بسم الله الرحمن الرحیم

آشنائی با آمار و احتمالات و کاربرد آن در بیمه

مدرس : احمد امینی آموزش مجازی بیمه (Lms) شرکت تعاونی پژوهشگران رایانگان فردیس

(این جزوه با هدف کاربرد آمار و احتمالات در بیمه در دست تدوین بوده و بعلت کمبود وقت در تنظیم مطالب دارای اشکالات احتمالی است . بدیهی است مطالب تکمیلی از منابع معرفی شده " آمار و کاربرد آن در مدیریت (جلد یک) دکتر عادل آذر و منصور مومنی ، نشر سمت ۱۴۰۰ " مطالعه فرمائید.)

اهداف دوره : کسب مهارت و آشنائی با آمار و احتمالات و کاربرد آن در بیمه

www.asia26090.ir

سرفصل های دوره :

۱. تعریف علم آمار و طبقه بندی انواع آمار ، آشنائی با مفاهیم اصلی آمار
۲. آمار توصیفی (توزیع ها ، فراوانی ها و نمودارهای مختلف)، قانون اعداد بزرگ
۳. شاخصهای اصلی تمایل مرکزی و کاربرد آن در صنعت بیمه
۴. شاخصهای پراکندگی در بیمه
۵. قوانین شمارش

فصل اول : تعریف علم آمار و طبقه بندی انواع آمار ، آشنائی با مفاهیم اصلی آمار

➢ تعریف علم آمار Statistics

آمار چیست : آمار در لغت به معنای شمارش و حساب است.

آمار علمی است که به روش گردآوری، خلاصه کردن و طبقه بندی، توصیف و تحلیل داده ها و استفاده از آنها برای پیش بینی، برآورد و تخمین رویدادها می پردازد. آمار یکی از مهمترین ابزارهای پردازش داده ها و تصمیم گیری است. آمار را می توان علم و هنر جمع آوری و تجزیه و تحلیل داده ها دانست به طوری که در نهایت بتوان آنچه استخراج شده است را تفسیر و استنباط کرد. امروزه به ندرت می توان بدون استفاده از تحلیل آماری اقدام به تفسیر و تحلیل نتایج به دست آمده از تحقیقات و پژوهش های علمی کرد.

به گزارشی که شامل جداول و نمودارهای آماری و تحلیل و تفسیر آن ها می شود تحلیل(تفسیر) آماری گفته می شود.

فرآیند تحلیل آماری را به دو بخش تقسیم می‌کنند: آمار استنباطی و آمار توصیفی. مبنا و پایه این تقسیم بندی به دو دسته استنباطی و توصیفی این است که بشر با توجه به محدودیت زمان و سایر امکانات تصمیم گرفت تا از استقراء استفاده کند. بر اساس استقراء می توان یافته هایی را که از «اجزا» به دست می آید را به «کل» نسبت داد.این بدین معناست که یافته های مربوط به نمونه ها، در مورد کل جامعه که احتمالا به دلیل محدودیت ها و مشکلات دیگر قابل بررسی به صورت کلی نبوده اند، حکم می شود.

➢ آمار توصیفی چیست؟

وقتی توده ای از اطلاعات برای تحقیق جمع آوری می شود باید ابتدا سازماندهی و خلاصه شود به طوری که به صورت معنی داری قابل درک و ارتباط باشند. برای این منظور از روش آمار توصیفی (Descriptive Statistics) استفاده می شود. در تجزیه و تحلیل توصیفی، پژوهشگر داده های جمع آوری شده را با کمک روش های آمار توصیفی از جمله تنظیم جداول توزیع فراوانی، خلاصه کرده و به کمک نمودار آن‌ها را نمایش می دهد و در نهایت با کمک سایر شاخص های آمار توصیفی آن ها را خلاصه می کند.



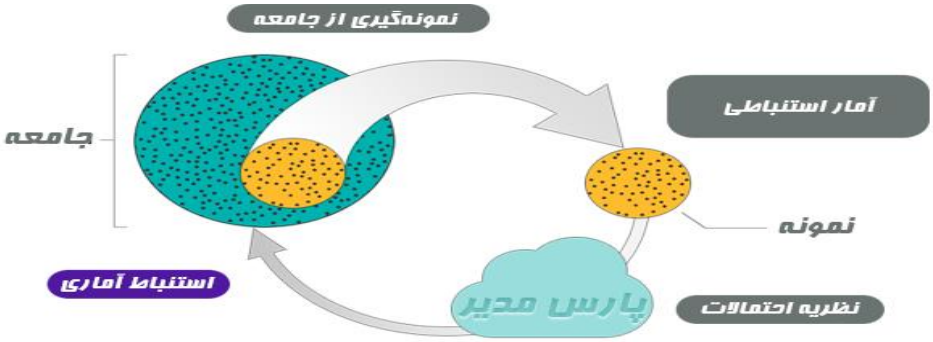
www.asia26090.ir

➢ روش های آمار توصیفی

- تشکیل جدول توزیع فراوانی(مطلق ، نسبی ، تجمعی)
- ترسیم نمودار (هیستوگرام ، چند ضلعی، میله‌ای ، تجمعی (اوجایو)، دایره‌ای ، نمودار شاخه و برگ پراکنش)
- محاسبه شاخص های مرکزی (نما(Mode) ، میانه (Median) و میانگین(Mean)
- محاسبه شاخص های پراکندگی(دامنه ، چارک ، واریانس، انحراف معیار(استاندارد)
- محاسبه همبستگی
- رگراسیون و پیش بینی

➢ آمار استباطی چیست؟

در آمار توصیفی به جمع آوری و تمرکز به روی خصوصیات و ویژگی های مهم داده ها می پردازیم و به نوعی می توان گفت نقش آمار توصیفی جمع آوری، خلاصه کردن و توصیف اطلاعات کمی به دست آمده از نمونه ها یا جامعه است ، اما قرار نیست کار یک محقق با توصیف اطلاعات به پایان برسد بلکه باید آنچه از بررسی گروه نمونه به دست آورده است را به گروه های مشابه بزرگ تر تعمیم دهد.



در آمار استنباطی ((Inferential، پژوهشگر با استفاده از مقادیر نمونه، آماره ها را محاسبه کرده و با کمک تخمین یا آزمون فرض آماری، آماره ها را به پارامترهای جامعه تعمیم می دهد. هدف آمار استنباطی، برآورد کردن خصوصیات و ویژگی های جامعه آماری است.

- آزمون های آمار استنباطی

آزمون های آماری مورد استفاده جهت تجزیه و تحلیل اطلاعات به دست آمده از یک گروه کوچک یا نمونه و تعمیم آن به جامعه مورد نظر با توجه به مقیاس اندازه گیری متغیرها، به دو گروه پارامتریک و ناپارامتریک تقسیم می شوند.

آزمون های پارامتریک، به تجزیه و تحلیل اطلاعات در سطح مقیاس فاصله ای و نسبتی می پردازند و حداقل شاخص آماری آ نها، میانگین و واریانس است.

در حالیکه آزمون های ناپارامتریک، به تجزیه و تحلیل اطلاعات در سطح مقیاس اسمی و رتبه ای می پردازند که شاخص آماری آنها میانه و نما است.

جامعه آماری و نمونه آماری

چنانکه گفته شد، جامعه آماری و نمونه آماری یکی از مباحث اولیه در تحقیق می‌باشد. پژوهشگران معمولاً کار خود را با توصیف اطلاعات شروع نموده (آمار توصیفی) و سعی می‌کنند آن‌چه را از بررسی گروه نمونه به دست آورده‌اند، به گروه‌های مشابه بزرگ‌تر یا جامعه آماری تعمیم دهند (امار استنباطی). بدین منظور در این مبحث با مفاهیم جامعه و نمونه آماری و و موضوعات مرتبط با آن‌ها آشنا می شویم.

- جامعه آماری
- نمونه آماری
- نمونه گیری
- انواع نمونه‌گیری
- تعیین حجم نمونه
- نکاتی که در تعیین حجم نمونه باید توجه نمود
- روشهای برآورد حجم نمونه

جامعه آماری

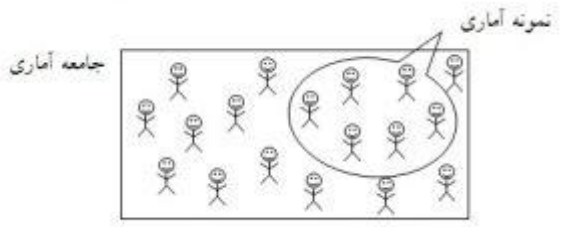
جامعه آماری عبارتست از مجموعه تمام افراد، گروه‌ها، اشیاء و یا رویدادهایی که دارای یک یا چند ویژگی مشترک باشند. تعداد اعضای جامعه را حجم یا اندازه جامعه می‌نامند و با حرف بزرگ N نشان می‌دهند.

مثال :جامعه کارکنان شاغل در بانک تجارت شهر تهران

نمونه آماری

نمونه آماری گروه کوچکتری از جامعه است که طبق ضابطه‌ای معین برای مشاهده و تجزیه و تحلیل انتخاب می‌شود و باید معرف جامعه باشد. نتایج نمونه ای را که معرف جامعه نباشد نمی‌توان به جامعه تعمیم داد. تعداد اعضای نمونه را با حرف کوچک n نشان می دهند.

مثال :کارکنان شاغل در بانک تجارت منطقه ۵ شهر تهران



نمونه‌گیری

نمونه‌گیری به منظور گردآوری داده‌های مورد نیاز درباره افراد جامعه و برآورد مقادیر جامعه به کمک مقادیر نمونه انجام می‌شود. نمونه‌گیری باعث صرفه‌جویی در هزینه و زمان است و کار تحقیق را ساده و امکان‌پذیر می‌سازد.

به طور کلی برای گردآوری اطلاعات دو روش وجود دارد:

الف) سرشماری: اگر محقق پژوهش خود را بر تمامی افراد جامعه اجرا کند روش او سرشماری خواهد بود. یعنی محقق باید تمامی افراد جامعه را تک تک مورد بررسی و آزمون قرار دهد. هزینه، نیروی انسانی و مدت زمان لازم برای انجام شمارش کامل (برای گردآوری داده ها) به میزانی است که معمولاً اجرای آن توصیه نمی شود.

ب) نمونه گیری: نمونه گیری عبارت است از «انتخاب افراد گروه نمونه از میان اعضای یک جامعه ی تعریف شده ی آماری براساس اصول و قواعد خاص». در این شیوه داده‌ها از همه افراد جامعه گردآوری نمی شود.

انواع نمونه‌گیری

نمونه‌گیری تصادفی یا احتمالی : در نمونه‌گیری تصادفی احتمال انتخاب شدن برای همه اعضای جامعه یکسان و معلوم است. هیچ عاملی جز شانس و تصادف در انتخاب شدن افراد نمونه از جامعه دخالت ندارد. نمونه‌گیری تصادفی انواع مختلفی دارد که عبارتند از: نمونه‌گیری تصادفی ساده، نمونه‌گیری منظم (سیستماتیک)، نمونه‌گیری طبقه‌ای (یا نسبی) و نمونه‌گیری خوشه‌ای (تک مرحله‌ای و چند مرحله‌ای). پژوهشگر می‌تواند بنا بر ویژگی‌های جامعه آماری خود یکی از این روش‌ها را برگزیند.

نمونه‌گیری غیر تصادفی یا غیر احتمالی : در نمونه‌گیری غیر تصادفی، احتمال انتخاب شدن برای همه اعضای جامعه نامعین و نامعلوم است. نمونه انتخاب شده به این روش معرف جامعه نیست و نمی‌توان نتایج حاصل از آن را به جامعه تعمیم داد. نمونه‌گیری غیرتصادفی شامل نمونه‌گیری اتفاقی (یا در دسترس)، هدفمند (یا قضاوتی)، سهمیه‌ای و شبکه‌ای (یا گلوله برفی) می‌باشد.

جدول ۱: روش‌های نمونه‌گیری براساس قابلیت تعمیم

روش‌های نمونه‌گیری تصادفی	روش‌های نمونه‌گیری غیر تصادفی
۱. تصادفی ساده: همه افراد شانس برابر و مستقل برای انتخاب شدن دارند.	۱. در دسترس: افراد فقط به دلیل سهولت، سادگی و در دسترس بودن انتخاب می‌شوند.
۲. منظم (سیستماتیک): شکل اصلاح شده روش تصادفی ساده است. اما در این روش، نمونه با نظم خاصی بر اساس فهرستی که از قبل تنظیم شده انتخاب می‌شود. یعنی انتخاب افراد مستقل از یکدیگر نیست.	۲. هدفمند (یا قضاوتی): نمونه براساس قضاوت شخصی و هدف‌های مطالعه انتخاب می‌شود.
۳. طبقه‌ای (یا نسبی): در این روش، نمونه به گونه‌ای انتخاب می‌شود که زیرگروه‌های آن به همان نسبتی که در جامعه وجود دارند، در نمونه نیز حضور داشته باشند.	۳. سهمیه‌ای: معادل نمونه‌گیری طبقه‌ای است که محقق سعی می‌کند نسبت یا ویژگی‌های جامعه در نمونه نیز وجود داشته باشد.
۴. خوشه‌ای: در این روش واحد نمونه‌گیری گروه یا خوشه‌ای از افراد است.	۴. شبکه‌ای (یا گلوله برفی): زمانی که شناخت اعضای یک جامعه دشوار باشد و نمونه‌ها از

خوشه‌ای چند مرحله‌ای: فهرست نمونه‌گیری دو بار یا بیش از دو بار تهیه می‌شود.	یکدیگر شناخت داشته باشند، هر یک از افراد جامعه عضو دیگر را به پژوهشگر معرفی می‌کند.

تعیین حجم نمونه

هر چه حجم نمونه بزرگتر باشد، میزان اشتباهات در نتیجه‌گیری کاهش می‌یابد و بر عکس. حجم نمونه ارتباط بسیار نزدیکی با آزمون فرضیه پوچ (صفر) دارد. بدین ترتیب که هر چه اندازه گروه نمونه بزرگتر باشد محقق با قاطعیت بیشتری فرض پوچ را که واقعاً نادرست است رد می‌کند.

نکاتی که در تعیین حجم نمونه باید توجه نمود:

۱. هر قدر حجم جامعه کوچکتر باشد نسبت بیشتری از جامعه باید در نمونه وجود داشته باشد و هر قدر حجم جامعه بزرگتر باشد نسبت کمتری از جامعه باید در نمونه وجود داشته باشد. اگر حجم جامعه ۳۰ نفر یا کمتر باشد محقق تقریباً باید کل جامعه را به عنوان نمونه انتخاب کند. یعنی از روش سرشماری استفاده نماید.

۲. اگر حجم جامعه بزرگ باشد، باید نمونه بزرگتری انتخاب شود. همچنین توجه داشته باشید که با افزایش حجم جامعه، حجم نمونه با میزان کمتری افزایش می یابد. در حجم جامعه بالاتر از ۳۸۰ نفر، حجم نمونه تقریباً ثابت می ماند.

۳. هر چه جامعه ناهمگون‌تر و یا به عبارت دیگر واریانس آن بیشتر باشد، محقق باید نمونه بزرگتری را انتخاب کند.

۴. محققان باید همیشه نمونه ای بزرگتر از آنچه که واقعاً می‌خواهند انتخاب کنند چرا که همیشه احتمال ریزش و افت آزمودنی‌ها وجود دارد. افت آزمودنی‌ها به ویژه در تحقیقات پانل (panel) روی می دهد. تحقیقاتی که در آن یک گروه از آزمودنی‌ها در طول زمان چندین بار مورد اندازه گیری قرار می‌گیرند. معمولاً محقق باید قبل از انجام تحقیق انتظار ۱۰ تا ۲۵ درصد ریزش نمونه را داشته باشد.

۵. حجم نمونه تا حد زیادی به هدف و روش تحقیق بستگی دارد. در تحقیقات قوم‌شناسی یا کیفی معمولاً از نمونه کوچک استفاده می‌شود. برای پژوهش‌های توصیفی، مانند مطالعات میدانی و زمینه‌یابی، نمونه‌ای به حجم حداقل ۱۰۰ نفر نیاز است. در پژوهش‌های همبستگی برخی منابع حداقل حجم نمونه را ۳۰ نفر و برخی دیگر ۵۰ نفر ذکر کرده‌اند. در پژوهش‌های از نوع آزمایشی و علّی- مقایسه‌ای، حجم نمونه حداقل ۱۵ نفر در هر گروه توصیه می‌شود. در تحقیقاتی که نیاز به طبقه‌بندی جامعه برای نمونه‌گیری می‌باشد، حداقل نمونه هر طبقه بین ۲۰ تا ۵۰ نفر است.

