه‌ها مقادیر ($\bar{x}_1=20$، $S_1=2$ و $\bar{x}_2=15$ و $S_2=3$) باشند و بدانیم واریانسهای دو جامعه باهم مساویند مطلوبست احتمال آنکه اختلاف میانگین‌های دو جامعه حداقل مساوی ۵ باشد.

۱۶ ــ سکه‌ای همتراز را ۲۵ بار به‌هوا پرتاب می‌کنیم. با استفاده از توزیع (Z) تعیین کنید.

الف ــ احتمال آنکه بین (۱۵ ــ ۷) بار شیر بیاید چقدر است؟

ب ــ سکه همتراز دیگری را نیز ۲۵ مرتبه پرتاب می‌کنیم. احتمال آنکه سکه اول حداقل ۸ مرتبه بیش از سکه دوم شیر بیاید چقدر است؟

۱۷ ــ اگرواریانس یک نمونه ۱۶ تایی از یک‌جامعه نرمال ($S^2=9$) باشدبا ٪۹۵ احتمال واریانس جامعه (σ^2) از چه عددی کوچکتر خواهد بود؟

۱۸ ــ واریانس یک‌جامعه نرمال ($\sigma^2=6$) می‌باشد. مطلوبست احتمال آنکه‌واریانس یک نمونه ۹ تایی از این جامعه در فاصله (۲ تا ۳) واقع شود.

۱۹ ــ واریانس دو نمونه ۲۱ و ۲۵ تایی که به‌ترتیب از دو جامعه نرمال با واریانس مجهول گرفته شده‌اند به‌ترتیب ۸و۵می‌باشد. آیا با ٪۹۰ احتمال می‌توان واریانس‌دوجامعه را مساوی دانست؟

۲۰ ــ واریانس دو جامعه به‌ترتیب ($\sigma_1^2=4$ و $\sigma_2^2=6$) می‌باشد. اگر از این دوجامعه دو نمونه ۱۵ و ۲۵ تایی اخذ کنیم مطلوبست:

الف ـ $P[0/313 \leqslant \frac{S_1^2}{S_2^2} \leqslant 1/57] = ?$

ب ـ $P[\frac{S_1^2}{S_2^2} \leqslant 0/22] = ?$

حل مسائل فرد فصل چهارم

۱ ــ الف ــ کل نمونه‌های مأخوذه و میانگین‌های آن به‌صورت جدول زیر می‌باشد .

ردیف	نمونه	$\bar{x}_i$
۱	۲،۲،۲،۲	۲
۲	۲،۲،۲،۴	۲/۵
۳	۲،۲،۴،۲	۲/۵
۴	۲،۴،۲،۲	۲/۵
۵	۴،۲،۲،۲	۲/۵
۶	۲،۲،۴،۴	۳
۷	۲،۴،۲،۴	۳
۸	۴،۲،۲،۴	۳
۹	۴،۲،۴،۲	۳
۱۰	۴،۴،۲،۲	۳
۱۱	۲،۴،۴،۲	۳
۱۲	۴،۴،۴،۲	۳/۵
۱۳	۴،۴،۲،۴	۳/۵
۱۴	۴،۲،۴،۴	۳/۵
۱۵	۲،۴،۴،۴	۳/۵
۱۶	۴،۴،۴،۴	۴

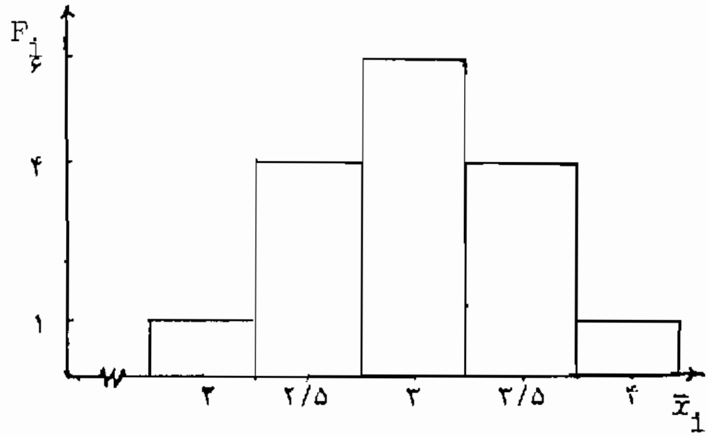
جدول ۱۲ ــ ۴

جدول توزیع نمونه‌گیری $\bar{X}$ به‌صورت زیر می‌باشد .

ردیف	$\bar{x}_i$	F
۱	۲	۱
۲	۲/۵	۴
۳	۳	۶
۴	۳/۵	۴
۵	۴	۱

جدول ۱۳ ــ ۴

هیستوگرام توزیع نمونه‌گیری $\bar{X}$ به‌صورت زیر خواهد بود .



شکل (۹-۴)

ب- $$\mu = \frac{\Sigma x_i}{N} = \frac{۲+۴}{۲} = ۳$$

تعداد نمونه‌ها را با m نشان می‌دهیم .

$$\mu_{\bar{X}} = \frac{\Sigma \bar{x}_i F_i}{m} = \frac{۲+ ۴(۲/۵) + ۶(۳)+ ۴(۳/۵)+۴}{۱۶} = \frac{۴۸}{۱۶} = ۳$$

$$\mu_{\bar{X}} = \mu_{\bar{X}} = ۳$$

$$\sigma_X = \frac{\Sigma (x_i - \mu)^۲}{N} = \frac{(۲-۳)^۲ + (۴-۳)^۲}{۲} = ۱$$

$$\sigma_{\bar{X}} = \frac{\Sigma (x_i - \mu)^۲}{m} = \frac{(۲-۳)^۲ + ۴(۲/۵-۳)^۲ + ۶(۰) + ۴(۳/۵-۳)^۲ + (۴-۳)^۲}{۱۶} = \frac{۱+۱+۰+۱+۱}{۱۶} = \frac{۴}{۱۶} = \frac{۱}{۴}$$

$$\sigma_{\bar{X}}^۲ = \frac{\sigma_X^۲}{n} = \frac{۱}{۴}$$

ج - چون توزیع نمونه‌گیری x ها نرمال است اگر دو مقدار از مقادیر Z را پیدا کنیم که ۶۸٪ مقادیر Z در وسط این دو مقدار قرار گیرند می‌توان نقاط متناظر با این دو مقدار را بر روی منحنی نرمال مربوط به منحنی توزیع نرمال $\bar{X}$ پیدا نمود . نقاط $(Z_{۰/۱۶})$ و $(Z_{۰/۸۴})$ نقاط مورد نظر می‌باشند از جدول Z داریم .

$$Z_{۰/۱۶} \simeq -۱ \quad و \quad Z_{۰/۸۴} \simeq ۱$$

$$Z = \frac{\bar{x} - \mu}{\sigma \sqrt{n}} \longrightarrow \bar{x} = \mu + Z\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$\bar{x}_{۰/۱۶} = \mu + Z_{۰/۱۶}(\sigma \sqrt{n}) = ۳ + (-۱)\left(\frac{۱}{۲}\right) = ۲/۵$$

$$\bar{x}_{۰/۸۴} = \mu + Z_{۰/۸۴}(\sigma \sqrt{n}) = ۳ + (۱)\left(\frac{۱}{۲}\right) = ۳/۵$$

بنابراین ٪۶۸ از مقادیر $\bar{X}$ بین دو نقطه ۲/۵ و ۳/۵ واقع می‌شوند و از اینجا می‌توان مقدار I را به‌دست آورد.

$$\mu + I = ۳/۵ \longrightarrow ۳ + I = ۳/۵ \longrightarrow I = ۰/۵$$

۳ــ اگر همه نمونه‌های ۸ تایی استخراج شوند توزیع $\bar{x}$ ها یک توزیع نرمال خواهد بود. بنابراین $\bar{x}$ مربوط به این نمونه نیز یکی از نقاط یک منحنی نرمال با میانگین ($E(\bar{X}) = \mu$) و واریانس ($\sigma_{\bar{X}}^{۲} = \frac{\sigma^{۲}}{n}$) خواهد بود. نتیجه اینکه احتمال مطلوب مسئله برابر با سطح زیر منحنی نرمال فوق در فاصله (۱/۵ـ۴/۵) می‌باشد. (چون مقادیر $\bar{X}$ به عدد صحیح تقریب شده‌اند) . میانگین و انحراف معیار این توزیع نرمال را از اطلاعات مسئله ۱ می‌توان محاسبه نموده و با استفاده از جداول (Z) احتمال مورد نظر را محاسبه نمود.

$$\mu_{\bar{X}} = \mu_X = ۳$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{۱}{\sqrt{۹}} = \frac{۱}{۳}$$

$$P(۱/۵ < \bar{X} < ۴/۵) = F_{\bar{X}}(۴/۵) - F_{\bar{X}}(۱/۵)$$

$$Z = \frac{\bar{x} - \mu}{\sigma \sqrt{n}}$$

$$Z(\bar{x} = ۴/۵) = \frac{۴/۵ - ۳}{\frac{۱}{۳}} = ۴/۵$$

$$Z(\bar{x} = ۱/۵) = \frac{۱/۵ - ۳}{\frac{۱}{۳}} = -۴/۵$$

از جدول تابع توزیع Z داریم:

$$P(Z < ۴/۵) = F_Z(۴/۵) \simeq ۱ \quad , F_Z(-۴/۵) \simeq ۰$$

$$P(۱/۵ < \bar{X} < ۴/۵) = [F_{\bar{X}}(۴/۵) - F_{\bar{X}}(۱/۵) = F_Z(۴/۵) - F_Z(-۴/۵)$$

$$P(۱/۵ < \bar{X} < ۴/۵) = ۱ - ۰ = ۱$$

۵ ـ بدون تردید انتخاب صد مغازه از پانصد مغازه با روش بازگردانی‌صورت‌نمی‌گیرد یعنی آمارگر از یک مغازه دو مرتبه سئوال نمی‌کند و چون حجم نمونه بزرگتر از ۳۰ و برابر صد می‌باشد لذا درصورتی که تمام‌نمونه‌های ممکن صدتایی انتخاب می‌شد میانگین‌وانحراف معیار همه نمونه‌ها به‌ترتیب زیر به‌دست می‌آمد .

$$\mu_{\bar{X}} = \mu = ۵/۰۲$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{۰/۳}{\sqrt{۱۰۰}} = ۰/۰۳$$

نمونه انتخابی ما یکی از هزاران نمونه صدتایی است که‌ممکن‌است‌ازجامعه‌فوق‌انتخاب شوند و میانگین آن ممکن‌است در هرنقطه‌ای از نقاط یک‌منحنی نرمال که توزیع میانگین‌های کلیه نمونه‌های ممکن را نشان می‌دهد قرار گیرد . بنابراین برای محاسبه احتمالات خواسته شده با یک توزیع نرمال با میانگین ۵/۰۲ و انحراف معیار ۲۷٪ سروکار خواهیم داشت و به‌ترتیب زیر عمل می‌کنیم .

الف ـ اگر میانگین قیمت صد مغازه بین ۴/۹۶ و ۵ باشد مجموع آنها بین دو عدد داده شده خواهد بود لذا این دو عدد را به‌استاندارد تبدیل می‌کنیم .

$$Z_1 = \frac{(۴/۹۶ - ۵/۰۲)}{۰/۰۳} = -۲$$

$$Z_2 = \frac{۵ - ۵/۱۲}{۰/۰۳} = -۰/۶۷$$

(سطح بین ۰ = Z و ۲- = Z) = (سطح بین ۲- = Z و ۰/۶۷- = Z) = احتمال موردنظر

— (Z = -۰/۶۷ و Z = ۰ سطح بین) = ۰/۴۷۷۲ − ۰/۲۴۵۴ = ۰/۲۳۱۸

ب ـ اگر میانگین قیمتهای داده شده از (۵/۱) تومان بیشتر باشد مجموع قیمتها از ۵۱۰ تومان بیشتر خواهد بود .

$$Z = \frac{۵/۱ - ۵/۰۲}{۰/۰۳} = ۲/۶۷$$

$$\text{احتمال مورد نظر} = (\text{سطح سمت راست } Z=۲/۶۷) = (\text{سطح سمت راست } Z=۰)$$
$$- (\text{سطح بین } Z=۰ \text{ و } Z=۲/۶۷)$$
$$= ۰/۵ - ۰/۴۹۶۲ = ۰/۰۰۳۸$$

۷ــ میانگین وزن برگهای دو گروه را با $\bar{x}_1$ و $\bar{x}_2$ نشان می‌دهیم .

$$\mu_{(\bar{x}_1-\bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = ۰/۵ - ۰/۵ = ۰$$
$$\sigma_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(۰/۰۲)^۲ + (۰/۰۲)^۲}{۱۰۰}} = ۰/۰۰۰۸۹۵$$

متغیر استاندارد شده برای تفاضل میانگین ها عبارتست از

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - ۰}{۰/۰۰۰۸۹۵}$$

اما تفاوت ۲ گرم در وزن کل دو گروه برابر $۰/۰۰۲ = \frac{۲}{۱۰۰۰}$ گرم تفاوت در میانگین‌ها می‌باشد . یعنی :

$$\bar{x}_1 - \bar{x}_2 \leqslant -۰/۰۰۲ \quad \text{ویا} \quad \bar{x}_2 - \bar{x}_1 \geqslant ۰/۰۰۲$$

که برحسب واحدهای استاندارد عبارات فوق چنین خواهند شد .

$$Z \leqslant \frac{-۰/۰۰۲-۰}{۰/۰۰۰۸۹۵} = -۲/۲۳ \quad \text{ویا} \quad Z \geqslant \frac{۰/۰۰۲-۰}{۰/۰۰۰۸۹۵} = ۲/۲۳$$

$$\text{احتمال مورنظر} = P\left\{Z \geqslant ۲/۲۳ \text{ ویا } Z \leqslant -۲/۲۳\right\} = P\,(Z \geqslant ۲/۲۳) + P\,(Z \leqslant -۲/۲۳)$$
$$= ۲(۰/۵ - ۰/۴۸۷۱) = ۰/۰۲۵۸$$

۹ــ مقادیر $\bar{X}$ با ۰/۱ تقریب بدست آمده‌اند و لذا تفاضل آنها نیز با همین تقریب بدست می‌آید . داریم .

$$P(۳/۴۵ < (\bar{x}_1 - \bar{x}_2) < ۵/۶۵) = ?$$

چون توزیع متغیر $(\bar{X}_1 - \bar{X}_2)$ یک توزیع نرمال است ، اگر فرض کنیم Z_1 نقطه متناظر با

نقطه $(\bar{x}_۱ - \bar{x}_۲ = ۳/۴۵)$ را در منحنی نرمال استاندارد نشان داده و به‌همین ترتیب $Z_۲$ نقطه متناظر با نقطه $۵/۶۵ = (\bar{x}_۱ - \bar{x}_۲)$ باشد. داریم:

$$P(۳/۴۵ < (\bar{x}_۱ - \bar{x}_۲) < ۵/۶۵) = P(Z_۱ < Z < Z_۲) = F_{(z_۲)} - F_{(z_۱)}$$

ابتدا مقادیر $Z_۱$ و $Z_۲$ را پیدا نموده وسپس $F(z_۱)$ و $F(z_۲)$ را ازجداول بدست‌می‌آوریم.

$$Z_۱ = \frac{(\bar{x}_۱ - \bar{x}_۲) - (\mu_۱ - \mu_۲)}{\sqrt{\frac{\sigma_۱^۲}{n_۱} + \frac{\sigma_۲^۲}{n_۲}}} = \frac{۳/۴۵ - (۸۰-۷۵)}{\sqrt{\frac{۲۵}{۲۵} + \frac{۹}{۳۶}}} = \frac{-۱/۵۵}{۱/۱} = -۱/۳۹$$

$$Z_۲ = \frac{(\bar{x}_۱ - \bar{x}_۲) - (\mu_۱ - \mu_۲)}{\sqrt{\frac{\sigma_۱^۲}{n_۱} + \frac{\sigma_۲^۲}{n_۲}}} = \frac{۵/۶۵ - ۵}{۱/۱} \simeq ۰/۵۸$$

از جدول $F(Z)$ داریم

$$P(Z < -۱/۳۹) = F(Z_۱) = ۰/۰۸۲۳$$

$$P(Z < -۰/۵۸) = F(Z_۲) = ۰/۷۱۹۰$$

$$P(۳/۴۵ < (\bar{x}_۱ - \bar{x}_۲) < ۵/۶۵) = F(Z_۲) - F(Z_۱) = ۰/۷۱۹۰ - ۰/۸۲۳ = ۰/۶۳۶۷$$

۱۱ـ الف $P(۴۷/۶ < \Sigma x_i < ۴۸ = ?$

با توجه به‌حجم نمونه عبارت فوق را به‌صورت زیر نیز می‌توان نوشت.

$$P(۴۷/۶ < \bar{x} < ۴/۸) = ?$$

$$P(۴/۷۶ < \bar{x} < ۴/۸) = P\left(\frac{\bar{x}_۱ - \mu}{S/\sqrt{n}} < \frac{\bar{x} - \mu}{S/\sqrt{n}} < \frac{\bar{x}_۲ - \mu}{S/\sqrt{n}}\right) = P(t_۱ < t_۹ < t_۲)$$

$$t_۱ = \frac{۴/۷۶ - ۵/۰۲}{۰/۴/\sqrt{۱۰}} = \frac{-۰/۲۶}{۰/۱۲۶} = -۲/۰۶۳$$

$$t_۲ = \frac{۴/۸ - ۵/۰۲}{۰/۱۲۶} = -۱/۷۴۶$$

$$P(-۲/۰۶۳ < t_۹ < -۱/۷۴۶) = P(t_۹ < -۱/۷۴۶) - P(t_۹ < -۲/۰۶۳)$$

از جدول t داریم .

$$P(t_{۹} < -۱/۷۴۶) = P(t_{۹} > ۱/۷۴۶) \simeq ۰/۰۶$$

$$P(t_{۹} < -۲/۰۶۳) = p(t_{۹} > ۲/۰۶۳) \simeq ۰/۰۳$$

$$(-۲/۰۶۳ < t_{۹} < -۱/۷۴۶) = ۰/۰۶ - ۰/۰۳ = ۰/۰۳$$

۱۳ ــ مسئله را با همان شرایط مسئله ۹ حل می‌کنیم زیرا شرایط کلی یکسانی بر هردو حاکم است .

$$P(۲/۴۵ < (\bar{x}_{۱} - \bar{x}_{۲}) < ۸/۶۵) = ?$$

چون واریانسهای جوامع مجهول هستند باید از توزیع t استفاده شود .

$$P(۱/۴۵ < (\bar{x}_{۱} - \bar{x}_{۲}) < ۸/۶۵) = P\left(\frac{۱/۴۵ - (\mu_{۱} - \mu_{۲})}{S_p\sqrt{\frac{۱}{n_{۱}} + \frac{۱}{n_{۲}}}} < t_{۱۴} < \frac{۸/۶۵ - (\mu_{۱} - \mu_{۲})}{S_p\sqrt{\frac{۱}{n_{۱}} + \frac{۱}{n_{۲}}}}\right)$$

$$= P(t_{۱} < t_{۱۴} < t_{۲}) = F_{t_{۲}} - F_{t_{۱}}$$

$$S_p = \frac{(n_{۱} - ۱)S_{۱}^{۲} + (n_{۲} - ۱)S_{۲}^{۲}}{n_{۱} + n_{۲} - ۲} = \frac{(۶)(۲۵) + (۸)(۹)}{۹ + ۷ - ۲} = ۱۵/۸۶$$

$$t_{۱} = \frac{۱/۴۵ - |۸۰ - ۷۵|}{\sqrt{۱۵/۸۶(\frac{۱}{۷} + \frac{۱}{۹})}} = \frac{-۳/۵۵}{۲} = -۱/۷۵۵$$

$$t_{۲} = \frac{۸/۶۵ - ۵}{۲} = ۱/۸۲۵$$

از جدول t داریم .