۶. هنگامیکه پیش‌بینی تفاوت یا همبستگی پایین است، اندازه نمونه باید بزرگ باشد. در تحقیقاتی که انتظار داریم برای گروههای مختلف تفاوت اندکی در متغیر وابسته بدست آوریم. یا در مطالعاتی که به منظور تعیین ارتباط صورت می‌گیرند و همبستگی پایین مورد انتظار است.

www.asia26090.ir

۷. زمانی که گروههای انتخاب شده باید به زیرگروههای دیگری تقسیم شوند و سپس این زیرگروهها مقایسه گردند، لازم است نمونه بزرگ باشد. تا زیرگروهها تعداد کافی آزمودنی را دربرگیرند.

۸. زمانی که در تحقیق متغیرهای کنترل نشده زیادی وجود دارند، انتخاب نمونه با اندازه بزرگ ضروری است.

۹. در برخی از تحقیقات، انتخاب نمونه ای با اندازه‌گیری کوچک مناسبتر از انتخاب یک نمونه با اندازه‌گیری بزرگ است. این بیشتر در مورد تحقیقاتی که هدف آنها اجرای نقش، مصاحبه های عمیق و اندازه گیریهای ذهنی است، صدق می‌کند.

۱۰. زمانی که وسیله پایایی برای اندازه گیری متغیر وابسته وجود ندارد. پایایی ابزار اندازه گیری بدان معنا است که هر گاه این ابزار در شرایط و زمانهای مختلف بکار رود، آزمودنیهای یکسان دارای نمره‌های مشابهی گردند.

۱۱. نوع مقیاس اندازه‌گیری در تعیین حجم نمونه موثر است. برای داده‌هایی که از مقیاس اسمی به دست می‌آیند، در مقایسه با مقیاس فاصله‌ای و نسبی به نمونه بزرگتری نیاز داریم.

۱۲. سطح اطمینان و خطای نمونه‌گیری در تعیین حجم نمونه موثر است. زمانی که محقق سطح بالاتری از اطمینان یا معنی دار بودن آماری مثلاً ۹۹ درصد اطمینان با خطای ۱ درصد را ملاک ارزیابی اطلاعات تحقیق خود قرار می‌دهد لازم است حجم نمونه او بزرگتر انتخاب شود.

۱۳. در تحقیقات چندمتغیره، حجم نمونه باید چند برابر (ترجیحاً ۱۰ برابر) تعداد متغیرها در پژوهش باشند.

- **اریبی:** (Bias) **اریبی** که در فارسی گاهی آن را «تورِشی» نیز می‌نامند، ابزار و معیاری برای نشان دادن میزان نزدیکی پارامتر جامعه به برآوردگر آن پارامتر است. از آنجایی که براساس هر نمونه تصادفی، برآوردگر مربوط به پارامتر مقدار متفاوتی خواهد داشت، انتظار داریم که با تکرار نمونه‌گیری‌ها، متوسط مقدار برآوردگرهای حاصل، با پارامتر واقعی جامعه تقریبا برابر شود. در این حالت میزان اریبی برآوردگر (فاصله آن از مقدار واقعی پارامتر) باید با مقدار صفر برابر باشد. چنین برآوردگری را یک برآوردگر «نااریب (Unbiased) «می‌گویند. در حقیقت وجود چنین خاصیتی برآوردگرهای پارامتر را به دو کلاس برآوردگر اریب و نااریب تفکیک می‌کند.

www.asia26090.ir

فصل دوم

آمار توصیفی (توزیع ها ، فراوانی ها و نمودارهای مختلف)، قانون اعداد بزرگ

نمایش داده ها :

مقادیر عددی در محاسبات ریاضی و همچنین آمار، نقش مهمی دارند. ولی در بسیاری از مواقع، باید رابطه‌ای بین این مقادیر را قبل از اندازه‌گیری آن‌ها ارائه کنیم. به این ترتیب رابطه یا فرمول مربوطه تعیین شده و پس از اندازه‌گیری کمیت‌های مورد نظر، محاسبات صورت می‌گیرد. برای اینکه بتوانیم قبل از اندازه‌گیری، برای این کمیت‌ها، از آن‌ها استفاده کنیم، لازم است که به آن‌ها اسمی بدهیم. این امر منجر به ظهور متغیرها در ریاضیات و علوم دیگر شد. معمولا متغیرها را با اسامی لاتین مثل X یا Y نشان می‌دهند

متغیرها، بستر یا مکانی برای قرارگیری مقادیر اندازه‌گیری شده هستند. فرمول‌ها و محاسبات براساس نمادهایی که برای متغیرها به کار می‌رود، نمایش داده شده و به کار گرفته می‌شوند. برای مثال فرمول زیر برای محاسبه میانگین ۵ مشاهده مناسب خواهد بود.

میانگین = 5 / X1+X2+X3+X4+X5

همانطور که می‌بینید، بخشی از فرمول (صورت کسر) با متغیرها و بخشی (مخرج) با عدد تکمیل شده است. واضح است که منظور از X1 مقدار مشاهده اول و X2 مقدار مشاهده دوم و الی آخر است. ولی از آنجایی که هنوز اندازه‌ای مشخص نشده برای نمایش این مقادیر از اسامی متغیرها استفاده شده است.

شاید بتوان فرمول بالا را به صورت زیر به بیان نوشتاری درآوریم.

مقدار اول را با مقدار دوم، سپس با مقدار سوم و چهارم و پنجم جمع و حاصل را بر ۵ تقسیم کن.

به این ترتیب یک عبارت فارسی طولانی به یک عبارت ریاضی ساده و کوتاه درآمد. این هنر استفاده از انواع متغیرها در آمار و فرمول‌های مربوطه محسوب می‌شود.

اینکه چه مقادیری در این متغیرها قرار می‌گیرند، باعث طبقه‌بندی متغیرها می‌شود. از این جهت می‌توان انواع متغیرها در آمار را به صورت زیر تفکیک کرد.

- متغیرها با مقادیر کیفی
- متغیرها با مقادیر کمی

از یک نظر ممکن است انواع متغیرها در آمار را براساس مجموعه اعدادی که می‌پذیرند، طبقه بندی کنند.

برای مثال اگر دامنه تغییرات مقادیر عددی، فقط «مجموعه اعداد صحیح» شامل مقادیر مثبت یا منفی (Whole Numbers)باشند، آن را «**متغیرهای گسسته**» (Discreet Variable) می‌نامند. بعبارتی مقدیر کسری یا اعشاری در مورد آنها وجود ندارد مثلا تعداد افراد خانواده ، کلاس ، روستا که الزاما" بصورت جدا یا گسسته هستند مانند : ۴ ،۹ ،۸ ،۲۶ ،۳۷۸۹

از طرفی اگر مجموعه مقادیر آن‌ها، همراه با اعشار بوده و شامل «اعداد حقیقی (Real Numbers)» باشد، متغیر را «پیوسته (Continuous)» نام‌گذاری می‌کنند. مانند معدل دانش آموز ۱۷/۵۷

از جنبه دیگر نقش هر متغیر در تحلیل و ایجاد مدل آماری نیز تعیین کننده ا ست. برای مثال در مدل های رگرسیونی، یک متغیر نقش پیش‌گو یا «مستقل (Independent)» و متغیر دیگر به عنوان متغیر «وابسته (Dependent)» نامیده می‌شوند.

مثال – تشویق کلامی کودک موجب می شود که آن رفتار بصورت عادت در آید

تشویق کلامی = متغیر مستقل

رفتار عادی = متغیر وابسته

➢ ارایه های ساده

فرض کنیم در برابر شما تعدادی بیمه نامه وسایل نقلیه باشد هر یک بیان کننده تعدادی از وسایل نقلیه بیمه شده توسط بیمه گذار است. شما به سرعت این اعداد را بر روی کاغذ می نویسید بدیهی است که جز لیستی از اعداد در مقابل شما چیزمعناداری وجود ندارد . ممکن است که شما این داده ها را به صورت آرایش ساده ارائه دهید که تا حدی وضوح بیشتری به داده ها میدهد . (جدول را ملاحظه نمایید)

تعداد وسایل نقلیه که بوسیله ۳۶ بیمه‌گزار، بیمه شده است

۸	۱۰	۳	۷	۴	۱
۷	۶	۶	۶	۲	۴
۵	۸	۵	۹	۹	۳
۶	۸	۳	۷	۵	۲
۱۱	۱۰	۷	۶	۷	۶
۵	۸	۴	۹	۴	۵

چنان که خواهیم دید هنوز نقیضه وجود دارد و علیرغم اینکه ساختار است مشخص شده لاکن همچنان فاقد وضوح و روشنی است.

➢ آرایه های مرتب

با دقت در جدول بالا آشکار می شود که برخی از اعداد پیش از یک بار آمده است . مثلاً عدد۸ جهار بار تکرار شده. برای آنکه منطقی روشنتر ارائه دهیم بهتر است جدول را به ترتیب صعودی مجدداً مرتب کنیم در آرایش ترتیبی جدید اعداد از کمترین تا بیشترین تعداد (گاهی بلعکس) مرتب شده است.

جدول پایین را ملاحظه نمایید و آن را جدولی با آرایه های مرتب می گوییم

تعداد وسایل نقلیه که بوسیله ۳۶ بیمه‌گزار، بیمه شده است

۱	۴	۵	۶	۷	۹
۲	۴	۵	۶	۷	۹
۲	۴	۵	۶	۸	۹
۳	۴	۶	۷	۸	۱۰
۳	۵	۶	۷	۸	۱۰
۳	۵	۶	۷	۸	۱۱

جدول توزیع فراوانی) :

جدول توزیع فراوانی گروهی نشده:

با توجه به جدول قبلی آنرا تهیه می کنیم . در این جدول داده یا متغیر یا مقادیر به X و فراوانی ها به Y نمایش داده می شود:

تعداد وسایل نقلیه در هربیمه‌نامه x	تعداد بیمه نامه f
۱	۱
۲	۲
۳	۳
۴	۴
۵	۵
۶	۶
۷	۵
۸	۴
۹	۳
۱۰	۲
۱۱	۱
۱۲	۰
جمع	۳۶

❖ توزیعهای فراوانی گروهی:

سازمان دادن داده‌ها یا مشاهدات به صورت طبقات همراه با فراوانی هر طبقه.

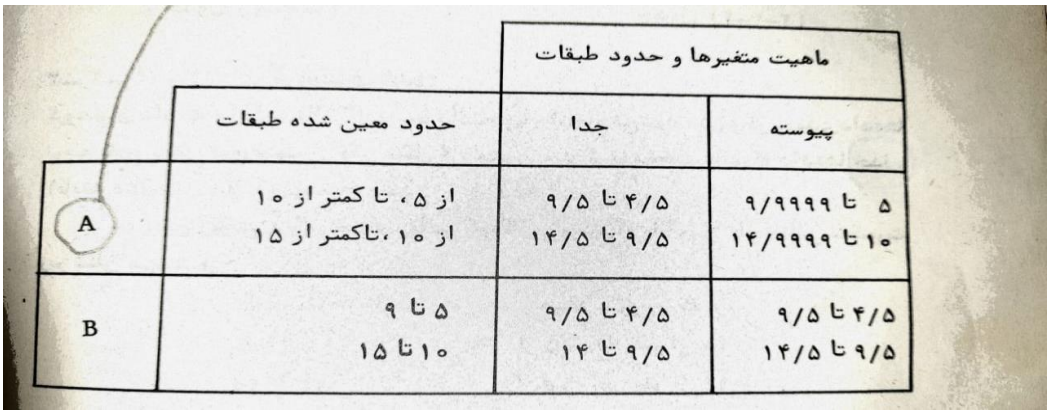
تعداد بیمه‌نامه‌ها	تعداد وسایل‌نقلیه درهربیمه‌نامه
۱۰	۱ تا ۴
۲۰	۵ تا ۸
۶	۹ تا ۱۲
۳۶	جمع

جدول (۲-۴-۲)

تعداد بیمه‌نامه‌ها	تعداد وسائل نقلیه در هربیمه‌نامه
۶	۱ تا ۳
۱۵	۴ تا ۶
۱۲	۷ تا ۹
۳	۱۰ تا ۱۲
۳۶	جمع

جدول (۱-۴-۲)

➢ داده های گروه بندی شده:



	ماهیت متغیرها و حدود طبقات		
	پیوسته	جدا	حدود معین شده طبقات
A	۵ تا ۹/۹۹۹۹ ۱۰ تا ۱۴/۹۹۹۹	۴/۵ تا ۹/۵ ۹/۵ تا ۱۴/۵	از ۵، تا کمتر از ۱۰ از ۱۰، تاکمتر از ۱۵
B	۴/۵ تا ۹/۵ ۹/۵ تا ۱۴/۵	۴/۵ تا ۹/۵ ۹/۵ تا ۱۴	۵ تا ۹ ۱۰ تا ۱۵

❖ مراحل طبقه بندی :

۱. تعداد طبقات

۲. دامنه تغییرات

۳. فاصله طبقات

۱. تعداد طبقات جدول K

محاسبه تعداد طبقات در اختیار محقق قرار دارد؛ اما معمولا برای جداول با داده های زیاد تعداد آن ها بین ۱۰ تا ۲۰ طبقه می باشد. دلیل این امر این می باشد که طبقات کمتر از ۱۰ موجب بزرگ تر شدن اندازه طبقات و در نتیجه از دست رفتن اطلاعات می گردند. همچنین طبقات بالاتر از ۲۰ موجب طولانی شدن تهیه و تنظیم جدول می شود.

اگر K تعداد طبقات و N تعداد کل داده ها باشد K را طوری بدست میآوریم که اولین عد صحیح باشد که به ازای آن رابطه $N <= 2^K$ برقرار باشد. مانند $n = 35$ باشد در اینصورت $k = 6$ خواهد بود بعبارتی دنبال اولین توانی هستیم که اگر عدد ۲ به توان آن عدد برسد از حجم جامعه بزرگتر شود. (اگر تعداد ۳۵ باشد در اینصورت ۲ به توان ۶ اولین عدد صحیح است که رابطه فوق برقرار است).

۲. دامنه تغییرات داده ها R

دامنه تغییرات را در جدول توزیع فراوانی می توان با استفاده از تفاضل بزرگ ترین عدد از کوچک ترین عدد به اضافه یک به دست آورد. بعبارتی فاصله کمترین داده از بیشترین داده در یک جامعه ی آماری، دامنه جامعه آماری نامیده شود و آن را با R نشان می دهند $R=X_{max}-X_{min}$

مثال : برای مشاهدات ۵،۸،۴،۱۲،۸،۲۱،۸ دامنه تغییرات بشرح زیر است:

$X_{max} = 21$ و $X_{min} = ۴$ **در نتیجه** $۲۱ - ۴ = ۱۷$

لازم به ذکر است که در صورت اعشاری بودن دامنه تغییرات عدد مورد نظر را باید گرد کرد.

www.asia26090.ir

۳. **فاصله طبقات C :**

۴. **فاصله طبقات را می توان با استفاده از فرمول زیر به دست آورد:**

تعداد طبقات K / دامنه تغییرات R = فاصله طبقات C

بعد از به دست آوردن K تعداد طبقات و R دامنه تغییرات و C فاصله طبقات ، طبقه بندی را شروع می کنیم. به طوری که از کوچکترین داده شروع کرده و به آن فاصله طبقات را اضافه کرده و این دو عدد را کران پایین و کران بالا طبقه اول می نامیم و سپس به این دو عدد، فاصله طبقه یعنی C را اضافه میکنیم تا کرانه های بعدی به دست آید . به عبارت دیگر اگر بخواهیم حدود طبقات را به دست آوریم از کوچکترین عدد یک واحد کم کرده و آن را حد پایین طبقه اول فرض می نماییم حال بین عدد فاصله طبقه را اضافه می کنیم تا حد پایین طبقه دوم به دست آید. از این عدد یک واحد دلخواه (با توجه به دقت اعداد اصلی یعنی اولین رقم سمت راست مربوط به فاصله طبقه) کم کرده و این عدد حاصل را حد بالای طبقه اول قرار می دهیم. بنابراین به این دو عدد فاصله طبقه یعنیC را اضافه کرده تا ستون حدود طبقات کاملگردد . در هر بررسی آماری، به یکی از دو ستون کران طبقات یا حدود طبقات را نیاز داریم .