$$P(t_{۱۴} < -۱/۷۷۵) = F_{t_{۱}} \simeq ۰/۰۴$$

$$P(t_{۱۴} < ۱/۸۲۵) = F_{t_{۲}} \simeq ۰/۹۶۲$$

$$P(۱/۴۵ < (\bar{x}_{۱} - \bar{x}_{۲}) < ۸/۶۵) = F_{t_{۲}} - F_{t_{۱}} = ۰/۹۶۲ - ۰/۰۴ = ۰/۹۲۲$$

-۱۵

$$P((\mu_۱-\mu_۲)>۵)=?$$

$$P((\mu_۱-\mu_۲)\geqslant ۵)=P\left[\frac{(\bar{x}_۱-\bar{x}_۲)-(\mu_۱-\mu_۲)}{S_p\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}}\leq\frac{(\bar{x}_۱-\bar{x}_۲)-۵}{S_p\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}}\right]=P(t_{۲۲}<t_۱)$$

$$S_p=\sqrt{\frac{(n_۱-۱)S_۱^۲+(n_۲-۱)S_۲^۲}{n_۱+n_۲-۲}}=\sqrt{\frac{(۱۳)(۴)+(۹)(۹)}{۲۲}}=۲/۴۴$$

$$t_۱=\frac{(\bar{x}_۱-\bar{x}_۲)-۵}{۲/۴۴\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}}=\frac{(۲۰-۱۵)-۵}{۰/۱۶۷}=۰$$

$$P((\mu_۱-\mu_۲)>۵)=P(t_{۲۲}<۰)=۰/۵=۵۰\ \%$$

۱۷ ـ $\sigma^۲$ مقادیر مختلفی می‌تواند داشته باشد که مقصود مسئله مقداری از $\sigma^۲$ است که ۹۵% مقادیر آن کوچکتر از این مقدار باشند . چون $\sigma^۲$ در مخرج کسر $\chi^۲$ قرار دارد با فرض ثبات سایر شرایط و تغییر مقدار واریانس این مقدار خاص از $\sigma^۲$ مقداری از $\chi^۲_{۱۵}$ را مشخص می‌کند که از ۹۵% مقادیر $\chi^۲$ با همین درجه آزادی کوچکتر است . یعنی داریم .

$$\text{موردنظر}\quad \sigma^۲\longrightarrow\chi^۲_{۱۵،۰/۹۵}$$

از جدول داریم

$$\chi^۲_{۱۵و۰/۹۵}=۷/۲۶۱$$

$$\chi^۲=\frac{(n-۱)S^۲}{\sigma^۲}\longrightarrow\sigma^۲=\frac{(n-۱)S^۲}{\chi^۲}=\frac{۱۵(۹)}{۷/۲۶۱}=۱۸/۶$$

۱۹ ـ این مسئله در واقع مربوط به‌مبحث آزمون فرضیه‌ها می‌باشد و در اینجا اجمالا" به‌عنوان زمینه‌ای برای مباحث آن فصل به‌راه حل آن اشاره می‌شود . مطلوب مسئله دراصل تحقیق در این مورد است که آیا می‌توان گفت که ۹۰% از مقادیری که واریانسهای این دو جامعه می‌گیرند با هم مساویند ؟ اگر واقعا" این طور باشد پس ۱۰% مقادیر واریانسها مساوی هم نیستند . معقول آن است که بگوییم این ده درصد وضعیت زیر را دارند .

$$P(\sigma_۲^۲>\sigma_۱^۲)=P\left(\frac{\sigma_۱^۲}{\sigma_۲^۲}>۱\right)=P(\sigma_۲^۲<\sigma_۲^۲)=P\left(\frac{\sigma_۱^۲}{\sigma_۲^۲}<۱\right)=۰/۰۵$$

می‌دانیم که با هریک جفت از واریانسهای دو جامعه و یک جفت واریانس دو نمونه می‌توان

یک مقدار برای متغیر (F) پیدا نمود که اگر تمام نمونه‌های ممکن از این دو جامعه‌اخذ شده و با تمام مقادیر واریانسهای دو جامعه اصلی ترکیب شوند اندازه‌هایی برای F پیدا خواهد شد . که از این تعداد مقادیر (F) ، ۹۰% مربوط به‌حالت تساوی واریانسها ، ۵% مربوط به‌حالت $(\sigma_1^2 > \sigma_2^2)$ و ۵% باقیمانده مربوط به‌عکس حالت اخیر می‌باشد . چون نسبت $\left(\frac{\sigma_2^2}{\sigma_1^2}\right)$ با توجه به‌رابطه‌مربوط به متغیر F با این متغیرنسبت مستقیم داردمی‌توان نوشت

$$F\left(\frac{\sigma_2^2}{\sigma_1^2} < 1\right) < F\left(\frac{\sigma_2^2}{\sigma_1^2}\right) < F\left(\frac{\sigma_2^2}{\sigma_1^2} > 1\right)$$

با توجه‌به‌نکات فوق می‌توان به‌طورتقریبی چنین‌گفت‌که ۵% از کوچکترین‌مقادیر (F ۲۴و۲۰) مربوط به‌حالت $\left(1 < \frac{\sigma_2^2}{\sigma_1^2}\right)$ و۹۰% از مقادیر وسطی این F مربوط به‌حالت $\left(1 = \frac{\sigma_2^2}{\sigma_1^2}\right)$ و ۵% باقیمانده مربوط به‌حالت سوم می‌باشد . اکنون با انتخاب یک جفت نمونه در واقع یکی از این مقادیر (F) انتخاب می‌شوند اگر این (F) در فاصله‌ای از مقادیر (۲۴ و ۲۰ F) قرار گیرد که ۹۰% از مقادیر وسطی این متغیر را در بر می‌گیرد فرض تساوی دوواریانس قابل رد نیست در غیراین صورت رد می‌شود .ابتدا فاصله فوق و سپس مقدار F مربوط به‌اطلاعات مسئله را محاسبه وآنها را با هم مقایسه می‌کنیم به‌دلیلی که بعدا" گفته خواهد شد معمولا" چنین‌عمل می‌کنندکه‌به‌جای‌مشخص‌کردن ناحیه‌قابل‌قبول ،ناحیه، غیرقابل قبول‌را مشخص می‌کنند و آنرا ناحیه بحرانی می‌نامند .

$$\text{ناحیه بحرانی} = \left(F > F_{5\%,20,24} \quad \text{محاسبه شده} \quad < F_{95\%,20,24}\right)$$

مقادیر دو حد فوق را از جدول پیدا می‌کنیم

$$F_{(5\%, 20, 24)} = 2$$

$$F_{(95\%, 20, 24)} = \frac{1}{F_{(5\%, 24, 20)}} = \frac{1}{2/02} \simeq 0/5$$

نواحی بحرانی $F < 0/5$ و $F > 2$

$$F = \frac{S_1^2}{S_1^2} \times \frac{\sigma_1^2}{2} = \frac{8}{5} = 1/6$$

F محاسبه شده در نواحی بحرانی نیست و فرض $\sigma_2^2 \overset{\sigma_1}{=} \sigma_1^2$ رد نمی‌شود .

پاسخ مسائل زوج فصل چهارم

۲ ــ الف : ۶ و ۶ ب ــ ۳/۲۹ و ۳/۲۹ ج ــ ۶ و ۶ د ــ ۲/۳۲ و ۲/۵۱

در هر بند عدد اول از سمت راست مربوط به‌حالت با جایگذاری و عدد بعدی مربوط به‌حالت بدون جایگذاری می‌باشد .

۴ ــ الف ــ ۲۷% ب ــ ۵ ج ــ ۸/۲

۶ ــ الف ــ ۹۷/۷۲% ب ــ ۰/۶۲%

۸ ــ الف ــ ۲/۷۴% ب ــ ۱۲/۵۱ %

۱۰ ــ ۹/۵%

۱۲ ــ الف ــ تقریبا" ۹۵% ب ــ تقریبا" ۲/۴%

۱۴ ــ الف ــ تقریبا" ۹ % ب ــ تقریبا (۰)

۱۶ ــ الف ــ ۰/۸۷۶۷ ب ــ ۱/۱۹ %

۱۸ ــ تقریبا" ۰/۶

۲۰ ــ الف ــ تقریبا" ۹۵% ب ــ تقریبا" ۱%

جداول

جدول شماره ۱

ضرایب دو جمله‌ای

$$\text{Entry}: \binom{n}{x} = \binom{n}{n-x} = \frac{n!}{x!\,(n-x)!}$$

n \ x	0	1	2	3	4	5	6	7	8	9	10
5	1	5	10								
6	1	6	15	20							
7	1	7	21	35							
8	1	8	28	56	70						
9	1	9	36	84	126						
10	1	10	45	120	210	252					
11	1	11	55	165	330	462					
12	1	12	66	220	495	792	924				
13	1	13	78	286	715	1287	1716				
14	1	14	91	364	1001	2002	3003	3432			
15	1	15	105	455	1365	3003	5005	6435			
16	1	16	120	560	1820	4368	8008	11440	12870		
17	1	17	136	680	2380	6188	12376	19448	24310		