مثال: در یک اداره بیمه از ۱۱ پرونده خسارت آتش سوزی اطلاعات خسارت پرداختی بشرح زیر به دست آمده است. برای این داده ها تعداد دسته ، دامنه تغییرات، فاصله طبقات، و ستون کران طبقه و تعداد خط نشان های هر طبقه را حساب کنید ۲۰ ،۱۱، ۱۸، ۱۳، ۸، ۱۶ ،۱۱، ۱۷، ۱۹،۱۰، ۱۸

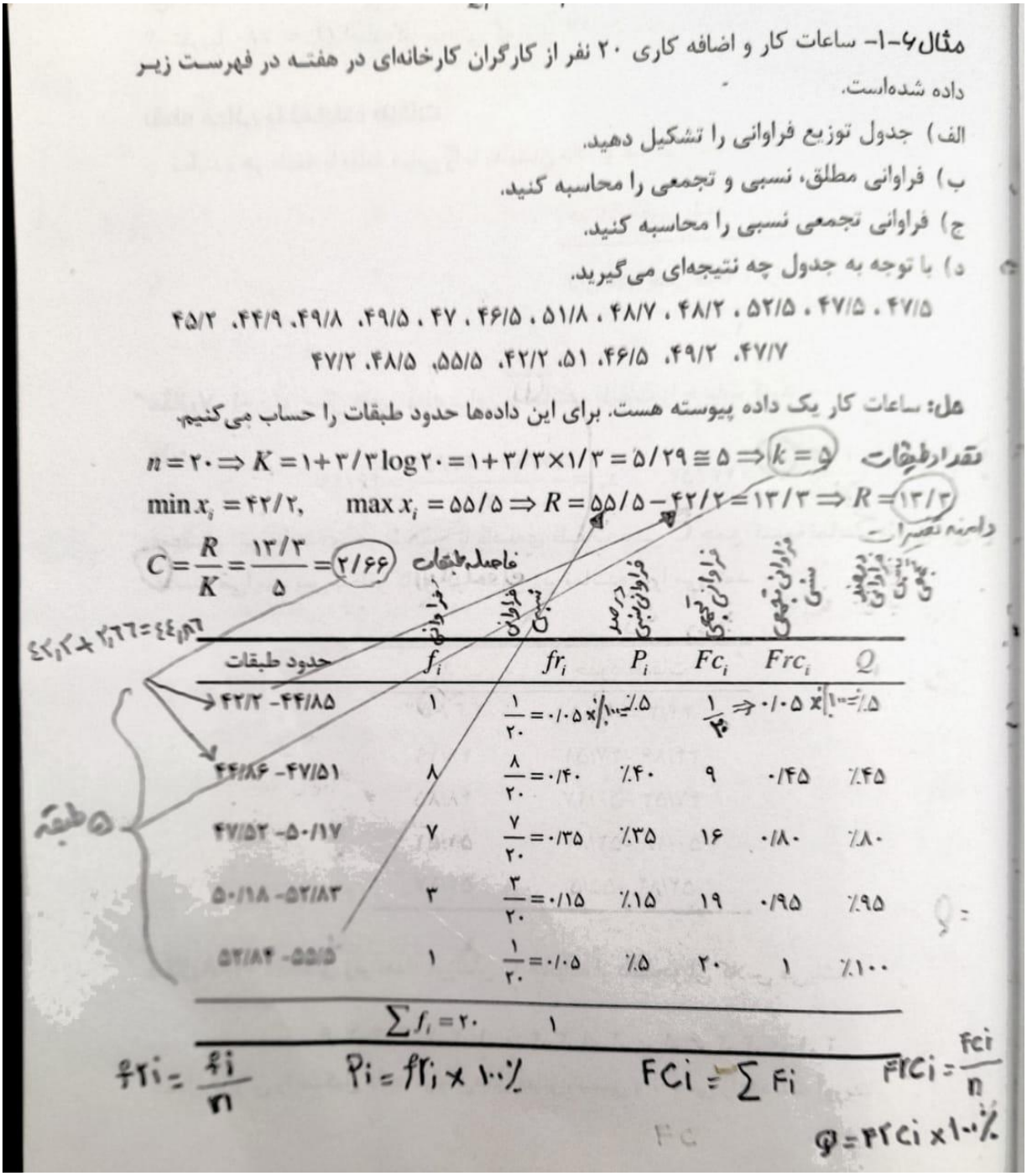
حل:

$$n=۱۱\ ,\ n\le ۲^k \Rightarrow n=۱۱<۲^۴ \Rightarrow k=۴$$

$$\min x_i = ۸\ ,\ \max x_i = ۲۰ \Rightarrow R=\max x_i - \min x_i = ۲۰-۸=۱۲$$

$$C=\frac{R}{K}=\frac{۱۲}{۴}=۳$$

کران طبقات	خط نشان
۸-۱۱	II
۱۱-۱۴	III
۱۴-۱۷	I
۱۷-۲۰	IIII



مثال۶-۱- ساعات کار و اضافه کاری ۲۰ نفر از کارگران کارخانه‌ای در هفته در فهرست زیر داده شده‌است.

الف) جدول توزیع فراوانی را تشکیل دهید.

ب) فراوانی مطلق، نسبی و تجمعی را محاسبه کنید.

ج) فراوانی تجمعی نسبی را محاسبه کنید.

د) با توجه به جدول چه نتیجه‌ای می‌گیرید.

۴۷/۵ ، ۴۷/۵ ، ۵۲/۵ ، ۴۸/۲ ، ۴۸/۷ ، ۵۱/۸ ، ۴۶/۵ ، ۴۷ ، ۴۹/۵ ، ۴۹/۸ ، ۴۴/۹ ، ۴۵/۲

۴۷/۷ ، ۴۹/۲ ، ۴۶/۵ ، ۵۱ ، ۴۲/۲ ، ۵۵/۵ ، ۴۸/۵ ، ۴۷/۲

حل: ساعات کار یک داده پیوسته هست. برای این داده‌ها حدود طبقات را حساب می‌کنیم.

$$n=۲۰ \Rightarrow K=۱+۳/۳\log ۲۰=۱+۳/۳\times ۱/۳=۵/۲۹\cong ۵ \Rightarrow k=۵$$

تعداد طبقات

$$\min x_i=۴۲/۲,\quad \max x_i=۵۵/۵ \Rightarrow R=۵۵/۵-۴۲/۲=۱۳/۳ \Rightarrow R=۱۳/۳$$

دامنه تغییرات

$$C=\frac{R}{K}=\frac{۱۳/۳}{۵}=۲/۶۶$$

فاصله طبقات

حدود طبقات	f_i	fr_i	P_i	Fc_i	Frc_i	Q
۴۲/۲ -۴۴/۸۵	۱	$\frac{۱}{۲۰}=۰/۰۵$	٪۵	$\frac{۱}{۲۰}$	۰/۰۵	٪۵
۴۴/۸۶ -۴۷/۵۱	۸	$\frac{۸}{۲۰}=۰/۴۰$	٪۴۰	۹	۰/۴۵	٪۴۵
۴۷/۵۲ -۵۰/۱۷	۷	$\frac{۷}{۲۰}=۰/۳۵$	٪۳۵	۱۶	۰/۸۰	٪۸۰
۵۰/۱۸ -۵۲/۸۳	۳	$\frac{۳}{۲۰}=۰/۱۵$	٪۱۵	۱۹	۰/۹۵	٪۹۵
۵۲/۸۴ -۵۵/۵	۱	$\frac{۱}{۲۰}=۰/۰۵$	٪۵	۲۰	۱	٪۱۰۰
	$\sum f_i=۲۰$	۱				

۵ طبقه

$fri=\frac{fi}{n}$ $\quad Pi=fri\times ۱۰۰٪$ $\quad FCi=\sum Fi$ $\quad FrCi=\frac{Fci}{n}$ $\quad Q=FrCi\times ۱۰۰٪$

➢ مرکز دسته

هر دسته به صورت یه بازه به شکل [a,b)نمایش داده می شود که در آن a کران (حد) پایین دسته و b کران(حد) بالای دسته نامیده می شود. برای پیدا کردن مرکز دسته [a,b)داریم:

$$x_i=\frac{a+b}{2}$$

مثال۷-۱- برای مثال فوق، نقطه میانی یا نماینده طبقات را حساب کنید.

حل:

$$x_1 = \frac{۴۲/۲ + ۴۴/۸۶}{۲} = ۴۳/۵۳, \qquad x_2 = \frac{۴۴/۸۶ + ۴۷/۵۲}{۲} = ۴۶/۱۹$$

همچنین اگر نماینده‌ی هر طبقه را با فاصله‌ی طبقات یعنی C جمع کنیم، نماینده‌های بعدی به‌دست می‌آیند. پس به‌طور کامل در جدول زیر نماینده‌ها را می‌بینید.

x_i	حدود طبقات
۴۳/۵۳	۴۲/۲ -۴۴/۸۵
۴۶/۱۹	۴۴/۸۶ -۴۷/۵۱
۴۸/۸۵	۴۷/۵۲ -۵۰/۱۷
۵۱/۵۱	۵۰/۱۸ -۵۲/۸۳
۵۴/۱۷	۵۲/۸۴ -۵۵/۵

www.asia2005.ir

➢ فراوانی یک دسته در داده های دسته بندی شده

تعداد داده هایی که در هر دسته قرار می گیرند را فراوانی مطلق آن دسته می گویند.

➢ فراوانی نسبی

فراوانی نسبی دستهi ام در داده های دسته بندی شده برابر است با:

$$F_i = \frac{f_i}{\sum_{i=1}^{n} f_i}$$

➢ فراوانی تجمعی

اگر فراوانی مربوط به هر دسته را با مجموع فراوانی های قبل از آن دسته جمع کنیم، فراوانی تجمعی مربوط به آن

دسته به دست می آید:

$$f_{c_i} = f_1 + f_2 + \ldots + f_i$$

➢ فراوانی تجمعی نسبی برابر است با:

$$F_{c_i} = \frac{f_{c_i}}{\sum_{i=1}^{n} f_i}$$

نکته :اختلاف دو فراوانی تجمعی نسبی برابر است با فراوانی مطلق اندیس بالاتر

$$F_{ci+1} - F_{ci} = f_{ci+1}$$

مثال: توزیع زیررا در نظر می گیریم، درصد فراوانی نسبی متناظر با ۵ xi = کدام است؟

x_i	۱	۲	۳	٤	۵	٦
f_i	٤	٦	۸	۷	۳	۲

پاسخ: طبق فرمول فراوانی نسبی داریم:

$$\text{درصد فراوانی نسبی} = \frac{۳}{۴+۶+۸+۷+۳+۲} \times ۱۰۰ = ۱۰٪$$

متغیر x_i	فراوانی داده f_i	فراوانی نسبی fr_i	درصد فراوانی نسبی P_i	فراوانی تجمعی Fc_i
۰	۳	$\frac{۳}{۲۰} = ۰/۱۵$	٪۱۵	۳
۱	۳	$\frac{۳}{۲۰} = ۰/۱۵$	٪۱۵	۶
۲	۴	$\frac{۴}{۲۰} = ۰/۲$	٪۲۰	۱۰
۳	۳	$\frac{۳}{۲۰} = ۰/۱۵$	٪۱۵	۱۳
۴	۴	$\frac{۴}{۲۰} = ۰/۲$	٪۲۰	۱۷
۵	۱	$\frac{۱}{۲۰} = ۰/۰۵$	٪۵	۱۸
۶	۲	$\frac{۲}{۲۰} = ۰/۱$	٪۱۰	۲۰
		n=۲۰ ← تعداد کل متغیرها یا مجموع f_i ها		

$fr_i = \frac{f_i}{n}$

$P_i = fr_i \times ۱۰۰$ ٪

$F_{ci} = \Sigma f_i$

➢ نمودارهای آماری

نمایش داده‌ها (Data Visualization) «روشی است که به کمک آن می‌توان به شکل سریع و موثر اطلاعات و داده‌ها را برای تصمیم‌گیری مدیران ارائه کرد. روش‌های رسم نمودار یا نمایش داده بیشتر به صورت ترسیم آن‌ها یا تشکیل جداول اطلاعاتی است .

طور کلی، داده‌ها را با توجه به نوع مقدارهای ثبتی یا جمع‌آوری شده، طبقه‌بندی می‌کنند. یکی از اصلی‌ترین طبقه‌بندی‌ها برای داده‌ها، «کیفی **(Qualitative)** »یا «کمی **(Quantitative)** «بودن آن‌ها است.

داده‌های کمی

اگر برای ثبت داده‌ها، از ابزار اندازه‌گیری یا شمارش استفاده شود، می‌گوییم آن داده‌ها کمی هستند. داده‌های کمی به راحتی توسط اعداد قابل ثبت‌اند. برای مثال اگر موضوع مورد بررسی شما اطلاعات مربوط به وزن افراد باشد، با استفاده از ابزار ترازو می‌توانید اطلاعات را ثبت و به شکل عددی نمایش دهید. همچنین تعداد افراد یک خانوار از طریق شمارش قابل محاسبه و ثبت است. مثال‌های دیگر برای داده‌های کمی می‌تواند داده‌های مربوط به دما، میزان مصرف سوخت، درصد رضایت شغلی و فشار خون باشد.

از آنجایی که ماهیت این نوع داده‌ها عددی است، قابلیت انجام محاسبات ریاضی روی آن‌ها وجود دارد. به همین علت از این گونه داده‌ها بیشتر برای محاسبه شاخص‌های آماری نظیر میانگین و انحراف معیار استفاده می‌شود.

داده‌های کیفی

برعکس داده‌های کمی که ماهیتی عددی داشتند، داده‌های کیفی نمایانگر کیفیت هستند و به راحتی نمی‌توان آن‌ها را به صورت عدد نمایش داد. اغلب در تعیین مقدار برای داده‌های کیفی اختلاف نظر یا اعمال سلیقه وجود دارد.

برای مثال داده‌هایی که برای نشان دادن میزان زبری یا نرمی یک پارچه تولید می‌شود به میزان حس یا سلیقه فرد ثبت کننده داده‌ها بستگی دارد. همچنین گروه خونی، رنگ پوست یا مدرک تحصیلی مثال‌هایی از داده‌های کیفی هستند.

البته کدگذاری یا رتبه‌بندی از روش‌هایی هستند که امکان می‌دهند، داده‌های کیفی به صورت عددی نمایش داده شوند. ولی باید توجه داشت که مقدار عددی برای این نوع داده‌ها، قابلیت انجام محاسبات ریاضی را نمی‌دهد و مثلا نمی‌توان مقدار داده‌های کیفی که توسط اعداد کدگذاری شده‌اند را با یکدیگر جمع یا تقسیم کرد.

برای مثال می‌توان مدرک تحصیلی را طبق جدول زیر با اعداد ۱ تا ۷ کدگذاری کرد.

نام مدرک تحصیلی	زیر دیپلم	دیپلم	کاردانی	کارشناسی	کارشناسی ارشد	دکتری	پسا دکتری
کد مدرک تحصیلی	۱	۲	۳	۴	۵	۶	۷

این کدها به منظور تفکیک مدرک‌های تحصیلی ایجاد شده‌اند و نمی‌توان انتظار داشت که سطح علمی دو نفر با مدرک تحصیلی دیپلم (۲+۲) برابر با کارشناسی (۴) باشد.

از داده‌های کیفی به منظور طبقه‌بندی یا مرتب‌سازی جامعه آماری استفاده می‌شود. برای مثال می‌توان افرادی که دارای گروه خونی **AB** هستند را در یک گروه قرار داد یا میانگین فشار خون برای این افراد را با اشخاصی که دارای گروه خونی **B** هستند مقایسه کرد. البته اگر اعضای جامعه براساس گروه‌های خونی طبقه‌بندی شوند، می‌توان رابطه‌ی بین میزان فشار خون و نوع گروه خونی را پیدا کرد.

ترسیم نمودار

یکی از نقاط ضعف نمایش داده‌ها به صورت جدول فراوانی عدم درک سریع اطلاعات جدول است. نمودارها ابزار مناسبی برای نمایش تصویری اطلاعات هستند. انواع مختلفی از نمودار وجود دارد که از جمله می‌توان به نمودار هیستوگرام، نمودار ستونی، نمودار چند ضلعی تراکمی، نمودار دایره‌ای، نمودار سریهای زمانی و ...اشاره کرد.

در آ مار تو صیفی هدف این ست که داده ها را سازماندهی و خلا صه سازی کن ند تا بتوان ند به شکل موثری آنها را ن مایش دهـ ند تا برای مخاط بان به سادگی قا بل درک با شد. یک ضرب الم ثل م عروف انگلی سی می گو ید: " یک ت صویر به ا ندازهٔ هزار کل مه می ارزد." هن گامی که صحبت از ارا ئه اطلا عات آ ماری در مجمو عه ای از داده ها می شود، ا ین ضرب الم ثل کاملاً صحیح ا ست در آ مار تو صیفی برای ن مایش داده ها به شکلی موثر از نمودار های مخت لف ا ستفاده می شود. چ گونگی ن مایش و سازماندهی داده ها در ا ین نمودار ها ن یز جزء مبا حث آ مار تو صیفی می با شد. ا گر مایل ید که در ارت باط با سایر مبا حث مرتبط با روانشنا سی بی شتر بداند ید، مجمو عه آ موزش روانشنا سی (Psychology) در فرادرس را به عنوان یک منبع آموزشی جامع به شما توصیه می کنم.

فرض کنید در کلاسی ۳۰ دانش آموز در یک آزمون آمار، نمرات زیر را کسب کرده اند.

86	80	25	77	73	76	100	90	69	93
90	83	70	73	73	70	90	83	71	95
40	58	68	69	100	78	87	97	92	74

آیا از روی داده های این جدول می توانید بگویید عملکرد کلاس در این آزمون آمار چگونه بوده است؟ مطمئناً با یک نگاه اجمالی، این مجموعه نمرات هیچ پاسخ روشنی را به شما نمی دهد. برای اینکه به پاسخ واضحی برسید، باید این مجموعهٔ داده ها را سازماندهی مجدد و بازنویسی کنید تا اطلاعات مرتبط با یکدیگر بهتر قابل مشاهده باشند. یک روش برای انجام این کار، ساختن نمودار ساقه و برگ می باشد. شکل زیر یک نمودار ساقه و برگ را نشان می دهد.

2	5
3	
4	0
5	8
6	9 8 9
7	7 3 6 0 3 3 0 1 8 4
8	6 0 3 3 7
9	0 3 0 0 5 7 2
10	0 0

برای تشکیل نمودار ساقه و برگ این گونه عمل می کنیم. اعدادی که در جایگاه دهگان قرار دارند، یعنی اعداد ۲ تا ۹، و البته عدد ۱۰ که به آنها اضافه می شود، ساقه ها (stems) هستند. این اعداد را به ترتیب عددی از بالا به پایین و البته در سمت چپ خط عمودی، سازماندهی می کنیم. عدد موجود در جایگاه یکان در هر اندازه گیری، یک برگ (leaf) می باشد و در یک ردیف در سمت راست ساقهٔ متناظرش قرار می گیرد.

به عنوان مثال در مورد ساقهٔ ۶، برگ های ۹، ۸ و ۹ در ردیف مقابل آن قرار می گیرند. در واقع در این ردیف، اعداد ۶۹، ۶۸ و ۶۹ به این شکل ساقه و برگ سازماندهی شده اند. سایر ساقه ها و برگ ها نیز از همین قانون تبعیت می کنند. سعی کنید خودتان برای تمامی این نمرات، نمودار ساقه و برگ را ترسیم کنید و سپس با تصویر بالا مقایسه کنید. بدین ترتیب درک شما از مطلب تضمین می شود. در شکل زیر این نمودار ساقه و برگ را مرتب سازی کرده ایم.

2	5
3	
4	0
5	8
6	8 9 9
7	0 0 1 3 3 3 4 6 7 8
8	0 3 3 6 7
9	0 0 0 2 3 5 7
10	0 0

همانطور که می بینید، مرتب سازی نمودار ساقه و برگ درک آن را ساده تر کرده است. هم اکنون با نگاهی به جدول می توانیم متوجه شویم که سه دانش آموز نمرات زیر ۶۰ کسب کره اند. همچنین از روی نمودار ساقه و برگ مرتب شده می توانیم متوجه شویم که دو نفر از دانش آموزان نمرهٔ کامل یعنی نمرهٔ ۱۰۰ را کسب کرده اند.

همچنین اگر دقت کنیم در خواهیم یافت که اکثر دانش آموزان نمراتشان در دهه های ۷۰، ۸۰ و ۹۰ می باشد. در این مثال، نمرات دارای ساقه طبیعی (جایگاه دهگان) و برگ طبیعی (جایگاه یکان) هستند. با این حال می توانیم این

نمودار را با تقسیم هر کدام از این جایگاه های دهگان به دو قسمت مجزا، بسط دهیم. به عنوان مثال تمامی امتیازات موجود در بخش ۸۰ را می توانیم به شکل زیر به دو ساقهٔ جداگانه بشکنیم

8	0 3 3
8	6 7

تعاریف ساقه و برگ در عمل انعطاف پذیر ا ست. هدف کلی نمودار ساقه و برگ اینست که یک نمایش سریع از چگونگی توزیع داده ها در امتداد بازهٔ مقادیر آنها داشته باشیم. برای بدست آوردن نموداری که به بهترین نحو این هدف را برآورده سازد، ممکن است گاهی لازم شود برخی ابتکارها را به کار برید. توجه داشته باشید که تمامی داده های اصلی را می توان از نمودار ساقه و برگ باز یا بی کرد.

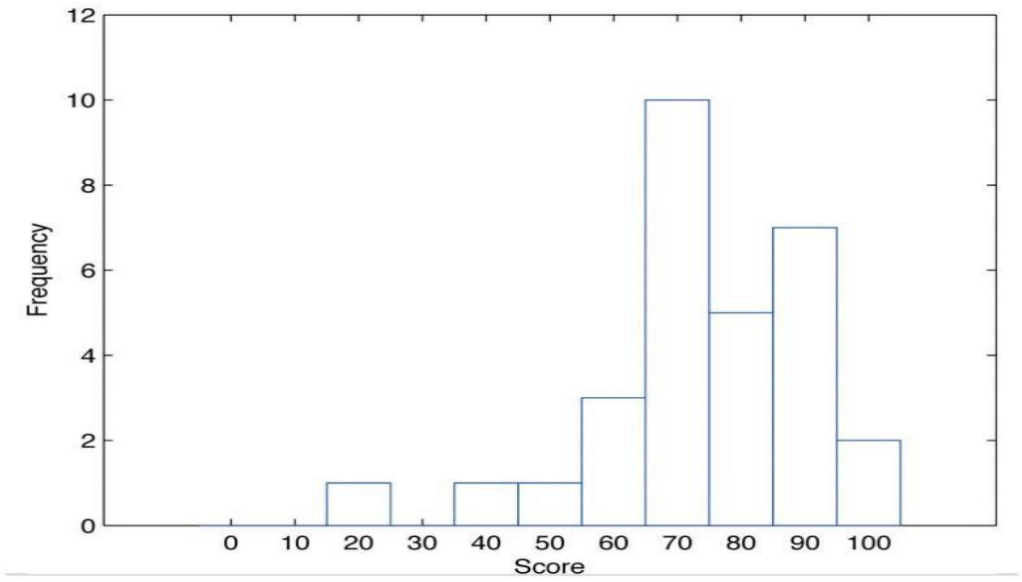
اما در انواع دیگری از نمودارها که در آمار توصیفی مورد استفاده قرار می گیرند، همچون نمودار هیستوگرام فراوانی نسبی، این کار را نمی توان انجام داد و در واقع داده ها به صورت تجمیعی خواهند بود.

هیستوگرام فرکانس (فراوانی) در آمار توصیفی

در این بخش از آموزش آمار توصیفی به نمودار های هیستوگرام فراوانی (Frequency Histograms) می پردازیم. هر چند در بخش پیشین آموزش مشاهده کردیم که نمودار ساقه و برگ در درک داده های نمرات آزمون به ما خیلی کمک کرد. اما در مجموعه داده های بزرگ این نمودار زیاد کاربردی نیست. طبیعتاً در دیتا ست های بزرگ به روش بهتری برای نمایش گرافیکی داده ها نیاز داریم.

ابزاری که به آن نیاز داریم هیستوگرام فراوانی نام دارد. برای این که بهتر بتوانید مقایسه ای بین نمودار ساقه و برگ و نمودار هیستوگرام فراوانی داشته باشید، از همان مجموعه داده هایی که در بخش قبلی استفاده کردیم، مجدداً بهره می بریم. این بار سازماندهی را به این شکل انجام می دهیم که به ازاء هر نمرهٔ موجود از ۰ تا ۱۰۰ که البته به صورت ده تا ده تا جدا شده اند، تعداد نمرات را شمارش می کنیم.

شکل زیر هیستوگرام فراوانی این داده ها می باشند. همانطور که در شکل می بینید از آنجا که دو نفر در این کلاس نمرهٔ ۱۰۰ گرفته اند، در نمودار هیستوگرام فراوانی، ستون مربوط روی ۲ قرار دارد. در این نمودار ۷ نفر نمراتی در دههٔ ۹۰ دارند. شش نفر نمراتشان در دههٔ ۸۰ می باشد و به همین ترتیب تعداد نمرات مربوط به هر دهه در نمودار سازماندهی شده اند. هر چند در این نمودار نمرات تک تک دانش آموزان را دیگر نداریم و داده هایشان را از دست داده ایم، با این حال تعداد نمرات هر دهه را داریم. این عدد فراوانی کلاس نامیده می شود، اساساً برای همین هم هست که نام این نمودار هیستوگرام فراوانی می باشد.

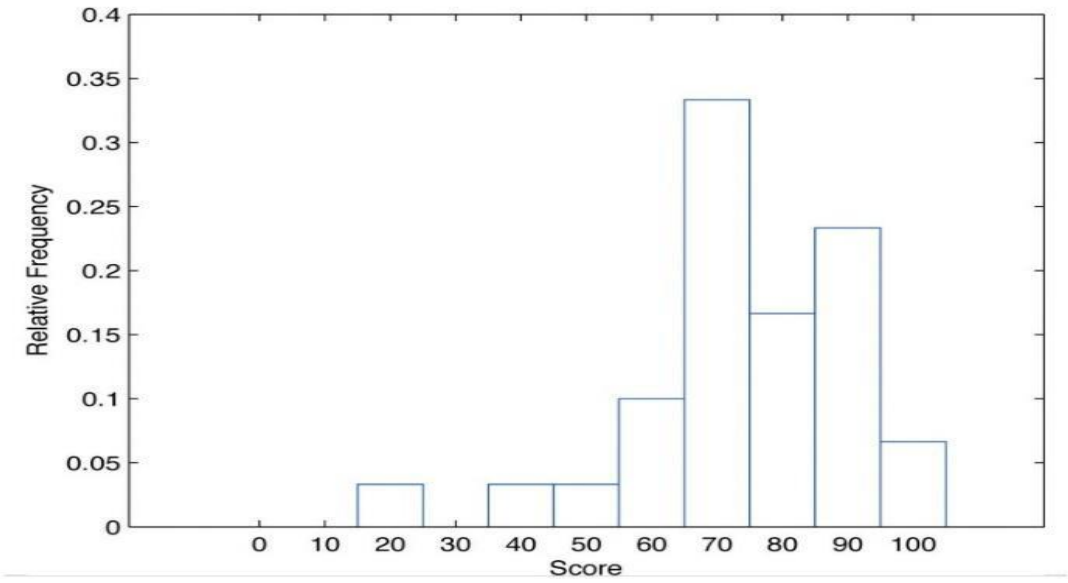


هیستوگرام فراوانی نسبی در آمار توصیفی

در ا ین ب خش از آ موزش آ مار تو صیفی به هی ستوگرام های فراوا نی ن سبی (Relative Frequency Histograms) می پردازیم. در ا ین ب خش، باز هم می خواهیم از داده های مر بوط به ن مرات کلاس در آز مون آ مار ا ستفاده ک نیم. در م ثال بالا م شاهده کرد ید که ن مرات پنج ن فر از دا نش آ موزان در د هۀ ۸۰ قرار داشتند. در واقع عدد ۵ فراوانی گروهی که نام آن دهۀ ۸۰ است، می باشد.

از آ نجا که ت عداد کل دا نش آ موزان ا ین کلاس ۳۰ ن فر می با شد، ن سبت ا فرادی که در ا ین د هه قرار دار ند، ن سبت به کل دا نش آ موزان، می با شد. ک سر را می توانیم به شکل ن یز ن شان داد. ا گر ا ین عدد اع شاری را به در صد ت بدیل ک نیم م قدار را خواهیم دا شت. م قدار فراوا نی ن سبی (relative frequency) این گروه می باشد. هر کدام از گروه های دیگر نیز یک فراوانی نسبی دارند.

از این رو می توانیم یک ن مودار تر سیم ک نیم که در آن برای هر گروه (یا کلاس) یک نوار ع مودی دا شته با شیم که م قدار آن برا بر با فراوا نی ن سبی آن گروه با شد. شکل ز یر ن مودار هی ستوگرام فراوا نی ن سبی داده های ن مرات کلاس را ن شان می د هد. ا ین ن مودار دقی قاً م شابه هی ستوگرام فراوا نی می با شد، با ا ین تفاوت که به جای فراوانی، فراوانی نسبی را نمایش می دهد.



➢ انواع نمودارهای آماری

نمودارهای آماری روشی عینی و سریع برای تجسم اطلاعات نهفته در داده‌ها هستند و ابزاری ساده برای نمایش و استباط از مشاهدات و روشی کارا در تجزیه و تحلیل‌های آماری محسوب می‌شوند و در دو دسته نمودارهای تک‌متغیره و نمودارهای چند متغیره تقسیم می‌شوند. نمودارهای شاخ و برگ دایره ای، میله‌ای و هیستوگرام از رایج‌ترین نمودارهای تک متغیره هستند.

۱ –نمودار شاخه و برگ

برای اینکه یک دید سریع از توزیع داده ها کسب کنیم بهترین روش استفاده از نمودار ساقه و برگ است. در نمودار ساقه و برگ هر مقدار مشاهده شده به دو بخش تقسیم می شود یک یا چند رقم اصلی که شاخه نامیده می شود و یک رقم دنباله که برگ نامیده می شود.

در نمودار ساقه و برگ هر ردیف مطابق با یک ساقه بوده و هر نمونه توسط برگ های آن نشان داده می شود. همچنین بیشتر از یک ردیف می توانند ساقه های یکسانی داشته باشند .این نوع نمودار همچنین روش مناسبی برای نمایش داده‌های پرت و مد ارائه می‌دهد. البته در صورتی که تعداد داده‌ها خیلی زیاد باشد نمودار شکل کاملا بهم ریخته ای خواهد داشت. بنابراین برای تعداد داده بیشتر از ۱۵۰ این نوع نمودار توصیه نمی‌شود. با تعداد کم داده نیز خواص توزیع بخوبی نمایش داده نخواهد شد که در این مورد نمودار نقطه– ای توصیه می شود.

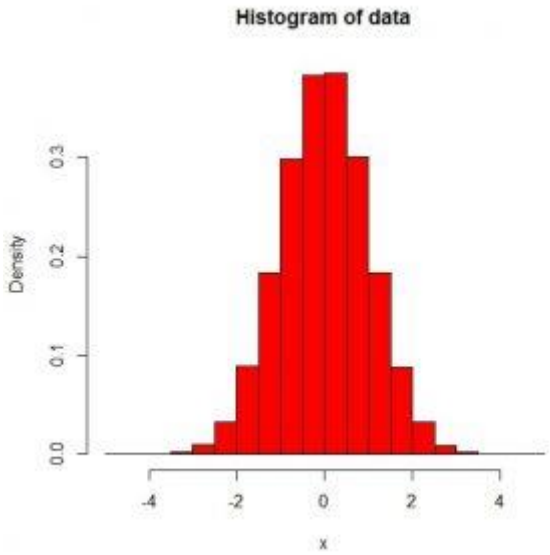
```
 0 | 7 7
 1 | 2 3 8 8
 2 | 3 5 7 7 7 7 7 7 8 8 9 9
 3 | 0 1 1 1 1 2 3 4 5 6 7 7 7 8 9
 4 | 0 0 1 2 2 2 2 3 3 3 4 4 4 5 6 7 8 8
 5 | 2 3 7 8 8
 6 | 1 3 8
 7 | 1
 8 |
 9 | 6 9
10 | 4 5 5 5 5 6 7 9 9 9
11 | 1 3 3 3 3 4 4 4 5 5 5 6 6 6 6 7 7 7 7 8 8 9 9
12 | 0 0 1 1 2 2 3 4 4 4 5 5 5 6 7 7 7 7 8 8
13 | 1 2 3 9
14 | 1 6
```

نمودار شاخه و برگ

www.asia26090.ir

۲ – هیستوگرام

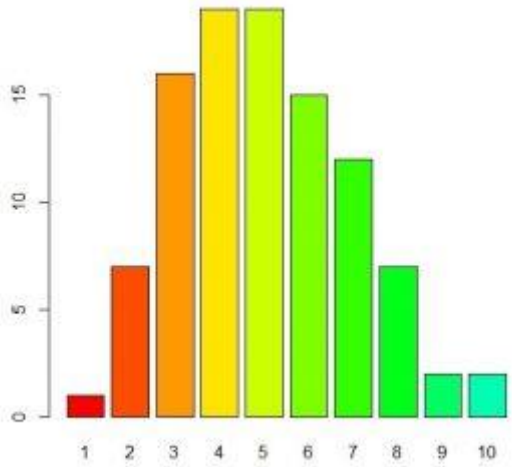
بافت نگار یا هیستوگرام از رایج‌ترین نمودارهای آماری برای داده‌های کمی فاصله‌ای و نسبی است که بهترین روش برای نمایش میزان گرایش به مرکز و پراکندگی در این نوع از داده‌ها و سنجش نرمال بودن توزیع است. در نمودارهیستوگرام ستون افقی شامل متغیرها و ستون عمودی فراوانی یا فراوانی نسبی آنهاست.