دنباله جدول ۱

n \ x	0	1	2	3	4	5	6	7	8	9	10	11
18	1	18	153	816	3060	8568	18564	31824	43758	48620		
19	1	19	171	969	3876	11628	27132	50388	75582	92378		
20	1	20	190	1140	4845	15504	38760	77520	125970	167960	184756	
21	1	21	210	1330	5985	20349	54264	116280	203490	293930	352716	
22	1	22	231	1540	7315	26334	74613	170544	319770	497420	646646	705432
23	1	23	253	1771	8855	33649	100947	245157	490314	817190	1144066	1352078
24	1	24	276	2024	10626	42504	134596	346104	735471	1307504	1961256	2496144
25	1	25	300	2300	12650	53130	177100	480700	1081575	2042975	3268760	4457100
26	1	26	325	2600	14950	65780	230230	657800	1562275	3124550	5311735	7726160
27	1	27	351	2925	17550	80730	296010	888030	2220075	4686825	8436285	13037895
28	1	28	378	3276	20475	98280	376740	1184040	3108105	6906900	13123110	21474180
29	1	29	406	3654	23751	118755	475020	1560780	4292145	10015005	20030010	34597290
30	1	30	435	4060	27405	142506	593775	2035800	5852925	14307150	30045015	54627300

دنباله جدول ۱

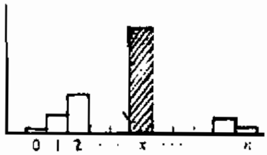
$n \backslash x$	12	13	14	15
24	2704156			
25	5200300			
26	9657700	10400600		
27	17383860	20058300		
28	30421755	37442160	40116600	
29	51895935	67863915	77558760	
30	86493225	119759850	145422675	155117520

*The author would like to thank Mr. Morty Yalovsky for computing this table, and the McGill University, Department of Computer Science for computer facilities.

جدول شماره ۲

احتمال دوجمله‌ای نقطـــه‌ای

جدول زیر احتمال مربوط به سطح هاشور خورده را ارائه می‌کند



n	x	.05	10	.15	.20	p 25	.30	.35	.40	.45	50
1	0	.9500	.9000	.8500	.8000	.7500	.7000	6500	.6000	.5500	5000
	1	.0500	.1000	.1500	.2000	.2500	.3000	.3500	4000	.4500	5000
2	0	.9025	.8100	7225	.6400	.5625	.4900	.4225	.3600	.3025	2500
	1	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	5000
	2	.0025	.0100	.0225	.0400	.0625	.0900	1225	.1600	.2025	.2500
3	0	.8574	.7290	.6141	5120	.4219	3430	2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	3750
	3	.0001	.0010	.0034	0080	0156	0270	.0429	.0640	.0911	1250
4	0	.8145	.6561	.5220	.4096	3164	.2401	.1785	1296	.0915	.0625
	1	.1715	2916	.3685	.4096	4219	.4116	.3845	3456	2995	2500
	2	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	3750
	3	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500
	4	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	0625
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312
	1	.2036	.3280	.3915	.4096	3955	.3602	.3124	.2592	.2059	1562
	2	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
	3	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
	4	.0000	.0004	.0022	.0064	.0146	0284	.0488	.0768	.1128	.1562
	5	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	0312
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
	2	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	2344
	3	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
	4	.0001	.0012	.0055	.0154	0330	.0595	.0951	.1382	.1861	2344
	5	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
	6	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	0156
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	0078
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1646	.1306	.0872	.0547
	2	.0406	.1240	.2097	2753	3115	.3177	.2985	.2613	.2140	1641
	3	.0036	.0230	.0617	1147	.1730	.2269	2679	.2903	.2918	2734
	4	.0002	.0026	.0109	.0287	.0577	0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	0000	.0001	.0002	.0006	.0016	.0037	0078
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312
	2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	2090	.1569	.1094
	3	.0064	0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	6	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	6	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039

دنباله جدول ۲

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	3874	3474	.2684	.1877	.1211	.0725	.0403	0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	2051
	7	0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
11	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0004
	1	.3293	.3835	.3248	.2362	.1549	.0932	0518	.0266	.0125	.0055
	2	.0867	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269
	3	.0137	.0710	.1517	.2215	.2581	.2568	.2254	.1774	.1259	.0806
	4	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611
	5	0001	.0025	.0132	.0388	.0803	.1321	1830	.2207	.2360	.2256
	6	.0000	.0003	0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256
	7	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	1611
	8	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	0462	.0806
	9	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	0269
	10	0000	.0000	.0000	.0000	.0000	.0000	.0002	.0007	.0021	.0054
	11	.0000	.0000	.0000	.0000	.0000	.0000	0000	.0000	.0002	.0005
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	0029
	2	.0988	.2301	2924	2835	.2323	.1678	.1088	.0639	0339	.0161
	3	0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0021	.0213	0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208
	5	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	2256
	7	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002

						p					
n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.1109	.2448	.2937	.2680	2059	.1388	.0836	.0453	.0220	.0095
	3	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
	4	.0028	.0277	.0838	.1535	.2097	.2337	.2222	.1845	.1350	.0873
	5	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571
	6	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2095
	7	0000	.0001	.0011	.0058	.0186	.0442	.0833	.1312	.1775	.2095
	8	.0000	.0000	.0001	.0011	.0047	.0142	.0336	.0656	.1089	.1571
	9	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	0495	.0873
	10	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	.0349
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0036	.0095
	12	.0000	.0000	.0000	.0000	.0000	0000	0000	.0001	.0005	.0016
	13	0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0000	.0001
14	0	.4877	2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009
	2	.1229	.2570	.2912	.2501	1802	.1134	.0634	.0317	.0141	.0056
	3	.0259	.1142	2056	2501	2402	.1943	1366	.0845	.0462	.0222
	4	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611
	5	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222
	6	0000	0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833
	7	0000	0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2095
	8	0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833
	9	0000	0000	0000	.0003	0018	0066	.0183	.0408	.0762	.1222
	10	0000	.0000	.0000	.0000	0003	.0014	.0049	.0136	0312	.0611
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	0093	.0222
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056
	13	0000	0000	.0000	.0000	.0000	0000	.0000	.0001	.0002	.0009
	14	.0000	0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	0001
5	0	4633	2059	.0874	.0352	.0134	0047	.0016	0005	.0001	.0000
	1	.3658	3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.1348	2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	0307	1285	.2184	.2501	2252	.1700	.1110	.0634	.0318	.0139
	4	0049	0428	.1156	1876	.2252	.2186	.1792	.1268	0780	.0417
	5	.0006	.0105	0449	.1032	1651	.2061	.2123	1859	.1404	.0916
	6	.0000	.0019	0132	0430	.0917	1472	.1906	.2066	.1914	.1527
	7	.0000	.0003	0030	0138	.0393	.0811	.1319	1771	.2013	.1964
	8	.0000	0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9	0000	.0000	.0001	0007	.0034	0116	.0298	.0612	.1048	.1527
	10	0000	0000	.0000	.0001	.0007	0030	.0096	.0245	.0515	.0916
	11	0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417
	12	0000	0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139
	13	0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032
	14	0000	0000	.0000	0000	.0000	.0000	.0000	.0000	.0001	.0035
	15	0000	0000	0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

دنبالـه جدول ۲

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002
	2	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018
	3	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085
	4	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278
	5	.0008	.0137	.0555	1201	.1802	.2099	.2008	.1623	.1123	.0667
	6	.0001	.0028	0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222
	7	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746
	8	.0000	.0001	0009	.0055	0197	.0487	.0923	.1417	.1812	.1964
	9	.0000	.0000	0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746
	10	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222
	11	.0000	0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667
	12	.0000	.0000	0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
	16	.0000	.0000	.0000	.0000	.0000	0000	.0000	.0000	.0000	.0000
17	0	.4181	.1668	.0631	.0225	0075	.0023	.0007	.0002	.0000	.0000
	1	.3741	.3150	.1893	.0957	0426	.0169	.0060	.0019	.0005	.0001
	2	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010
	3	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052
	4	.9076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182
	5	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	0875	.0472
	6	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944
	7	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1927	.1841	.1484
	8	.0000	.0001	.0014	.0084	0279	.0644	.1134	.1606	.1883	.1855
	9	.0000	.0000	.0003	.0021	.0093	0276	.0611	.1070	.1540	.1855
	10	.0000	.0000	.0000	.0004	.0025	0095	.0263	.0571	.1008	.1484
	11	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944
	12	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472
	13	.0000	.0000	.0000	.0000	.0000	0001	.0005	.0021	.0068	.0182
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
18	0	.3972	.1501	.0536	.0180	0056	.0016	.0004	.0001	.0000	.0000
	1	.3763	.3002	.1704	.0811	.0338	.0126	.0042	0012	.0003	.0001
	2	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006
	3	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031
	4	.0093	.0700	.1592	.2153	2130	.1681	.1104	.0614	.0291	.0117
	5	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327
	6	.0002	.0052	.0301	.0816	.1436	.1873	.1941	.1655	.1181	.0708
	7	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214
	8	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669
	9	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1894	.1855
	10	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669
	11	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214

دنبالـــه جدول ۲

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
18	12	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.3774	.2852	.1529	.0685	.0268	.0093	.0029	.0008	.0002	.0000
	2	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003
	3	.0533	.1796	.2428	.2182	.1517	.0869	.0422	.0175	.0082	.0018
	4	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074
	5	.0018	.0266	.0907	.1636	.2023	.1916	.1468	.0933	.0497	.0222
	6	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518
	7	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961
	8	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442
	9	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762
	10	.0000	.0000	.0001	.0013	.0066	.0220	.0528	.0976	.1449	.1762
	11	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0532	.0970	.1442
	12	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961
	13	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0085	.0233	.0518
	14	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
20	0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000
	1	.3774	.2702	.1368	.0576	.0211	.0068	.0020	.0005	.0001	.0000
	2	.1887	.2852	.2293	.1369	.0669	.0278	.0100	.0031	.0008	.0002
	3	.0596	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011
	4	.0133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0139	.0046
	5	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148
	6	.0003	.0089	.0454	.1091	.1656	.1916	.1712	.1244	.0746	.0370
	7	.0000	.0020	.0160	.0545	.1124	.1643	.1844	.1659	.1221	.0739
	8	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201
	9	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602
	10	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762
	11	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602
	12	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201
	13	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0146	.0366	.0739
	14	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0049	.0148
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