هیستوگرام

۳ –نمودار میله ای

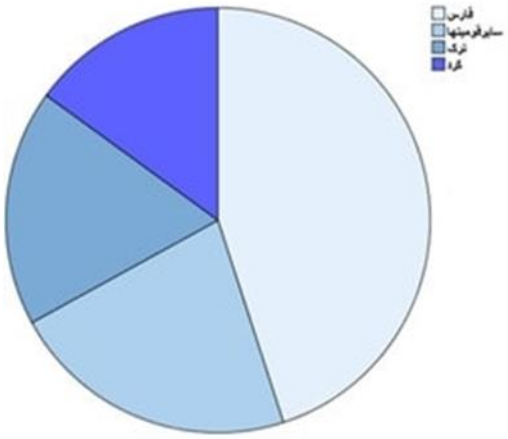
نمودار میله‌ای برای نمایش داده‌های کیفی اسمی و ترتیبی و کمی با تعداد طبقات کم بکار می‌رود شامل مجموعه ای از ستون‌ها یا میله-ها با فاصله‌های یکنواخت است و در آن هر مقدار به عنوان رده یا طبقه‌ای از یک متغیر و فراوانی نسبی آن با ارتفاع میله‌ها نشان داده می شوند.



نمودار میله ای

۴ –نمودار دایره ای

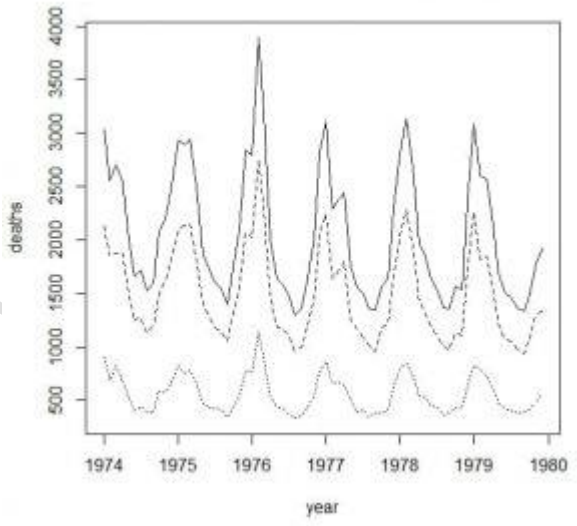
یک روش جالب برای نمایش متغیرهای مرتبط استفاده از نمودار دایره ای است. در این نوع نمایش سهم مشارکت متغیرهای خاص در موضوع مورد بررسی نشان داده می شود که در آن متغیرها به تناسب فراوانی نسبی، سهمی از دایره را پر می کنند. معمولاً بخش های تشکیل دهنده در این نوع نمودار در جهت عقربه های ساعت از بزرگ به کوچک تنظیم می شوند.



نمودار دایره ای

۵ –نمودار سری زمانی

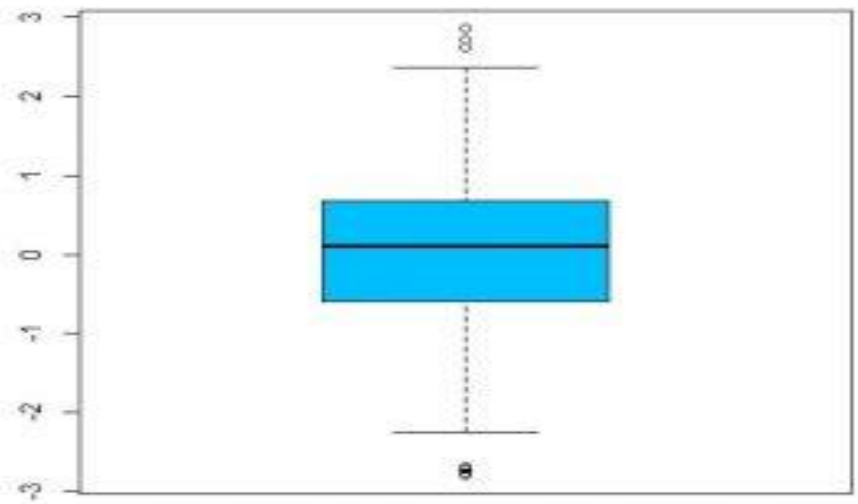
این نوع نمودار برای نمایش گرافیکی پدیده هایی که در طول زمان تغییر کرده و اندازه گیری شده اند، استفاده می شود. بدین ترتیب که مقدار مشاهده شده متغییر محور عمودی و زمان نیز محور افقی قرار می گیرد.



نمودار سری زمانی

۶ –نمودار جعبه ای

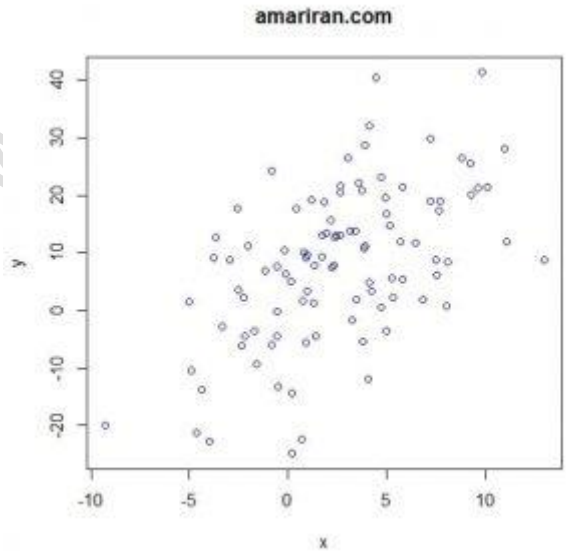
این نمودار روشی مناسب برای نمایش نحوه پراکندگی داده های کمی است که بر اساس داده‌های مرتب شده به صورت صعودی و تعیین چارک‌ها و دامنه است در نمودار جعبه ای از چارک ها برای نشان دادن فاصله استفاده م ی‌شود و خط داخل جعبه چارک دوم یا میانه را مشخص می‌کند . دامنه داده ها نیز با خطوطی خارج از جعبه، حد بالایی و پایینی داده ها در رده‌ای خاص را نشان می دهند. داده های پرت نیز در صورت وجود با نقاطی اطراف جعبه نشان داده می شود.



نمودار جعبه ای

۷ –نمودار پراکنش

رابطه بین دو متغییر با استفاده از نمودار پراکنش بررسی می شود. که در آن مقادیر هر متغیر با نقاطی بر روی صفحه نمودار مشخص می‌شوند .و به این صورت می توان وجود الگوهای مختلف در ارتباط دو متغییر را جستجو کرد. تفسیر نمودار پراکنش آسان بوده و اطلاعات زیادی در مورد داده ها در اختیار می گذارد.



نمودار پراکنش

فصل سوم

شاخصهای اصلی تمایل مرکزی و کاربرد آن در صنعت بیمه

- **دو دسته از روش های آماری که برای توصیف داده ها لازمند نام ببرید : شاخص های مرکزی و شاخص های پراکندگی**

شاخص های گرایش مرکزی:

شاخص های مرکزی به مقدار یا عددی که به بهترین شکل ویژگی گروه را به عنوان یک کل نشان می دهد اشاره می کنند . مترادف شاخص مرکزی نمره ی متوسط یا معدل ((averageاست.بنابراین شاخص های مرکزی یا گرایش به مرکز شاخص هایی هستندکه با استفاده ازآنها مجموعه ای ازداده ها در یک مقدار یا عدد که نماینده ی آن مجموعه است خلاصه می شود . نما ومیانه و میانگین((mean,median,modeسه شاخص مرکزی هستند که درزیر به آنها اشاره می کنیم .

میانگین: میانگین n داده آماری $x1.x2 \dots xn$ برابر است با:

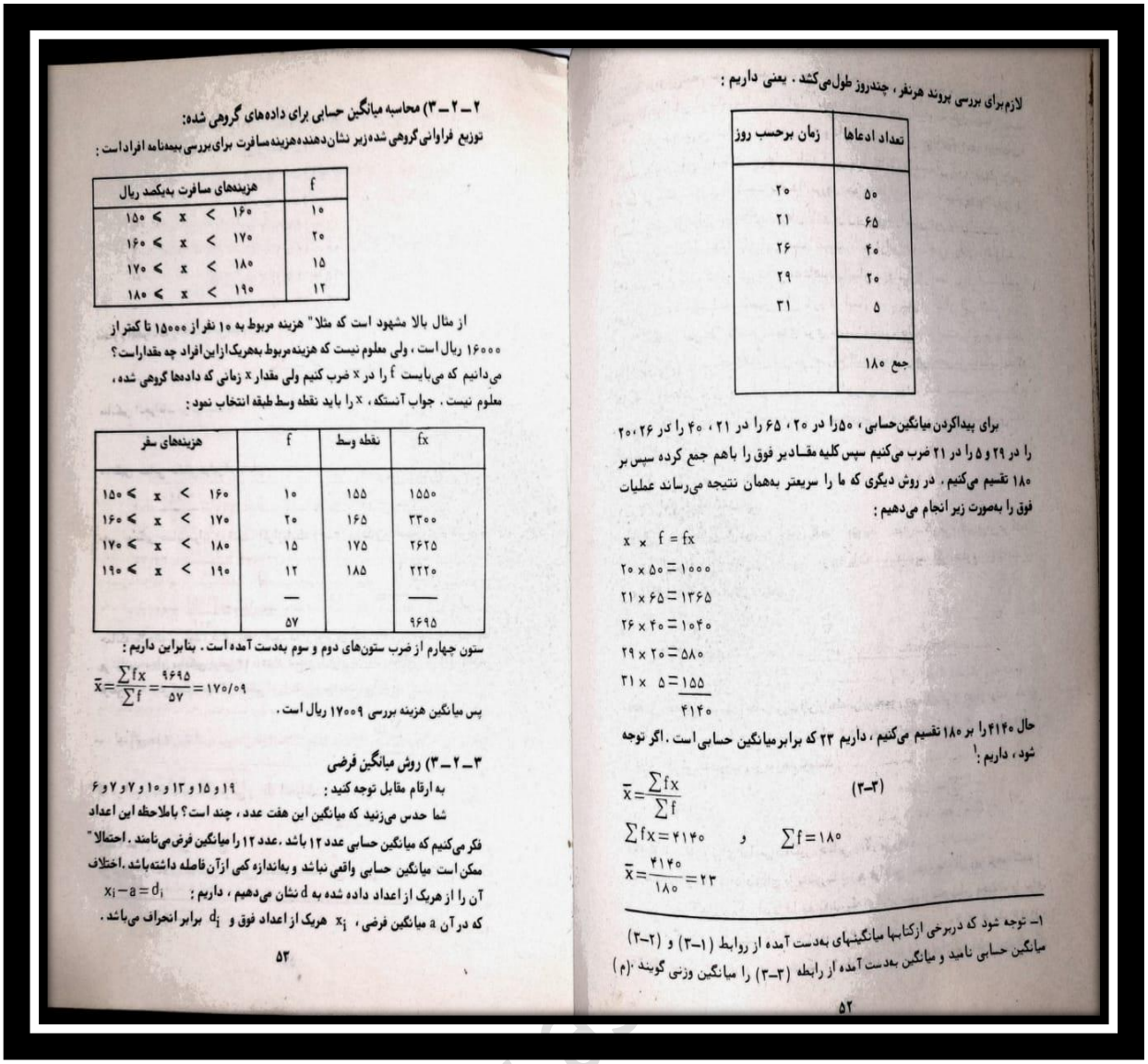
$$\overline{X}=\frac{x_1+x_2+\dots+x_n}{n}=\frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow \sum_{i=1}^{n} x_i = n\overline{X}$$

در یک جامعه آماری، میانگین عددی منحصر به فرد است که همواره بین بزرگ ترین داده و کوچکترین داده قرار دارد.

محاسبه میانگین در جدول فراوانی: (میانگین وزن دار(

میانگین n داده ی آماری که در آن داده ها دارای فراوانی باشند برابر است با

$$\overline{x}=\frac{f_1x_1+f_2x_2+\cdots+f_nx_n}{f_1+f_2+\cdots+f_n}$$



لازم برای بررسی بروند هرنفر، چندروز طول می‌کشد. یعنی داریم :

تعداد ادعاها	زمان برحسب روز
۵۰	۲۰
۶۵	۲۱
۴۰	۲۶
۲۰	۲۹
۵	۳۱
جمع ۱۸۰	

برای پیداکردن میانگین‌حسابی، ۵۰را در ۲۰، ۶۵ را در ۲۱، ۴۰ را در ۲۶، ۲۰ را در ۲۹ و ۵ را در ۳۱ ضرب می‌کنیم سپس کلیه مقادیر فوق را باهم جمع کرده سپس بر ۱۸۰ تقسیم می‌کنیم. در روش دیگری که ما را سریعتر بهمان نتیجه می‌رساند عملیات فوق را به‌صورت زیر انجام می‌دهیم :

$x \times f = fx$

$20 \times 50 = 1000$

$21 \times 65 = 1365$

$26 \times 40 = 1040$

$29 \times 20 = 580$

$31 \times 5 = 155$

4140

حال ۴۱۴۰را بر ۱۸۰ تقسیم می‌کنیم، داریم ۲۳ که برابر میانگین حسابی است. اگر توجه شود، داریم :[1]

$$\bar{x} = \frac{\sum fx}{\sum f} \qquad (۳-۳)$$

$$\sum fx = 4140 \quad \text{و} \quad \sum f = 180$$

$$\bar{x} = \frac{4140}{180} = 23$$

۱- توجه شود که دربرخی ازکتابها میانگینهای به‌دست آمده از روابط (۳-۱) و (۳-۲) میانگین حسابی نامید و میانگین به‌دست آمده از رابطه (۳-۳) را میانگین وزنی گویند. (م)

۵۲

۲ــ۲ــ۳) محاسبه میانگین حسابی برای داده‌های گروهی شده:

توزیع فراوانی گروهی شده‌زیر نشان‌دهنده‌هزینه‌مسافرت برای‌بررسی بیمه‌نامه افراد است :

هزینه‌های مسافرت به‌یکصد ریال	f
۱۵۰ ≤ x < ۱۶۰	۱۰
۱۶۰ ≤ x < ۱۷۰	۲۰
۱۷۰ ≤ x < ۱۸۰	۱۵
۱۸۰ ≤ x < ۱۹۰	۱۲

از مثال بالا مشهود است که مثلا" هزینه مربوط به ۱۰ نفر از ۱۵۰۰۰ تا کمتر از ۱۶۰۰۰ ریال است، ولی معلوم نیست که هزینه‌مربوط به‌هریک‌ازاین‌افراد چه مقداراست؟ می‌دانیم که می‌بایست f را در x ضرب کنیم ولی مقدار x زمانی که داده‌ها گروهی شده، معلوم نیست. جواب آنستکه، x را باید نقطه وسط طبقه انتخاب نمود :

هزینه‌های سفر	f	نقطه وسط	fx
۱۵۰ ≤ x < ۱۶۰	۱۰	۱۵۵	۱۵۵۰
۱۶۰ ≤ x < ۱۷۰	۲۰	۱۶۵	۳۳۰۰
۱۷۰ ≤ x < ۱۸۰	۱۵	۱۷۵	۲۶۲۵
۱۹۰ ≤ x < ۱۹۰	۱۲	۱۸۵	۲۲۲۰
	—		—
	۵۷		۹۶۹۵

ستون چهارم از ضرب ستون‌های دوم و سوم به‌دست آمده است. بنابراین داریم :

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{9695}{57} = 170/09$$

پس میانگین هزینه بررسی ۱۷۰۰۹ ریال است.

۳ــ۲ــ۳) روش میانگین فرضی

به ارقام مقابل توجه کنید : ۱۹ و ۱۵ و ۱۳ و ۱۰ و ۷ و ۷ و ۶

شما حدس می‌زنید که میانگین این هفت عدد، چند است؟ باملاحظه این اعداد فکر می‌کنیم که میانگین حسابی عدد ۱۲ باشد. عدد ۱۲ را میانگین فرض می‌نامند. احتمالا" ممکن است میانگین حسابی واقعی نباشد و به‌اندازه کمی ازآن فاصله داشته‌باشد. اختلاف آن را از هریک از اعداد داده شده به d نشان می‌دهیم، داریم : $x_i - a = d_i$

که در آن a میانگین فرضی، x_i هریک از اعداد فوق و d_i برابر انحراف می‌باشد.

۵۳

نما mode:

نما ساده ترین شاخص گرایش مرکزی است، نما عددی است که دارای بیشترین فراوانی است،یا عددی که بیشتر از اعداد دیگر تکرار شده است.نماازطریق مشاهده ی توزیع فراوانی وتعیین عددی که دارای بیشترین فراوانی است تعیین می گردد.عددی که دارای بیشترین فراوانی است،غالباًنزدیک مرکزتوزیع فراوانی قراردارد.درچنین شرایطی نمایک شاخص مرکزی است.امانماهمیشه درمرکزتوزیع فراوانی قرارندارد.به همین دلیل نمی توان به عنوان یک شاخص مرکزی به آن اطمینان داشت.درواقع نما درمیان شاخص های گرایش به مرکز،شاخصی بی ثبات است.به توزیع فراوانی که فقط یک نمادارد،یک نمایی و به توزیعی که دونمادارد، توزیع دونمایی می گویند وبه توزیع بیش ازدونما، توزیع چند نمایی گفته می شود.

نما یا مُد

داده یا داده‌هایی که دارای بیشترین تکرار باشند مُد یا نما می‌نامیم و با Mo نمایش می‌دهیم و با دو حالت زیر قابل محاسبه است. توجه کنید که داده‌ها ممکن است بیش از یک نما داشته باشند.

۱- اگر داده معین و تعداد آنها محدود باشند، آن داده یا داده‌هایی که بیشتر فراوانی مطلق را دارد، نما یا مُد درنظر می‌گیریم.

۲- اگر داده‌های جامعه آماری همراه کران یا حدود در جدولی طبقه‌بندی شده باشد، طبقه یا طبقاتی که دارای بیشترین فراوانی مطلق‌اند را طبقه نمادار فرض کرده، سپس از رابطه‌ی زیر نما یا مُد را حساب می‌کنیم.

$$Mo = L + \frac{d_1}{d_1 + d_2} \times C$$

L = کران پائین طبقه‌ی نمادار

d_1 = اختلاف فراوانی مطلق طبقه‌ی نمادار با طبقه قبل

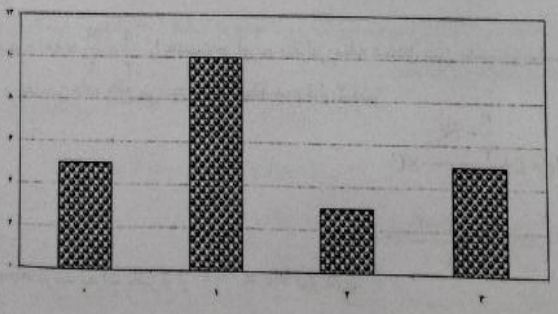
d_2 = اختلاف فراوانی مطلق طبقه‌ی نمادار با طبقه بعد

C = فاصله طبقه

مثال۱۳-۲- تعداد مشروطی دانشجویان کلاسی را سؤال کرده‌ایم، نما را حساب کنید.

تعداد مشروطی (x_i)	۰	۱	۲	۳
تعداد دانشجو (f_i)	۵	۱۰	۳	۵

حل: طبقه دوم بیشترین تکرار را دارد، پس Mo = ۱ خواهد بود.



مثال۱۴-۲- سن دانشجویان کلاسی ۲۵, ۳۰, ۲۵, ۲۷, ۲۸, ۲۰, ۲۵, ۲۰, ۲۱, ۲۰, ۱۹ می‌باشد، نما یا مُد را حساب کنید.

حل: در این داده‌ها دو عدد بیشترین تکرار را دارد پس در این داده‌ها دو مد وجود دارد.

$Mo_1 = ۲۰$, $Mo_2 = ۲۵$

مثال۱۵-۲- برای جدول زیر نما یا مُد را حساب کنید.

نمرات	۸-۱۰	۱۰-۱۲	۱۲-۱۴	۱۴-۱۶	۱۶-۱۸	۱۸-۲۰
تعداد دانشجو	۱۲	۸	۱۰	۱۱	۵	۷

حل: دراین داده‌ها طبقه اول طبقه نمادار است .

$C = ۲$, $d_1 = ۱۲ - ۰ = ۱۲$, $d_2 = ۱۲ - ۸ = ۴$

$$\Rightarrow Mo = L + \frac{d_1}{d_1 + d_2} \times C$$

$$= ۸ + \frac{۱۲}{۱۲ + ۴} \times ۲ = ۸ + \frac{۲۴}{۱۶} = ۹/۵$$

یعنی اکثر نمرات کلاس نزدیک ۹/۵ = Mo قرار دارند.

نکته۴: اگر در جدولی کران پائین طبقه اول و یا کران بالای طبقه آخر معین نباشد، میانگین حسابی را نمی‌توان حساب کرد، و باید با محاسبه میانه و نما این داده را تحلیل کنیم.

مثال ۱۶-۲- برای جدول زیر معیارهای تمرکز را محاسبه کنید.

نمرات	۸-۱۰	۱۰-۱۲	۱۲-۱۴	۱۴-۱۶	بیش از ۱۶
تعداد دانشجو f_i	۵	۸	۷	۸	۲
Fc_i	۵	۱۳	۲۰		

حل: میانگین حسابی را با توجه به نکته فوق نمی‌توان حساب کرد، ولی می‌توانیم میانه و مُد را به‌دست آوریم، این داده دارای دو نما می‌باشد .

$n = \sum f_i = ۳۰$, $\frac{n}{۲} = \frac{۳۰}{۲} = ۱۵$ $\Rightarrow$ داده‌ی پانزدهم میانه است

خلاصه ی ویژگی های نما عبارتند از:

۱. نما یک شاخص اسمی است.

۲. عددی است که دارای بیشترین فراوانی است.

۳. غالباًدروسط توزیع نمره هاقراردارد.

۴. دربعضی ازتوزیع ها بیش ازیکی وجوددارد.

۵. اعمال ریاضی رانمی توان باآن انجام داد.

۶. دارای پایداری یاثبات خیلی کمی است.

۷. برآوردخیلی ضعیفی ازپارامترجامعه است.

۸. به ندرت به کاربرده می شود.

میانه median

میانه نقطه ی وسط یک توزیع فراوانی است.هنگامی که نمرات توزیع فراوانیبه ترتیب ازبالاترین به پایین ترین نمره مرتب شود،نصف نمرات بالای میانه ونصف نمرات پایین میانه قرارمی گیرد.بنابراین،میانه نقطه ۵۰درصداست.یعنی عددی است که توزیع نمره هارابه دوقسمت مساوی تقسیم می کند.

بعبارت دیگر اگر داده های آماری را به صورت صعودی یا نزولی مرتب کنیم، در صورتی که تعداد داده ها فرد باشد، عددی که در وسط قرار می گیرد و اگر تعداد داده ها زوج باشد، نصف مجموع دو عددی که در وسط قرار گرفته، میانه نام دارد. معمولاً میانه مقداری است که تعداد اعضایی از جامعه که از آن بیشترند برابر تعداد اعضایی است که از آن کمترند.

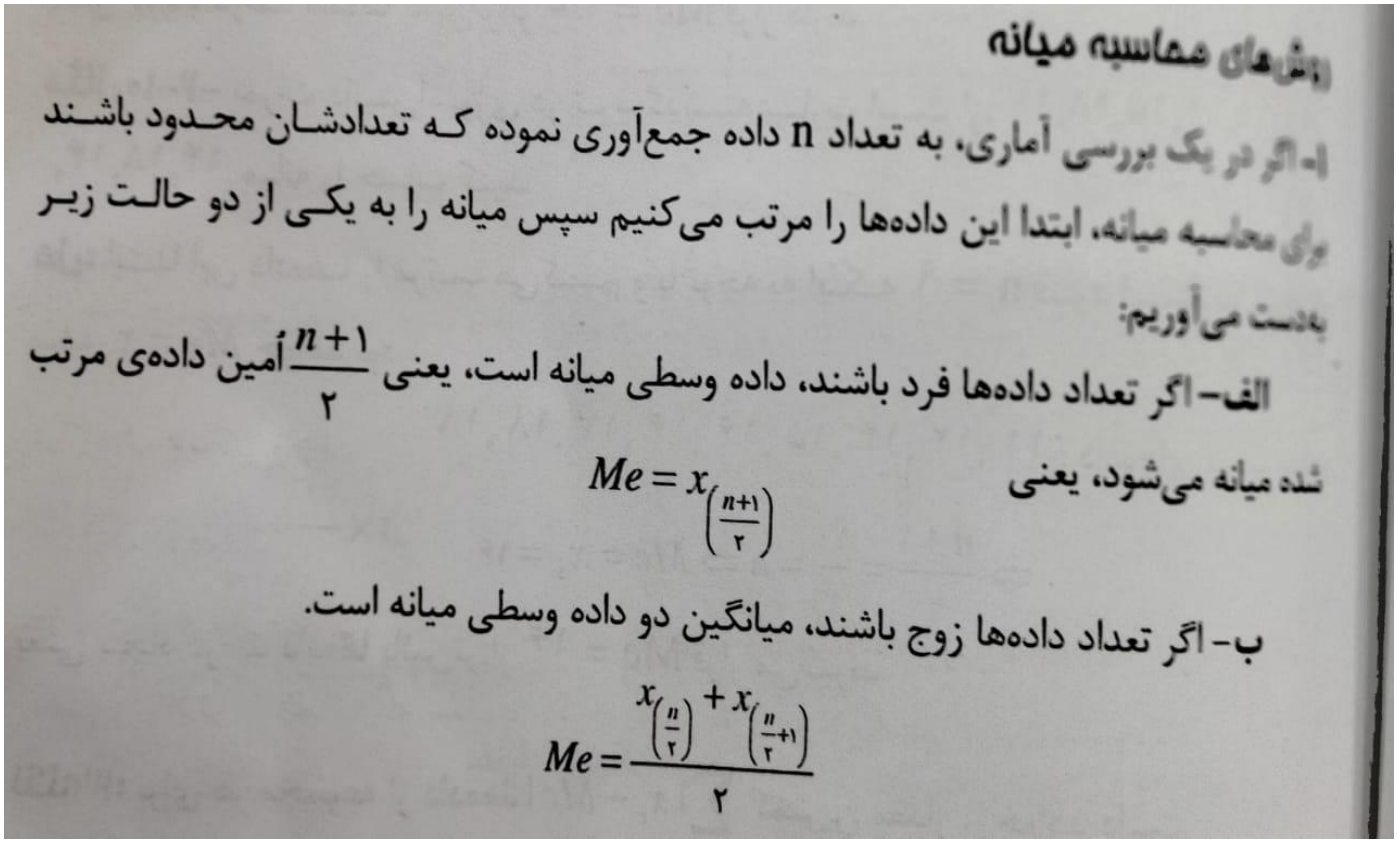
مثال: میانه داده های زیر را بدست آورید:

۴، ۷، ۷، ۳، ۱۲، ۱۱، ۱۷، ۱۵، ۱۴، ۱۷، ۱۹، ۱۴، ۱۰، ۹، 5

پاسخ: ابتدا داده ها را مرتب می کنیم.

3 4 5 7 7 9 10 11 12 14 14 15 17 17 19

چون ۱۵داده داریم، داده ۸ ام میانه است. پس عدد ۱۱ میانه است.



روش‌های محاسبه میانه

۱- اگر در یک بررسی آماری، به تعداد n داده جمع‌آوری نموده که تعدادشان محدود باشند برای محاسبه میانه، ابتدا این داده‌ها را مرتب می‌کنیم سپس میانه را به یکی از دو حالت زیر بدست می‌آوریم:

الف- اگر تعداد داده‌ها فرد باشند، داده وسطی میانه است، یعنی $\frac{n+1}{2}$ اُمین داده‌ی مرتب شده میانه می‌شود، یعنی

$$Me = x_{\left(\frac{n+1}{2}\right)}$$

ب- اگر تعداد داده‌ها زوج باشند، میانگین دو داده وسطی میانه است.

$$Me = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}$$

۲- اگر داده‌ها همراه کران یا حدود در جدولی آمده باشند، ابتدا عدد $\frac{n}{2}$ را با ستون فراوانی تجمعی مقایسه کرده، تا طبقه‌ای که $\frac{n}{2}$ امین عدد در آن طبقه واقع شده است، مشخص گردد. این طبقه را طبقه‌ی میانه‌دار می‌نامیم، سپس میانه را با فرمول زیر به‌دست می‌آوریم.

$$Me = L + \frac{\frac{n}{2} - Fc_{i-1}}{f_i} \times C$$

L = کران پائین طبقه‌ی میانه‌دار

Fc_{i-1} = فراوانی تجمعی طبقه‌ی ماقبل میانه‌دار

f_i = فراوانی مطلق طبقه‌ی میانه‌دار

C = فاصله طبقات

مثال۹-۲- از چند دانشجوی کلاس سؤال کرده‌ایم که تاکنون چند بار مشروط شده‌اید. داده‌های ۲, ۲, ۱, ۲, ۲, ۰, ۰, ۱ به‌دست آمد. میانه را حساب کنید.

حل: ابتدا این داده‌ها را مرتب می‌کنیم. با توجه به اینکه $n = 8$ زوج است پس $\frac{n}{2} = 4$ و میانه به‌صورت زیر خواهد بود.

داده‌های مرتب شده: ۰, ۰, ۱, ۱, ۲, ۲, ۲, ۲ $\Rightarrow$ $Me = \frac{x_4 + x_5}{2} = \frac{1+2}{2} = 1/5$

یعنی پنجاه درصد داده‌ها پائین‌تر از $Me = 1/5$ قرار دارند.

مثال۱۰-۲- نمرات دانش‌آموزی در ترم گذشته‌عبارت است از: ۱۲,۱۱,۱۶,۱۵,۱۸,۱۷, ۱۴,۱۸,۱۶. میانه را حساب کنید.

حل: ابتدا این داده‌ها را مرتب می‌کنیم و با توجه به اینکه $n = 9$ فرد است، در نتیجه میانه $Me = x_5$ خواهد بود.

داده‌های مرتب شده: ۱۱, ۱۲, ۱۴, ۱۵, ۱۶, ۱۶, ۱۷, ۱۸, ۱۸

$$\Rightarrow \frac{n+1}{2} = \frac{10}{2} = 5 \Rightarrow Me = x_5 = 16$$

یعنی پنجاه درصد داده‌ها پائین‌تر از $Me = 16$ قرار می‌گیرند.

نکته۳: برای هر مجموعه از داده‌ها، $\sum |x_i - Me|$ کمترین مقدار را خواهد داشت.

مثال۱۱-۲- برای جدول زیر میانه را حساب کنید.

تعداد فرزندان پسر	۰	۱	۲	۳	۴
تعداد خانوار	۵	۱۰	۱۵	۸	۲

حل: تعداد خانوار را f_i درنظر می‌گیریم، چون $n = \sum f_i = 40$ زوج می‌باشد

$$Me = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} = \frac{x_{20} + x_{21}}{2} = \frac{2+2}{2} = 2$$

یعنی پنجاه درصد خانواده‌ها تعداد فرزندان پسرشان پایین‌تر از $Me = 2$ می‌باشد.

مثال ۱۲-۲- اگر در یک کلاس، از دانشجویان نمرات درس ریاضی (۱) را سؤال کرده و به‌صورت جدول زیر طبقه‌بندی شده باشد، میانه را حساب کنید.

نمرات	f_i	Fc_i
۸-۱۰	۱۲	۱۲
۱۰-۱۲	۸	۲۰
۱۲-۱۴	۱۰	۳۰
۱۴-۱۶	۱۱	
۱۶-۱۸	۵	
۱۸-۲۰	۷	
	n=۵۳	

حل: ابتدا $\frac{n}{2} = \frac{53}{2} = 26/5$ را به‌دست می‌آوریم. بنابراین میانه عددی بین بیست و ششم و بیست و هفتمین داده واقع می‌شود که در طبقه سوم قرار گرفت.