جدول شماره ۱

ضرایب دو جمله‌ای

$$\sum_{r=0}^{x}\binom{n}{r}p^r(1-p)^{n-r}$$

n	x	p= .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500
	1	.9975	.9900	.9775	.9600	.9375	.9100	.8775	.8400	.7975	.7500
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.9927	.9720	.9392	.8960	.8437	.7840	.7182	.6480	.5747	.5000
	2	.9999	.9990	.9966	.9920	.9844	.9730	.9571	.9360	.9089	.8750
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	.8145	.6561	.5220	.4096	.3164	.2401	17.85	.1296	.0915	.0625
	1	.9860	.9477	.8905	.8192	.7383	.6517	.5630	.4752	.3910	.3125
	2	.9995	.9963	.9880	.9728	.9492	.9163	.8735	.8208	.7585	.6875
	3	1.0000	.9999	.9995	.9984	.9961	.9919	.9850	.9744	.9590	.9375
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول ۳

5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0313
	1	.9774	.9185	.8352	.7373	.6328	.5282	.4284	.3370	.2562	.1875
	2	.9988	.9914	.9734	.9421	.8965	.8369	.7648	.6826	.5931	.5000
	3	1.0000	.9995	.9978	.9933	.9844	.9692	.9460	.9130	.8688	.8125
	4	1.0000	1.0000	.9999	.9997	.9990	.9976	.9947	.9898	.9815	.9688
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.9672	.8857	.7765	.6554	.5339	.4202	.3191	.2333	.1636	.1094
	2	.9978	.9841	.9527	.9011	.8306	.7443	.6471	.5443	.4415	.3438
	3	.9999	.9987	.9941	.9830	.9624	.9295	.8826	.8208	.7447	.6562
	4	1.0000	.9999	.9996	.9984	.9954	.9891	.9777	.9590	.9308	.8906
	5	1.0000	1.0000	1.0000	.9999	.9998	.9993	.9982	.9959	.9917	.9844
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.9556	.8503	.7166	.5767	.4449	.3294	.2338	.1586	.1024	.0625
	2	.9962	.9743	.9262	.8520	.7564	.6471	.5323	.4199	.3164	.2266
	3	.9998	.9973	.9879	.9667	.9294	.8740	.8002	.7102	.6083	.5000
	4	1.0000	.9998	.9988	.9953	.9871	.9712	.9444	.9037	.8471	.7734
	5	1.0000	1.0000	.9999	.9996	.9987	.9962	.9910	.9812	.9643	.9375
	6	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9984	.9963	.9922
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.9428	.8131	.6572	.5033	.3671	.2553	.1691	.1064	.0632	.0352
	2	.9942	.9619	.8948	.7969	.6785	.5518	.4278	.3154	.2201	.1445
	3	.9996	.9950	.9786	.9437	.8862	.8059	.7064	.5941	.4770	.3633
	4	1.0000	.9996	.9971	.9896	.9727	.9420	.8939	.8263	.7396	.6367
	5	1.0000	1.0000	.9998	.9988	.9958	.9887	.9747	.9502	.9115	.8555

دنباله جدول ۳

n	x	p= .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
	6	1.0000	1.0000	1.0000	.9999	.9996	.9987	.9964	.9915	.9819	.9648
	7	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9993	.9983	.9961
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.9288	.7748	.5995	.4362	.3003	.1960	.1211	.0705	.0385	.0195
	2	.9916	.9470	.8591	.7382	.6007	.4628	.3373	.2318	.1495	.0898
	3	.9994	.9917	.9661	.9144	.8343	.7297	.6089	.4826	.3614	.2539
	4	1.0000	.9991	.9944	.9804	.9511	.9012	.8283	.7334	.6214	.5000
	5	1.0000	.9999	.9994	.9969	.9900	.9747	.9464	.9006	.8342	.7461
	6	1.0000	1.0000	1.0000	.9997	.9987	.9957	.9888	.9750	.9502	.9102
	7	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9986	.9962	.9909	.9805
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9992	.9980
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.9139	.7361	.5443	.3758	.2440	.1493	.0860	.0464	.0233	.0107
	2	9885	.9298	.8202	.6778	.5256	.3828	.2616	.1673	.0996	.0547
	3	.9990	.9872	.9500	.8791	.7759	.6496	.5138	.3823	.2660	.1719
	4	.9999	.9984	.9901	.9672	.9219	.8497	.7515	.6331	.5044	.3770
	5	1.0000	.9999	.9986	.9936	.9803	.9527	.9051	.8338	.7384	.6230
	6	1.0000	1.0000	.9999	.9991	.9965	.9894	.9740	.9452	.8980	.8281
	7	1.0000	1.0000	1.0000	.9999	.9996	.9984	.9952	.9877	.9726	.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9983	.9955	.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول

11	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005
	1	.8981	.6974	.4922	.3221	.1971	.1130	.0606	.0302	.0139	.0059
	2	.9848	.9104	.7788	.6174	.4552	.3127	.2001	.1189	.0652	.0327
	3	.9984	.9815	.9306	.8389	.7133	.5696	.4256	.2963	.1911	.1133
	4	.9999	.9972	.9841	.9496	.8854	.7897	.6683	.5328	.3971	.2744
	5	1.0000	.9997	.9973	.9883	.9657	.9218	.8513	.7535	.6331	.5000
	6	1.0000	1.0000	.9997	.9980	.9924	.9784	.9499	.9006	.8262	.7256
	7	1.0000	1.0000	1.0000	.9998	.9988	.9957	.9878	.9707	.9390	.8867
	8	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9980	.9941	.9852	.9673
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9993	.9978	.9941
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9995
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.8816	.6590	.4135	.2749	.1584	.0850	.0424	.0196	.0083	.0032
	2	.9804	.8891	.7358	.5583	.3907	.2528	.1513	.0834	.0421	.0193
	3	.9978	.9744	.9078	.7946	.6488	.4925	.3467	.2253	.1345	.0730
	4	.9998	.9957	.9761	.9274	.8424	.7237	.5833	.4382	.3044	.1938
	5	1.0000	.9995	.9954	.9806	.9456	.8822	.7873	.6652	.5269	.3872
	6	1.0000	.9999	.9993	.9961	.9857	.9614	.9154	.8418	.7393	.6128
	7	1.0000	1.0000	.9999	.9994	.9972	.9905	.9745	.9427	.8883	.8062
	8	1.0000	1.0000	1.0000	.9999	.9996	.9983	.9944	.9847	.9644	.9270
	9	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9992	.9972	.9921	.9807
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9989	.9968
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول ۳

n	x	$p=$ .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.8646	.6213	.3983	.2336	.1267	.0637	.0296	.0126	.0049	.0017
	2	.9755	.8661	.6920	.5017	.3326	.2025	.1132	.0579	.0269	.0112
	3	.9969	.9658	.8820	.7473	.5843	.4206	.2783	.1686	.0929	.0461
	4	.9997	.9935	.9658	.9009	.7940	.6543	.5005	.3530	.2279	.1334
	5	1.0000	.9991	.9925	.9700	.9198	.8346	.7159	.5744	.4268	.2905
	6	1.0000	.9999	.9987	.9930	.9757	.9376	.8705	.7712	.6437	.5000
	7	1.0000	1.0000	.9998	.9988	.9944	.9818	.9538	.9023	.8212	.7095
	8	1.0000	1.0000	1.0000	.9998	.9990	.9960	.9874	.9679	.9302	.8666
	9	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9975	.9922	.9797	.9539
	10	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9987	.9959	.9888
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9983
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.8470	.5846	.3567	.1979	.1010	.0475	.0205	.0081	.0029	.0009
	2	.9699	.8416	.6479	.4481	.2811	.1608	.0839	.0398	.0170	.0065
	3	.9958	.9559	.8535	.6982	.5213	.3552	.2205	.1243	.0632	.0287
	4	.9996	.9908	.9533	.8702	.7415	.5842	.4227	.2793	.1672	.0898
	5	1.0000	.9985	.9885	.9561	.8883	.7805	.6405	.4859	.3373	.2120
	6	1.0000	.9998	.9978	.9884	.9617	.9064	.8164	.6925	.5461	.3953
	7	1.0000	1.0000	.9997	.9976	.9897	.9685	.9247	.8499	.7414	.6047
	8	1.0000	1.0000	1.0000	.9996	.9978	.9917	.9757	.9417	.8811	.7880
	9	1.0000	1.0000	1.0000	1.0000	.9997	.9983	.9940	.9825	.9574	.9102
	10	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9989	.9961	.9886	.9713
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9978	.9935
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9991
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول