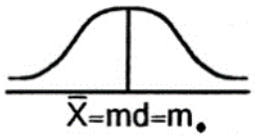
$$Me = L + \frac{\frac{n}{2} - Fc_{i-1}}{f_i} \times C$$

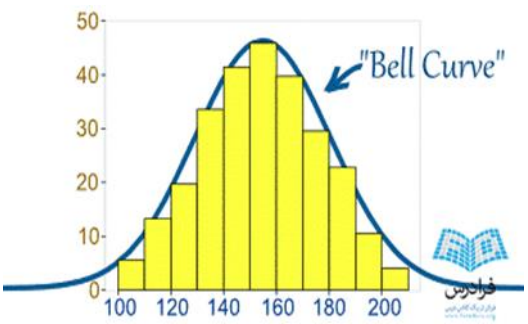
$$= 12 + \frac{26/5 - 20}{10} \times 2 = 13/3$$

یعنی پنجاه درصد نمرات پائین‌تر از $Me = 13/3$ قرار دارند.

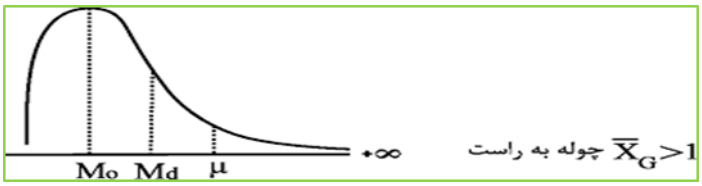
رابطه بین میانگین حسابی ، میانه، نما

اگر X=MD=MO برابر باشد داده ها را متقارن یا نرمال گویند.

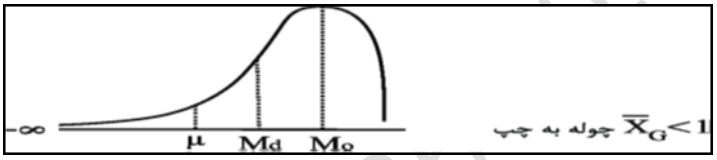


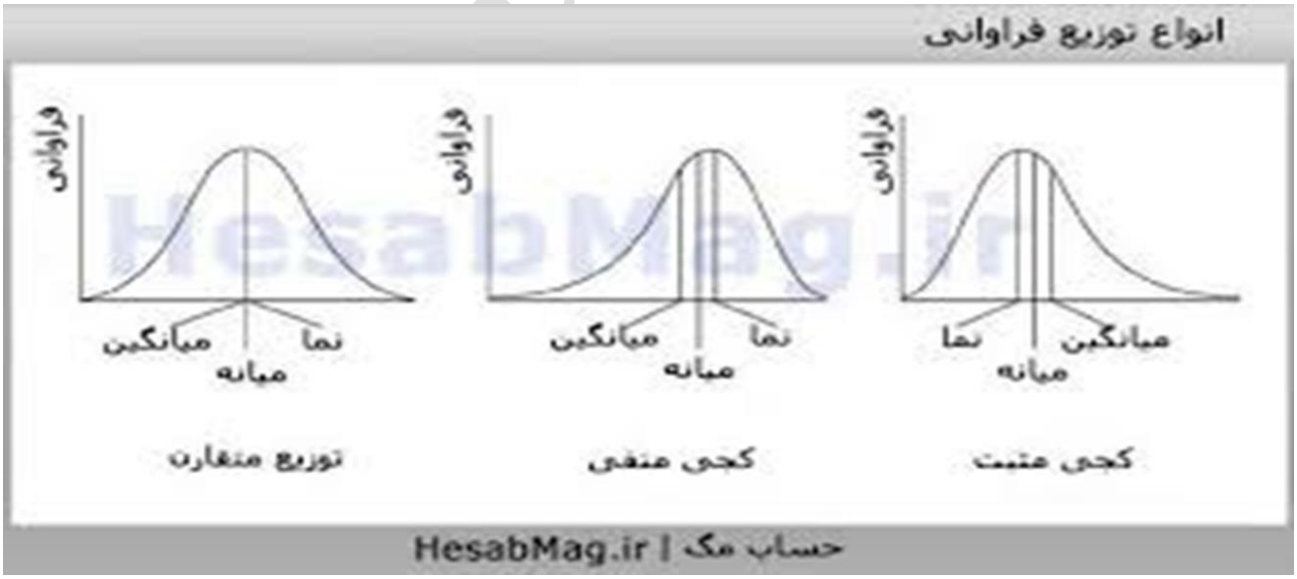


- اگر X>MD>MO باشند داده ها کشیدگی (چولگی) به راست دارند. تعداد خسارات کوچک زیاد و مد است و تعداد خسارات بزرگ(میانگین ها) کم ولی مبالغ بزرگ را تشکیل داده



- اگر X<ME<MO باشند داده ها کشیدگی به چپ دارند. خسارات بزرگ (مد) تعداد بیشتری هستند. و تعداد میانگین خسارات کم است





هرگاه میانگین هندسی کمتر از یک برابر باشد چولگی منفی است. و بالعکس

➢ چارک ها ، دهک ها ، صدکها(درصدها)

میانه ، داده های توزیع را به دوقسمت تقسیم می کند. ولی گاهی ترجیح می دهیم که ۲۵ درصد بالا و یا ده درصد زیر داده ها را داشته باشیم . این مقصود با تقسیم توزیع به صورت های زیر به دست می آید

یک چهارم ها = چهارکها

یک دهم ها = دهک ها

یک صدم ها = صدک ها

چهارک ها ، دهک‌ها و صدک ها به وسیله فرمولی شبیه فرمول ۸-۳ با قرار دادن عدد مناسب به جای ۲ / N در آن به دست می آید این مقادیر به شرح زیر است:

$Q_1 = \frac{N}{4}$ = **برای اولین چارک**

$Q_3 = \frac{3N}{4}$ = **برای سومین چارک**

$D_1 = \frac{N}{10}$ = **برای اولین دهک**

$D_2 = \frac{2N}{10}$ = **برای دومین دهک**

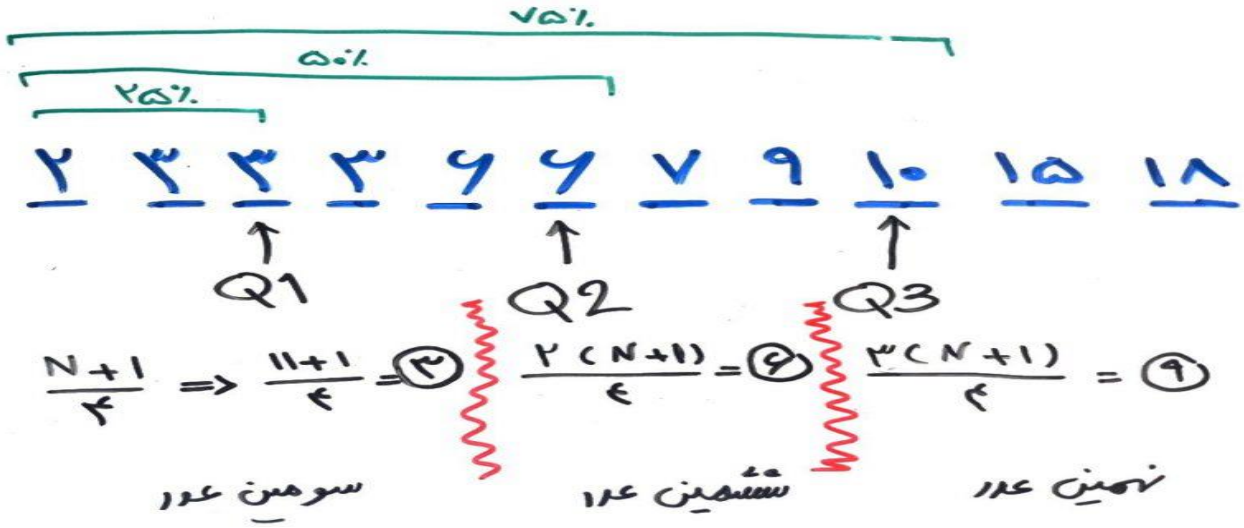
$P_1 = \frac{N}{100}$ = **برای اولین صدک**

$P_2 = \frac{2N}{100}$ = **برای دومین صدک**

توجه شود که دومین چارک $Q_2 = \frac{2N}{4}$ **فرمول و پنجمین دهک** $D_5 = \frac{5N}{10}$ **و پنجاهمین صدک** $P_{50} = \frac{50N}{100}$ **همگی همان میانه می باشند. یعنی هر یک توزیع را به دو قسمت تقسیم می کنند.**

مثال : فرض کنید داده‌های زیر، مبالغ خسارت بیمه حوادث انفرادی در ۱۱ پرونده خسارت باشد (اعداد به میلیون ریال) : ۶ – ۳ – ۲ – ۶ – ۳ – ۷ – ۱۵ – ۳ – ۹ – ۱۰ – ۱۸

برای اینکه این داده‌ها را چارک‌بندی کنیم، ابتدا آن‌ها را به ترتیب از کوچک به بزرگ مرتب کرده و داده‌ها به بخش‌های ۲۵ درصدی تقسیم می‌کنیم. چیزی مانند شکل زیر



داده‌ها را به علاوه‌ی یک کرده و سپس تقسیم بر چهار می‌کنیم تا به چارک اول برسیم. مثلاً در این مثال ۱۱ + ۱ برابر ۱۲ می‌شود و سپس ۱۲ را تقسیم بر ۴ می‌کنیم که به عدد ۳ می‌رسیم. پس باید داده‌ها سه تا سه تا تقسیم شود.

در کل سه چارک داریم، Q1، Q2و Q3. چارک اول یا همان Q1 ، عددی است که ۲۵ درصد داده‌ها کمتر از آن هستند و ۷۵ درصد داده‌ها بیشتر از آن هستند Q2 .که چارک دوم است، میانه‌ی (median) داده‌ها را نشان می‌دهد .میانه که در دروس قبلی به آن اشاره کردیم همان عددی است که ۵۰ درصد داده‌ها کمتر از آن هستند و ۵۰ درصد داده‌ها بیشتر از آن Q3 .هم که همان‌طور که حدس می‌زنید، داده‌ای است که ۷۵ درصد داده‌ها کمتر از آن و ۲۵ درصد داده‌ها بیشتر از آن هستند.

مثال:

جدول خسارات پرداختی بیمه بدنه بشرح زیر مفروض است :

فراوانی تجمعی	فراوانی	هزینه خسارات به تومان
۷	۷	۱۴۰۰ تا ۱۵۹۰
۲۷	۲۰	۱۶۰۰ تا ۱۷۹۰
۶۰	۳۳	۱۸۰۰ تا ۱۹۹۰
۸۵	۲۵	۲۰۰۰ تا ۲۱۹۰
۹۶	۱۱	۲۲۰۰ تا ۲۳۹۰
۱۰۰	۴	۲۴۰۰ تا ۲۵۹۰
	جمع = ۱۰۰	

برای بدست آوردن چارک یا دهک یا صدک این جدول از فرمول زیر استفاده میکنیم:

$$L_i + \frac{np - Fc_{i-1}}{f_i} \times C_i$$

L_i = کران پایین (به طور مثال عدد ۱۴۰۰ در جدول فوق کران پایین ردیف اول میباشد)

Np = (N = تعداد / p = برای مثال چهارک سوم برابر است با $\frac{3}{4}$ یا صدک هفدهم برابر است با $\frac{17}{100}$ یا دهک نهم برابر است با $\frac{9}{10}$)

Fc_{i-1} = فراوانی تجمعی (در این فرمول به اندیس *i-1* توجه ویژه شود) (اگر i یا همان شماره ردیف بدست امده طبق توضیحات زیر عدد ۴ شد پس فراوانی تجمعی ردیف ۳ رو لحاظ میکنیم)

f_i = فراوانی

i = شماره ردیف (شماره ردیف از طریق فرمول Np بدست می آید. در ابتدا مقدار Np را با توجه به توضیحات بالا بدست میاوریم. باید دید که عدد بدست امده زیر مجموعه کدام ردیف از فراوانی تجمعی قرار میگیرد برای نمونه سومین چارک در جدول فوق برابر ۷۵ شده است. عدد ۷۵ بیشتر از ۶۰ (یعنی فراوانی تجمعی ردیف سوم) پس

www.asia26090.ir

خارج از این محدوده قرار دارد و کمتر از ۸۵ (فراوانی تجمعی ردیف چهارم)یعنی زیر مجموعه این ردیف قرار گرفته است. پس مبنا ردیف چهارم است)

C = اختلاف کران (در جدول فوق برابر است با ۱۹۰=۱۴۰۰-۱۵۹۰)

نحوه محاسبه این مقادیر را با مثال زیر نشان میدهیم :

اگر خسارت پرداختی در بیمه بدنه را در نظر بگیریم ، سومین چارک برابر است با :

$$Q_3 = 1990 + \frac{(\frac{3}{4} \times 100 - 60)}{25} 200 = 2110$$

***** (طبق توضیحات مربوط به i مبنای این جدول ردیف ۴ (۸۵) میباشد اما دقت شود طبق فرمول Fc_{i-1}**

فراوانی تجمعی ردیف سوم (۶۰) مبنا قرار میگیرد زیرا اندیس منهای یک شده است) ***

در این حالت ۲۵درصد خسارات پرداختی مساوی یا بالای ۲۱۱۰ تومان و ۷۵ درصد خسارات پرداختی مساوی یا کمتر از ۲۱۱۰ تومان است و یا به عبارت دیگر توزیع چارک سوم از ۲۱۱۰ تا ۲۵۹۰ (کرانه بالای آخرین طبقه). دهکها و صدکها به طریقه همانندی محاسبه می شوند به کمک آنها می توان در بالای ۱۰ درصد خسارات و یا زیر ۲۴ درصد خسارات و یا سایر تقسیم بندی ها بحث کرد .

www.asia26090.ir

فصل چهارم :

شاخصهای پراکندگی در بیمه

شاخص های پـراکندگی

شاخص های پراکندگی، میزان پراکندگی (تغییرات) مقادیر هر متغیر را در اطراف میانگین نشان می دهند. زمانی که در شاخص های گرایش به مرکز، داده ها را در یک اندازه واحد خلاصه کنیم، طبیعتا بخشی از جزئیات و اطلاعات حذف خواهد شد. از این رو باید به دنبال شاخص هایی برای اندازه گیری تفاوت موردها در یک متغیر باشیم تا از نحوه و میزان پراکندگی و تغییر داده ها در اطراف میانگین (یا میانه و مد) مطلع شویم.

به عنوان مثال نمرات درس های دو دانشجو را با یکدیگر مقایسه می کنیم. معدل (میانگین نمرات) هر دو دانشجو یکسان و برابر با ۱۵ است. نمرات دانشجوی الف بدین صورت است: ۱۷، ۱۶، ۱۵، ۱۴ و ۱۳ . و نمرات دانشجوی ب بدین صورت است: ۲۰، ۱۸، ۱۶، ۱۱

و ۱۰ است. نگاهی به نمرات دو دانشجو نشان می دهد که دامنه نمرات دانشجوی الف محدودتر است و داده ها از نمره میانگین (۱۵) فاصله کمتری دارند. نتیجه می گیریم در حالتی که میانگین، یک شاخص گرایش به مرکز است در هر دو دانشجو برابر با ۱۵ است اما پراکندگی نمرات در دانشجوی ب بیشتر از دانشجوی الف است.

- **شاخص های پراکندگی چه چیز را نشان می دهند؟ میزان پراکندگی یا تغییراتی را که در بین (نمره ها ،قد ، وزن ، مبالغ حق بیمه، مبالغ خسارت و ....) وجود دارد، نشان می دهند.**
- **شاخص های پراکندگی را نام ببرید : ۱– دامنه تغییرات ۲– انحراف چارکی ۳– واریانس ۴– انحراف استاندار**
- **وجود یا نبود پراکندگی خوب است یا بد؟ مثال : قضاوت در این مورد باید بر اساس هدف پژوهش انجام بگیرد مثلا : تعیین میزان معلومات یک گروه: پراکندگی زیاد خوبه و سطح یادگیری پس از تدریس یا آموزش: پراکندگی کم خوبه**

شاخص های گرایش به مرکز اطلاعاتی از نحوه پراکندگی داده ها و نحوه توزیع‌شان به ما نمی دهند و جهت اطلاع از نحوه پراکندگی داده ها، باید از شاخص های پراکندگی استفاده کنیم. از مهم ترین شاخص های پراکندگی می توان به انحراف معیار / استاندارد، واریانس، ضریب تغییرات و دامنه تغییرات اشاره کرد. ساده‌ترین شاخص پراکندگی دامنه است که تفاوت بین بزرگ‌ترین و کوچک‌ترین مقدار است .