15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000
	1	.8290	.5490	.3186	.1671	.0802	.0353	.0142	.0052	.0017	.0005
	2	.9638	.8159	.6042	.3980	.2361	.1268	.0617	.0271	.0107	.0037
	3	.9945	.9444	.8227	.6482	.4613	.2969	.1727	.0905	.0424	.0176
	4	.9994	.9873	.9383	.8358	.6865	.5155	.3519	.2173	.1204	.0592
	5	.9999	.9978	.9832	.9389	.8516	.7216	.5643	.4032	.2608	.1509
	6	1.0000	.9997	.9964	.9819	.9434	.8689	.7548	.6098	.4522	.3036
	7	1.0000	1.0000	.9994	.9958	.9827	.9500	.8868	.7869	.6535	.5000
	8	1.0000	1.0000	.9999	.9992	.9958	.9848	.9578	.9050	.8182	.6964
	9	1.0000	1.0000	1.0000	9999	.9992	.9963	.9876	.9662	.9231	.8491
	10	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9972	.9907	.9745	.9408
	11	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9981	.9937	.9824
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9989	.9963
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.8108	.5147	.2839	.1407	.0635	.0261	.0098	.0033	.0010	.0003
	2	.9571	.7892	.5614	.3518	.1971	.0994	.0451	.0183	.0066	.0021
	3	.9930	.9316	.7899	.5981	.4050	.2459	.1339	.0651	.0281	.0106
	4	.9991	.9830	.9209	.7982	.6302	.4499	.2892	.1666	.0853	.0384
	5	.9999	.9967	.9765	.9183	.8103	.6598	.4900	.3288	.1976	.1051
	6	1.0000	.9995	.9944	.9733	.9204	.8247	.6881	.5272	.3660	.2272
	7	1.0000	.9999	.9989	.9930	.9729	.9256	.8406	.7161	.5629	.4018
	8	1.0000	1.0000	.9998	.9985	.9925	.9743	.9329	.8577	.7441	.5982
	9	1.0000	1.0000	1.0000	.9998	.9984	.9929	.9771	.9417	.8750	.7728
	10	1.0000	1.0000	1.0000	1.0000	.9997	.9984	.9938	.9809	.9514	.8949
	11	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987	.9951	.9851	.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9991	.9965	.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول ۳

n	x	$p=$ .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000
	1	.7922	.4818	.2525	.1182	.0501	.0193	.0067	.0021	.0006	.0001
	2	.9497	.7618	.5198	.3096	.1637	.0774	.0327	.0123	.0041	.0012
	3	.9912	.9174	.7556	.5489	.3530	.2019	.1028	.0464	.0184	.0064
	4	.9988	.9779	.9013	.7582	.5739	.3887	.2348	.1260	.0596	.0245
	5	.9999	.9953	.9681	.8943	.7653	.5968	.4197	.2639	.1471	.0717
	6	1.0000	.9992	.9917	.9623	.8929	.7752	.6188	.4478	.2902	.1662
	7	1.0000	.9999	.9983	.9891	.9598	.8954	.7872	.6405	.4743	.3145
	8	1.0000	1.0000	.9997	.9974	.9876	.9597	.9006	.8011	.6626	.5000
	9	1.0000	1.0000	1.0000	.9995	.9969	.9873	.9617	.9081	.8166	.6855
	10	1.0000	1.0000	1.0000	.9999	.9994	.9968	.9880	.9652	.9174	.8338
	11	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9970	.9894	.9699	.9283
	12	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9975	.9914	.9755
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9981	.9936
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9988
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000
	1	.7735	.4503	.2241	.0991	.0395	.0142	.0046	.0013	.0003	.0001
	2	.9419	.7338	.4797	.2713	.1353	.0600	.0236	.0082	.0025	.0007
	3	.9891	.9018	.7202	.5010	.3057	.1646	.0783	.0328	.0120	.0038
	4	.9985	.9718	.8794	.7164	.5187	.3327	.1886	.0942	.0411	.0154
	5	.9998	.9936	.9581	.8671	.7175	.5344	.3550	.2088	.1077	.0481
	6	1.0000	.9988	.9882	.9487	.8610	.7217	.5491	.3743	.2258	.1189
	7	1.0000	.9998	.9973	.9837	.9431	.8593	.7283	.5634	.3915	.2403
	8	1.0000	1.0000	.9995	.9957	.9807	.9404	.8609	.7368	.5778	.4073
	9	1.0000	1.0000	.9999	.9991	.9946	.9790	.9403	.8653	.7473	.5927
	10	1.0000	1.0000	1.0000	.9998	.9988	.9939	.9788	.9424	.8720	.7597

دنباله جدول ۳

	11	1.0000	1.0000	1.0000	1.0000	.9998	.9986	.9938	.9797	.9463	.8811
	12	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9986	.9942	.9817	.9519
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987	.9951	.9846
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9990	.9962
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9993
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.7547	.4203	.1985	.0829	.0310	.0104	.0031	.0008	.0002	.0000
	2	.9335	.7054	.4413	.2369	.1113	.0462	.0170	.0055	.0015	.0004
	3	.9868	.8850	.6841	.4551	.2631	.1332	.0591	.0230	.0077	.0022
	4	.9980	.9648	.8556	.6733	.4654	.2822	.1500	.0696	.0280	.0096
	5	.9998	.9914	.9463	.8369	.6678	.4739	.2968	.1629	.0777	.0318
	6	1.0000	.9983	.9837	.9324	.8251	.6655	.4812	.3081	.1727	.0835
	7	1.0000	.9997	.9959	.9767	.9225	.8180	.6656	.4878	.3169	.1796
	8	1.0000	1.0000	.9992	.9933	.9713	.9161	.8145	.6675	.4940	.3238
	9	1.0000	1.0000	.9999	.9984	.9911	.9674	.9125	.8139	.6710	.5000
	10	1.0000	1.0000	1.0000	.9997	.9977	.9895	.9653	.9115	.8159	.6762
	11	1.0000	1.0000	1.0000	1.0000	.9995	.9972	.9886	.9648	.9129	.8204
	12	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9969	.9884	.9658	.9165
	13	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9969	.9891	.9682
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9972	.9904
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9978
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9996
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول ۳

n	x	$p=$ .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
	0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000
	1	.7358	.3917	.1756	.0692	.0243	.0076	.0021	.0005	.0001	.0000
	2	.9245	.6769	.4049	.2061	.0913	.0355	.0121	.0036	.0009	.0002
	3	.9841	.8670	.6477	.4114	.2252	.1071	.0444	.0160	.0049	.0013
	4	.9974	.9568	.8298	.6296	.4148	.2375	.1182	.0510	.0189	.0059
	5	.9997	.9887	.9327	.8042	.6172	.4164	.2454	.1256	.0553	.0207
	6	1.0000	.9976	.9781	.9133	.7858	.6080	.4166	.2500	.1299	.0577
	7	1.0000	.9996	.9941	.9679	.8982	.7723	.6010	.4159	.2520	.1316
	8	1.0000	.9999	.9987	.9900	.9591	.8867	.7624	.5956	.4143	.2517
	9	1.0000	1.0000	.9998	.9974	.9861	.9520	.8782	.7553	.5914	.4119
	10	1.0000	1.0000	1.0000	.9994	.9961	.9829	.9468	.8725	.7507	.5881
	11	1.0000	1.0000	1.0000	.9999	.9991	.9949	.9804	.9435	.8692	.7483
	12	1.0000	1.0000	1.0000	1.0000	.9998	.9987	.9940	.9790	.9420	.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9985	.9935	.9786	.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9984	.9936	.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9985	.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

*The author would like to thank Mr. Morty Yalovsky for computing this table, and the McGill University, Department of Computer Science for computer facilities.

جدول شماره ۴

احتمال تجمعی پواسن

$$\sum_{r=0}^{x} \frac{\lambda^r}{r!} e^{-\lambda}$$

$x \backslash \lambda$	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.9953	.9825	.9631	.9384	.9098	.8781	.8442	.8088	.7725	.7358
2	.9998	.9989	.9964	.9921	.9856	.9769	.9659	.9526	.9371	.9197
3	1.0000	.9999	.9997	.9992	.9982	.9966	.9942	.9909	.9865	.9810
4	1.0000	1.0000	1.0000	.9999	.9998	.9996	.9992	.9986	.9977	.9963
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9997	.9994
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$x \backslash \lambda$	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.6990	.6626	.6268	.5918	.5578	.5249	.4932	.4628	.4337	.4060
2	.9004	.8795	.8571	.8335	.8088	.7834	.7572	.7306	.7037	.6767
3	.9743	.9662	.9569	.9463	.9344	.9212	.9068	.8913	.8747	.8571
4	.9946	.9923	.9893	.9857	.9814	.9763	.9704	.9636	.9559	.9473
5	.9990	.9985	.9978	.9968	.9955	.9940	.9920	.9896	.9868	.9834
6	.9999	.9997	.9996	.9994	.9991	.9987	.9981	.9974	.9966	.9955
7	1.0000	1.0000	.9999	.9999	.9998	.9997	.9996	.9994	.9992	.9989
8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9998	.9998
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول ۴

x \ λ	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.3796	.3546	.3309	.3084	.2873	.2674	.2487	.2311	.2146	.1991
2	.6496	.6227	.5960	.5697	.5438	.5184	.4936	.4695	.4460	.4232
3	.8386	.8194	.7993	.7787	.7576	.7360	.7141	.6919	.6696	.6172
4	.9379	.9275	.9162	.9041	.8912	.8774	.8629	.8477	.8318	.8153
5	.9796	.9751	.9700	.9643	.9580	.9510	.9433	.9349	.9258	.9161
6	.9941	.9925	.9906	.9884	.9858	.9828	.9794	.9756	.9713	.9665
7	.9985	.9980	.9974	.9967	.9958	.9947	.9934	.9919	.9901	.9881
8	.9997	.9995	.9994	.9991	.9989	.9985	.9981	.9976	.9969	.9962
9	.9999	.9999	.9999	.9998	.9997	.9996	.9995	.9993	.9991	.9989
10	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9998	.9998	.9997
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ λ	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	.0450	.0408	.0369	.0334	.0302	.0273	.0247	.0224	.0202	.0183
1	.1847	.1712	.1586	.1468	.1359	.1257	.1162	.1074	.0992	.0916
2	.4012	.3799	.3594	.3397	.3208	.3027	.2854	.2689	.2531	.2381
3	.6248	.6025	.5803	.5584	.5366	.5152	.4942	.4735	.4532	.4335
4	.7982	.7806	.7626	.7442	.7254	.7064	.6872	.6678	.6481	.6288
5	.9057	.8946	.8829	.8705	.8576	.8441	.8301	.8156	.8006	.7851
6	.9612	.9554	.9490	.9421	.9347	.9267	.9182	.9091	.8995	.8893
7	.9858	.9832	.9802	.9769	.9733	.9692	.9648	.9599	.9546	.9489
8	.9953	.9943	.9931	.9917	.9901	.9883	.9863	.9840	.9815	.9786
9	.9986	.9982	.9978	.9973	.9967	.9960	.9952	.9942	.9931	.9919
10	.9996	.9995	.9994	.9992	.9990	.9987	.9984	.9981	.9977	.9972
11	.9999	.9999	.9998	.9998	.9997	.9996	.9995	.9994	.9993	.9991
12	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999	.9998	.9998	.9997
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

دنباله جدول ۴

$x \backslash \lambda$	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
0	.0166	.0150	.0136	.0123	.0111	.0101	.0091	.0082	.0074	.0067
1	.0845	.0780	.0719	.0663	.0611	.0563	.0518	.0477	.0439	.0404
2	.2238	.2102	.1974	.1851	.1736	.1626	.1523	.1425	.1333	.1247
3	.4142	.3954	.3772	.3594	.3423	.3257	.3097	.2942	.2793	.2650
4	.6093	.5898	.5704	.5512	.5321	.5132	.4946	.4763	.4582	.4405
5	.7693	.7531	.7367	.7199	.7029	.6858	.6684	.6510	.6335	.6160
6	.8786	.8675	.8558	.8436	.8311	.8180	.8046	.7908	.7767	.7622
7	.9427	.9361	.9290	.9214	.9134	.9049	.8960	.8867	.8769	.8666
8	.9755	.9721	.9683	.9642	.9597	.9549	.9497	.9442	.9382	.9319
9	.9905	.9889	.9871	.9851	.9829	.9805	.9778	.9749	.9717	.9682
10	.9966	.9959	.9952	.9943	.9933	.9922	.9910	.9896	.9880	.9863
11	.9989	.9986	.9983	.9980	.9976	.9971	.9966	.9960	.9953	.9945
12	.9997	.9996	.9995	.9993	.9992	.9990	.9988	.9986	.9983	.9980
13	.9999	.9999	.9998	.9998	.9997	.9997	.9996	.9995	.9994	.9993
14	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999	.9999	.9998	.9998
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999

$x \backslash \lambda$	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	.0061	.0055	.0050	.0045	.0041	.0037	.0033	.0030	.0027	.0025
1	.0372	.0342	.0314	.0289	.0266	.0244	.0224	.0206	.0189	.0174
2	.1165	.1088	.1016	.0948	.0884	.0824	.0768	.0715	.0666	.0620
3	.2513	.2381	.2254	.2133	.0217	.1906	.1800	.1700	.1604	.1512
4	.4231	.4061	.3895	.3733	.3575	.3422	.3272	.3127	.2987	.2851

دنباله جدول ۴

x \ λ	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
5	.5984	.5809	.5635	.5461	.5289	.5119	.4950	.4783	.4619	.4457
6	.7474	.7324	.7171	.7017	.6860	.6703	.6544	.6384	.6224	.6063
7	.8560	.8449	.8335	.8217	.8095	.7970	.7841	.7710	.7576	.7440
8	.9252	.9181	.9106	.9027	.8944	.8857	.8766	.8672	.8574	.8472
9	.9644	.9603	.9559	.9512	.9462	.9409	.9352	.9292	.9228	.9161
10	.9844	.9823	.9800	.9775	.9747	.9718	.9686	.9651	.9614	.9574
11	.9937	.9927	.9916	.9904	.9890	.9875	.9859	.9841	.9821	.9799
12	.9976	.9972	.9967	.9962	.9955	.9949	.9941	.9932	.9922	.9912
13	.9992	.9990	.9988	.9986	.9983	.9980	.9977	.9973	.9969	.9964
14	.9997	.9997	.9996	.9995	.9994	.9993	.9991	.9990	.9988	.9986
15	.9999	.9999	.9999	.9998	.9998	.9998	.9997	.9996	.9996	.9995
16	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999	.9999	.9999	.9998
17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999

x \ λ	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	.0022	.0020	.0018	.0017	.0015	.0014	.0012	.0011	.0010	.0009
1	.0159	.0146	.0134	.0123	.0113	.0103	.0095	.0087	.0080	.0073
2	.0577	.0536	.0498	.0463	.0430	.0400	.0371	.0344	.0320	.0296
3	.1425	.1342	.1264	.1189	.1118	.1052	.0988	.0928	.0871	.0818
4	.2719	.2592	.2469	.2351	.2237	.2127	.2022	.1920	.1823	.1730
5	.4298	.4141	.3988	.3837	.3690	.3547	.3406	.3270	.3137	.3007

دنباله جدول ۴

6	.5902	.5742	.5582	.5423	.5265	.5108	.4953	.4799	.4647	.4497
7	.7301	.7160	.7017	.6873	.6728	.6581	.6433	.6285	.6136	.5987
8	.8367	.8259	.8148	.8033	.7916	.7796	.7673	.7548	.7420	.7291
9	.9090	.9016	.8939	.8858	.8774	.8686	.8596	.8502	.8405	.8305
10	.9531	.9486	.9437	.9386	.9332	.9274	.9214	.9151	.9084	.9015
11	.9776	.9750	.9723	.9693	.9661	.9627	.9591	.9552	.9510	.9467
12	.9900	.9887	.9873	.9857	.9840	.9821	.9801	.9779	.9755	.9730
13	.9958	.9952	.9945	.9937	.9929	.9920	.9909	.9898	.9885	.9872
14	.9984	.9981	.9978	.9974	.9970	.9966	.9961	.9956	.9950	.9943
15	.9994	.9993	.9992	.9990	.9988	.9986	.9984	.9982	.9979	.9976
16	.9998	.9997	.9997	.9996	.9996	.9995	.9994	.9993	.9992	.9990
17	.9999	.9999	.9999	.9999	.9998	.9998	.9998	.9997	.9997	.9996
18	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999	.9999	.9999
19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$x \backslash \lambda$	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	.0008	.0007	.0007	.0006	.0006	.0005	.0005	.0004	.0004	.0003
1	.0067	.0061	.0056	.0051	.0047	.0043	.0039	.0036	.0033	.0030
2	.0275	.0255	.0236	.0219	.0203	.0188	.0174	.0161	.0149	.0138
3	.0767	.0719	.0674	.0632	.0591	.0554	.0518	.0485	.0453	.0424
4	.1641	.1555	.1473	.1395	.1321	.1249	.1181	.1117	.1055	.0996
5	.2881	.2759	.2640	.2526	.2414	.2307	.2203	.2103	.2006	.1912
6	.4349	.4204	.4060	.3920	.3782	.3646	.3514	.3384	.3257	.3134
7	.5838	.5689	.5541	.5393	.5246	.5100	.4956	.4812	.4670	.4530
8	.7160	.7027	.6892	.6757	.6620	.6482	.6343	.6204	.6065	.5925
9	.8202	.8096	.7988	.7877	.7764	.7649	.7531	.7411	.7290	.7166
10	.8942	.8867	.8788	.8707	.8622	.8535	.8445	.8352	.8257	.8159
11	.9420	.9371	.9319	.9265	.9208	.9148	.9085	.9020	.8952	.8881
12	.9703	.9673	.9642	.9609	.9573	.9536	.9496	.9454	.9409	.9362
13	.9857	.9841	.9824	.9805	.9784	.9762	.9739	.9714	.9687	.9658
14	.9935	.9927	.9918	.9908	.9897	.9886	.9873	.9859	.9844	.9827
15	.9972	.9969	.9964	.9959	.9954	.9948	.9941	.9934	.9926	.9918

$x \backslash \lambda$	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
16	.9989	.9987	.9985	.9983	.9980	.9978	.9974	.9971	.9967	.9963
17	.9996	.9995	.9994	.9993	.9992	.9991	.9989	.9988	.9986	.9984
18	.9998	.9998	.9998	.9997	.9997	.9996	.9996	.9995	.9994	.9993
19	.9999	.9999	.9999	.9999	.9999	.9999	.9998	.9998	.9998	.9997
20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999
21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$x \backslash \lambda$	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
0	.0003	.0003	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001
1	.0028	.0025	.0023	.0021	.0019	.0018	.0016	.0015	.0014	.0012
2	.0127	.0118	.0109	.0100	.0093	.0086	.0079	.0073	.0068	.0062
3	.0396	.0370	.0346	.0323	.0301	.0281	.0262	.0244	.0228	.0212
4	.0940	.0887	.0837	.0789	.0744	.0701	.0660	.0621	.0584	.0550
5	.1822	.1736	.1653	.1573	.1496	.1422	.1352	.1284	.1219	.1157
6	.3013	.2896	.2781	.2670	.2562	.2457	.2355	.2256	.2160	.2068
7	.4391	.4254	.4119	.3987	.3856	.3728	.3602	.3478	.3357	.3239
8	.5786	.5647	.5507	.5369	.5231	.5094	.4958	.4823	.4689	.4557
9	.7041	.6915	.6788	.6659	.6530	.6400	.6269	.6137	.6006	.5874
10	.8058	.7955	.7850	.7743	.7634	.7522	.7409	.7294	.7178	.7060
11	.8807	.8731	.8652	.8571	.8487	.8400	.8311	.8220	.8126	.8030
12	.9313	.9261	.9207	.9150	.9091	.9029	.8965	.8898	.8829	.8758
13	.9628	.9595	.9561	.9524	.9486	.9445	.9403	.9358	.9311	.9261
14	.9810	.9791	.9771	.9749	.9726	.9701	.9675	.9647	.9617	.9585
15	.9908	.9898	.9887	.9875	.9862	.9848	.9832	.9816	.9798	.9780
16	.9958	.9953	.9947	.9941	.9934	.9926	.9918	.9909	.9899	.9889
17	.9982	.9979	.9977	.9973	.9970	.9966	.9962	.9957	.9952	.9947

دنباله جدول ۴

18	.9992	.9991	.9990	.9989	.9987	.9985	.9983	.9981	.9978	.9976
19	.9997	.9997	.9996	.9995	.9995	.9994	.9993	.9992	.9991	.9989
20	.9999	.9999	.9998	.9998	.9998	.9998	.9997	.9997	.9996	.9996
21	1.0000	1.0000	.9999	.9999	.9999	.9999	.9999	.9999	.9998	.9998
22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999

x \ λ	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0
0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0000
1	.0011	.0010	.0009	.0009	.0008	.0007	.0007	.0006	.0005	.0005
2	.0058	.0053	.0049	.0045	.0042	.0038	.0035	.0033	.0030	.0028
3	.0198	.0184	.0172	.0160	.0149	.0138	.0129	.0120	.0111	.0103
4	.0517	.0486	.0456	.0429	.0403	.0378	.0355	.0333	.0312	.0293
5	.1098	.1041	.0986	.0935	.0885	.0838	.0793	.0750	.0710	.0671
6	.1978	.1892	.1808	.1727	.1649	.1574	.1502	.1433	.1366	.1301
7	.3123	.3010	.2900	.2792	.2687	.2584	.2485	.2388	.2294	.2202
8	.4426	.4296	.4168	.4042	.3918	.3798	.3676	.3558	.3442	.3328
9	.5742	.5611	.5479	.5349	.5218	.5089	.4960	.4832	.4705	.4579
10	.6941	.6820	.6699	.6576	.6453	.6329	.6205	.6080	.5955	.5830
11	.7932	.8732	.7730	.7626	.7520	.7412	.7303	.7193	.7081	.6968
12	.8684	.8607	.8529	.8448	.8364	.8279	.8191	.8101	.8009	.7916
13	.9210	.9156	.9100	.9042	.8981	.8919	.8853	.8786	.8716	.8645
14	.9552	.9517	.9480	.9441	.9400	.9357	.9312	.9265	.9216	.9165
15	.9760	.9738	.9715	.9691	.9665	.9638	.9609	.9579	.9546	.9513
16	.9878	.9865	.9852	.9838	.9823	.9806	.9789	.9770	.9751	.9730
17	.9941	.9934	.9927	.9919	.9911	.9902	.9892	.9881	.9870	.9857
18	.9973	.9969	.9966	.9962	.9957	.9952	.9947	.9941	.9935	.9928
19	.9988	.9986	.9985	.9983	.9980	.9978	.9975	.9972	.9969	.9965
20	.9995	.9994	.9993	.9992	.9991	.9990	.9989	.9987	.9986	.9984
21	.9998	.9998	.9997	.9997	.9996	.9996	.9995	.9995	.9994	.9993

دنباله جدول ۴

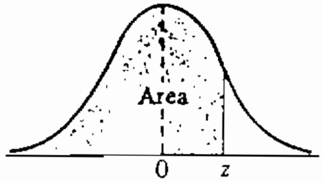
x \ λ	9.1	9.2	9/3	9.4	9.5	9·6	9.7	9.8	9.9	10.0
22	.9999	.9999	.9999	.9999	.9999	.9998	.9998	.9998	.9997	.9997
23	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999	.9999	.9999
24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ λ	11	12	13	14	15	16	17	18	19	20
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2	.0012	.0005	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
3	.0049	.0023	.0011	.0005	.0002	.0001	.0000	.0000	.0000	.0000
4	.0151	.0076	.0037	.0018	.0009	.0004	.0002	.0001	.0000	.0000
5	.0375	.0203	.0107	.0055	.0028	.0014	.0007	.0003	.0002	.0001
6	.0786	.0458	.0259	.0142	.0076	.0040	.0021	.0010	.0005	.0003
7	.1432	.0895	.0540	.0316	.0180	.0100	.0054	.0029	.0015	.0008
8	.2320	.1550	.0998	.0621	.0374	.0220	.0126	.0071	.0039	.0021
9	.3405	.2424	.1658	.1094	.0699	.0433	.0261	.0154	.0080	.0050
10	.4599	.3472	.2517	.1757	.1185	.0774	.0491	.0304	.0183	.0108
11	.5793	.4616	.3532	.2600	.1848	.1270	.0847	.0549	.0347	.0214
12	.6887	.5760	.4631	.3585	.2676	.1931	.1350	.0917	.0606	.0390
13	.7813	.6815	.5730	.4644	.3632	.2745	.2009	.1426	.0984	.0661
14	.8540	.7720	.6751	.5704	.4657	.3675	.2808	.2081	.1497	.1049
15	.9074	.8444	.7636	.6694	.5681	.4667	.3715	.2867	.2148	.1565
16	.9441	.8987	.8355	.7559	.6641	.5660	.4677	.3751	.2920	.2211
17	.9678	.9370	.8905	.8272	.7489	.6593	.5640	.4686	.3784	.2970
18	.9823	.9626	.9302	.8826	.8195	.7423	.6550	.5622	.4695	.3814
19	.9907	.9787	.9573	.9235	.8752	.8122	.7363	.6509	.5606	.4703
20	.9953	.9884	.9750	.9521	.9170	.8682	.8055	.7307	.6472	.5591

21	.9977	.9939	.9859	.9712	.9469	.9108	.8615	.7991	.7255	.6437
22	.9990	.9970	.9924	.9833	.9673	.9418	.9047	.8551	.7931	.7206
23	.9995	.9985	.9960	.9907	.9805	.9633	.9367	.8989	.8490	.7875
24	.9998	.9993	.9980	.9950	.9888	.9777	.9594	.9317	.8933	.8432
25	.9999	.9997	.9990	.9974	.9938	.9869	.9748	.9554	.9269	.8878
26	1.0000	.9999	.9995	.9987	.9967	.9925	.9848	.9718	.9514	.9221
27	1.0000	.9999	.9998	.9994	.9983	.9959	.9912	.9827	.9687	.9475
28	1.0000	1.0000	.9999	.9997	.9991	.9978	.9950	.9897	.9805	.9657
29	1.0000	1.0000	1.0000	.9999	.9996	.9989	.9973	.9941	.9881	.9782
30	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9986	.9967	.9930	.9865
31	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9993	.9982	.9960	.9919
32	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9990	.9978	.9953
33	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9995	.9988	.9973
34	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9985
35	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9992
36	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9996
37	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998
38	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
39	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999

*The author would like to thank Mr. Morty Yalovsky for computing this table, and the McGill University, Department of Computer Science for computer facilities.

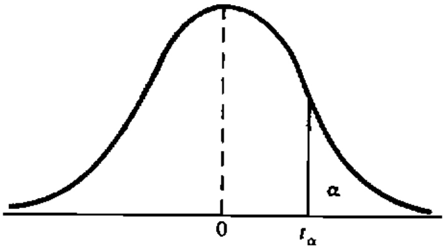
جدول شماره ۵
سطح زیر منحنی نرمال



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

جدول شماره ۶

مقادیر بحرانی توزیع t

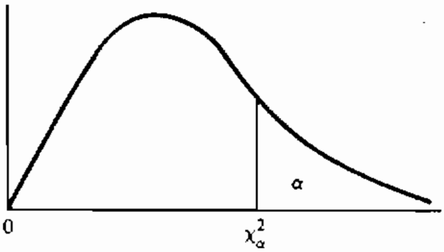


ν	α 0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf	1.282	1.645	1.960	2.326	2.576

* Table A.5 is taken from Table IV of R. A. Fisher: *Statistical Methods for Research Workers*, published by Oliver & Boyd Ltd, Edinburgh, by permission of the author and publishers.

جدول شماره ۷

مقادیر بحرانی توزیع مربع‌کای

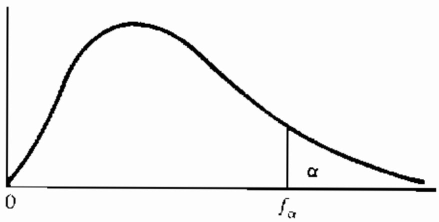


ν	α							
	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	3.841	5.024	6.635	7.879
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

* Abridged from Table 8 of *Biometrika Tables for Statisticians*, Vol. I, by permission of E. S. Pearson and the Biometrika Trustees

جدول شماره ۸

مقادیر بحرانی توزیع F



$f_{0.05}(\nu_1, \nu_2)$

ν_2	ν_1 = 1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

جدول شماره ۸

ادامه مقادیر بحرانی توزیع F

f ۰/۰۵ (ν_1, ν_2)

ν_2	ν_1									
	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

جدول شماره ۸

ادامه مقادیر بحرانی توزیع F

F ۰/۰۱ (ν_1 و ν_2)

ν_2	ν_1 = 1	2	3	4	5	6	7	8	9
1	4052	4999.5	5403	5625	5764	5859	5928	5981	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

جدول شماره ۸
ادامه مقادیر بحرانی توزیع F
$f_{0/01}(\gamma_1 \text{ و } \gamma_2)$

ν_2	ν_1 10	12	15	20	24	30	40	60	120	∞
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	[illegible]	14.37	14.28	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

* Reproduced from Table 18 of *Biometrika Tables for Statisticians*, Vol. I, by permission of E. S. Pearson and the Biometrika Trustees.

FERDOWSI UNIVERSITY OF MASHHAD

Publication No. 122

STATISTICS

for Economics and Business Students

by:

Sayyid Ali Akbar Razmi

Lecturer in Economics

FERDOWSI UNIVERSITY PRESS

1992