امنه :روشی ساده برای بررسی پراکندگی داده‌ها یافتن دامنه یا به عبارتی اختلاف بزرگترین و کوچکترین مشاهده است. اما هنگامی که داده‌های بسیار بزرگ یا بسیار کوچک وجود داشته باشند این معیار رضایت بخش نبوده و معرف خوبی برای بیان میزان انحراف داده‌ها نیست. در این‌گونه موارد استفاده از واریانس مطلوب است.

دامنه

سریعترین و ساده ترین روش برای پراکندگی، محاسبه اختلاف بین بیشترین و کمترین مقدار متغیر بوده که به دامنه موسوم است که. اگر دامنه را با R و مقدار حد پائین متغیر را به L و حد بالا را به U نشان دهیم در این صورت داریم : $R=U-L$

توزیع فراوانی زیر را در نظر میگیریم : تعداد کارکنان استخدام شده در ادارات مختلف شرکتهای A و B بصورت زیر است:

تعداد کارکنان	اداره‌های شرکت A	اداره‌های شرکت B
$۵ \leqslant n < ۱۰$	۰	۵
$۱۰ \leqslant n < ۱۵$	۳	۸
$۱۵ \leqslant n < ۲۰$	۱۳	۱۰
$۲۰ \leqslant n < ۲۵$	۲۴	۱۲
$۲۵ \leqslant n < ۳۰$	۱۷	۱۴
$۳۰ \leqslant n < ۳۵$	۳	۵
$۳۵ \leqslant n < ۴۰$	۰	۳
$۴۰ \leqslant n < ۴۵$	۰	۳
	جمع ۶۰	جمع ۶۰

که با توجه به مقدار n برای شرکتA برابر است با ۲۴ = ۱۰ – ۳۴ و برای شرکت B برابر است با ۳۹ = ۵ – ۴۴

واریانس

با به توان دو رساندن انحراف استاندارد، واریانس به دست می آید (به بیان صحیح تر، جذر واریانس برابر با انحراف استاندارد است). اما چون تفسیر واریانس کمی دشوار است در گزارش ها و پژوهش ها از انحراف استاندارد به جای واریانس استفاده می شود. انحراف استاندارد بزرگ‌نمایی حاصل از مجذورکردن (به توان دو رساندن) را تا حد زیادی خنثی می سازد و تفسیر میزان پراکندگی را آسان تر

و قابل فهم تر می سازد. در دنیای کسب‌وکار واریانس یکی از پرکاربردترین شاخص‌ها برای نشان دادن میزان پراکندگی داده‌ها است.

انحراف معیار یا انحراف استاندارد

با توجه استفاده از توان ۲ در محاسبه واریانس، واحد اندازه‌گیری برای این شاخص به صورت مربع در خواهد آمد. برای مثال، اگر داده‌ها برحسب گرم باشند، واریانس برحسب گرم مربع (گرم به توان ۲) خواهد بود. این موضوع باعث می‌شود که نتوان واریانس را با خود داده‌ها مقایسه کرد.

برای رفع این مشکل کافی است که جذر واریانس را محاسبه کنیم تا واحد اندازه‌گیری این شاخص پراکندگی با واحد اندازه‌گیری داده‌ها یکی شود. به حاصل این کار «انحراف معیار (Standard Deviation)» می‌گویند. انحراف معیار جامعه آماری با σσو انحراف معیار نمونه S نشان داده می‌شود.

پس شیوه محاسباتی برای انحراف معیار به صورت زیر خواهد بود؛

σ=√σ2σ=σ2

S=√S2

ضریب تغییرات و خطای نسبی

یکی از شاخص‌های آمار توصیفی که اغلب برای نمایش و مقایسه میزان پراکندگی بین دو جامعه یا دو متغیر به کار می‌رود، «ضریب تغییرات (Coefficient of Variation)» است. از آنجایی که کم بودن پراکندگی، نشانگر همگن بودن جامعه است، هر چه میزان ضریب تغییرات کمتر باشد، میانگین را معیار بهتری برای نقطه تمرکز می‌یابیم. بنابراین در بین دو جامعه، آن که دارای ضریب تغییرات کمتری باشد، جامعه بهتری بوده، زیرا نتایج گرفته شده از شاخص میانگین، دقت بیشتری دارند.

از آنجایی که ضریب تغییرات بوسیله یک نسبت از کمیت‌های هم واحد، محاسبه و ساخته می‌شود، هیچ واحد اندازه‌گیری نداشته و به صورت درصدی مورد استفاده قرار می‌گیرد. همین موضوع نیز اهمیت استفاده از این شاخص را برای مقایسه بین جوامع مختلف، مشخص می‌کند. انحراف معیار

$$CV = \frac{\text{انحراف معیار}}{\text{میانگین}}$$

مثال – شاخص پراکندگی

مثال : اعداد حق بیمه بشرح زیر مفروض است مطلوبست محاسبه میانگین، واریانس و انحراف معیار را پیدا کنید.

گام اول یافتن میانگین است:

۶۰۰- ۴۷۰- ۱۷۰- ۴۳۰- ۳۰۰

$$\text{میانگین} = \frac{600 + 470 + 170 + 430 + 300}{5} = \frac{1970}{5} = 394$$

گام دوم : اکنون اختلاف مبلغ هرکدام از خسارات را از مقدار میانگین حساب می کنیم . مانند :

۶۰۰ -۳۹۴=۲۰۶

۴۷۰ – ۳۹۴ = ۷۶

الی آخر

گام سوم : برای محاسبه واریانس، اختلاف تک تک داده ها را به توان دو رسانده و سپس میانگین می گیریم:

بنابراین متوسط توان دوم انحراف داده ها از میانگین را واریانس که با S2 / زیگما ۲ نمایش داده می شود.

$$\text{واریانس: } \sigma^2 = \frac{206^2 + 76^2 + (-224)^2 + 36^2 + (-94)^2}{5} = \frac{42436 + 5776 + 50176 + 1296 + 8836}{5}$$

$$= \frac{108520}{5} = 21704$$

گام چهارم : انحراف معیار که همان جذر واریانس است، پس:

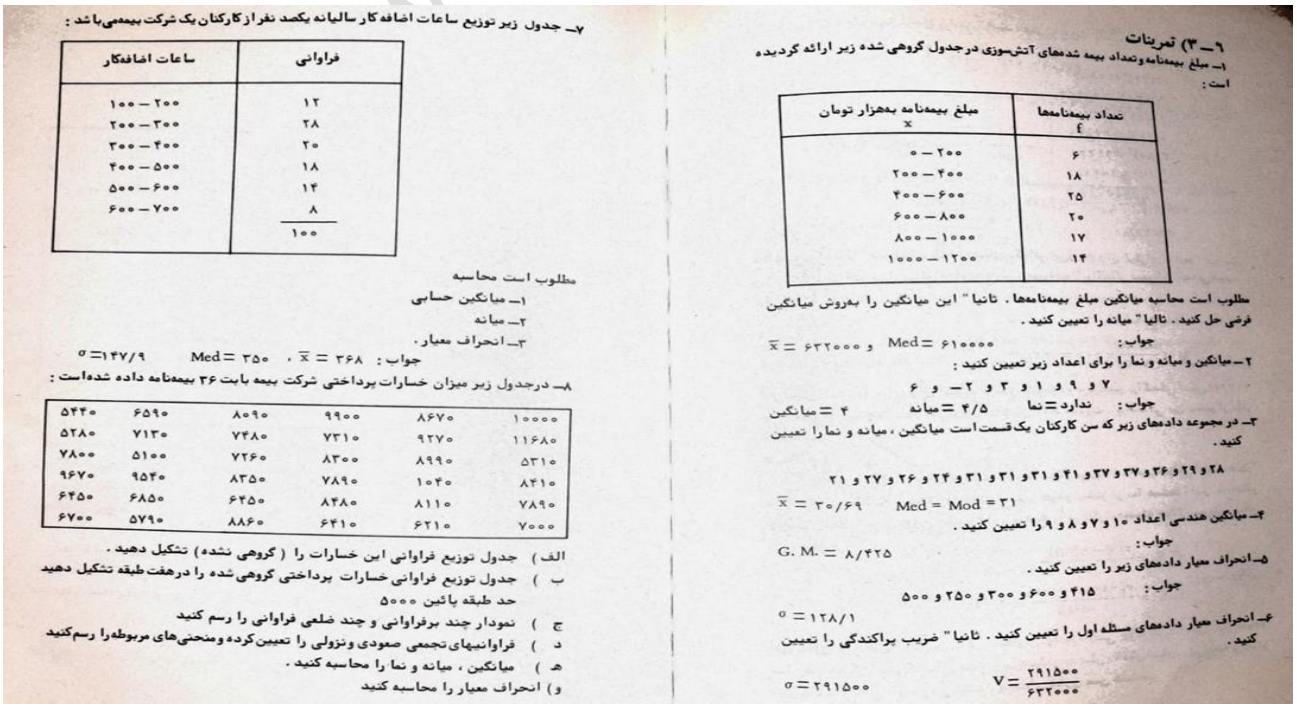
$$\sigma = \sqrt{21704} = 147.32 \ldots = 147 \text{ (نزدیک ترین داده)}$$

فرمول انحراف معیار

و اما نکته خوب درمورد انحراف معیار، سودمند بودن آن است. اکنون می توانیم بفهمیم مبلغ کدام خسارت در محدوده یک انحراف معیار میانگین(۱۴۷میلیون ریال) است:

پس با استفاده از انحراف معیار، ما یک راه "استاندارد" برای یافتن مقدار نرمال، مقدار بیش از نرمال و مقدار کمتر از نرمال در دست داریم.

تمرینات :



۶ــ۳) تمرینات

۱ــ مبلغ بیمه‌نامه و تعداد بیمه شده‌های آتش‌سوزی در جدول گروهی شده زیر ارائه گردیده است :

تعداد بیمه‌نامه‌ها f	مبلغ بیمه‌نامه به‌هزار تومان x
۶	۰ – ۲۰۰
۱۸	۲۰۰ – ۴۰۰
۲۵	۴۰۰ – ۶۰۰
۲۰	۶۰۰ – ۸۰۰
۱۷	۸۰۰ – ۱۰۰۰
۱۴	۱۰۰۰ – ۱۲۰۰

مطلوب است محاسبه میانگین مبلغ بیمه‌نامه‌ها . ثانیا " این میانگین را به‌روش میانگین فرضی حل کنید . ثالثا " میانه را تعیین کنید .

جواب : $\bar{x} = 632000$ و $Med = 610000$

۲ــ میانگین و میانه و نما را برای اعداد زیر تعیین کنید :

۷ و ۹ و ۱ و ۳ و ۲ – و ۶

جواب : ندارد = نما ۴/۵ = میانه ۴ = میانگین

۳ــ در مجموعه داده‌های زیر که سن کارکنان یک قسمت است میانگین ، میانه و نما را تعیین کنید .

۲۸ و ۲۹ و ۳۶ و ۳۷ و ۴۱ و ۳۱ و ۳۱ و ۲۴ و ۲۶ و ۲۷ و ۲۱

$\bar{x} = 30/69$ $Med = Mod = 31$

۴ــ میانگین هندسی اعداد ۱۰ و ۷ و ۸ و ۹ را تعیین کنید .

جواب : $G.M. = 8/425$

۵ــ انحراف معیار داده‌های زیر را تعیین کنید .

۴۱۵ و ۶۰۰ و ۳۰۰ و ۲۵۰ و ۵۰۰

جواب : $\sigma = 128/1$

۶ــ انحراف معیار داده‌های مسئله اول را تعیین کنید . ثانیا " ضریب پراکندگی را تعیین کنید .

$\sigma = 291500$ $V = \frac{291500}{632000}$

۷ــ جدول زیر توزیع ساعات اضافه کار سالیانه یکصد نفر از کارکنان یک شرکت بیمه‌می‌باشد :

فراوانی	ساعات اضافه‌کار
۱۲	۱۰۰ – ۲۰۰
۲۸	۲۰۰ – ۳۰۰
۲۰	۳۰۰ – ۴۰۰
۱۸	۴۰۰ – ۵۰۰
۱۴	۵۰۰ – ۶۰۰
۸	۶۰۰ – ۷۰۰
۱۰۰	

مطلوب است محاسبه

۱ــ میانگین حسابی

۲ــ میانه

۳ــ انحراف معیار .

جواب : $\bar{x} = 368$ ، $Med = 350$ $\sigma = 147/9$

۸ــ درجدول زیر میزان خسارات پرداختی شرکت بیمه بابت ۳۶ بیمه‌نامه داده شده‌است :

۵۴۴۰	۶۵۹۰	۸۰۹۰	۹۹۰۰	۸۶۷۰	۱۰۰۰۰
۵۲۸۰	۷۱۳۰	۷۴۸۰	۷۳۱۰	۹۲۷۰	۱۱۶۸۰
۷۸۰۰	۵۱۰۰	۷۲۶۰	۸۳۰۰	۸۹۹۰	۵۳۱۰
۹۶۷۰	۹۵۴۰	۸۳۵۰	۷۸۹۰	۱۰۴۰	۸۴۱۰
۶۴۵۰	۶۸۵۰	۶۴۵۰	۸۴۸۰	۸۱۱۰	۷۸۹۰
۶۷۰۰	۵۷۹۰	۸۸۶۰	۶۴۱۰	۶۲۱۰	۷۰۰۰

الف) جدول توزیع فراوانی این خسارات را (گروهی نشده) تشکیل دهید .

ب) جدول توزیع فراوانی خسارات پرداختی گروهی‌شده را درهفت طبقه تشکیل دهید حد طبقه پائین ۵۰۰۰

ج) نمودار چند برفراوانی و چند ضلعی فراوانی را رسم کنید

د) فراوانیهای تجمعی صعودی ونزولی را تعیین‌کرده ومنحنی‌های مربوطه‌را رسم‌کنید

هـ) میانگین ، میانه و نما را محاسبه کنید .

و) انحراف معیار را محاسبه کنید

فصل پنجم :

قوانین شمارش و احتمال

➢ قواعد شمارش در آمار

چنانچه گفتیم برای محاسبه احتمال وقوع یک پیشامد نیاز به محاسبه تعداد اعضای آن و تعداد اعضای فضای نمونه داریم . اغلب او قات شمارش اعضای یک مجموعه کار دشواری است و برای انجام این کار نیاز به شناسایی برخی اصول و قوانین داریم که به بررسی آنها می پردازیم.

اصل ضرب فرض کنید یک کار را بتوان با دو عمل پیاپی A,B انجام داد اگر عمل A به m طریق و به دنبال آن عمل B بتواند به n طریق انجام پذیرد آنگاه این کار به mn طریق انجام می پذیرد.

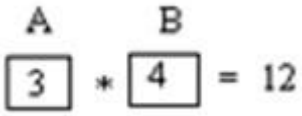
مثال : با ارقام ۰,۱,۲,۳ چند عدد دو رقمی می توان نوشت در صورتی که :

الف) تکرار ارقام مجاز باشد.

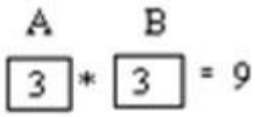
ب) تکرار ارقام مجاز نباشد.

حل : نوشتن یک عدد دو رقمی شامل دو عمل انتخاب رقم دهگان (A) و انتخاب رقم یکان (B) می باشد. بنابراین

الف) رقم دهگان می نواند یکی از ارقام ۳,۲,۱ به سه طریق و رقم یکان یکی از ارقام ۳,۲,۱,۰ به چهار طریق باشد. بنابراین



ب) رقم دهگان می تواند یکی از ارقام ۳.۲.۱ به سه طریق و رقم یکان می تواند رقم ۰ یا یکی از دو رقم باقی مانده از ارقام ۳,۲,۱ باشد. بنابراین



اصل جمع فرض کنید یک کار را بتوان با دو عملB یا Aانجام داد.اگر عمل A به m طریق و عمل B بهn طریق انجام پذیرد و این دو عمل نتوانند همزمان اتفاق بیفتند آن گاه این کار به m + n طریق انجام می پذیرد.

مثال به چند طریق می نوان از بین ۴ دانشجوی کامپیوتر و ۵ دانشجوی ریاضی دو دانشجو را انتخاب کرد به طوری که نفر اول به عنوان

سرگروه و نفر دوم به عنوان دستیار باشد و هر دو نفر از یک رشته باشند.

حل: طبق اصل ضرب دو دانشجو از رشته کامپیوتر به ۴ * ۳ = ۱۲ طریق یا از رشته ریاضی به ۵ * ۴= ۲۰ طریق انتخاب می شوند . بنابراین طبق اصل جمع این دو نفر را می توان به ۲۰ + ۱۲ = ۳۲ طریق انتخاب کرد.

جایگشتها

در بعضی مسائل می خواهیم تعداد طریق قرارگرفتن ۶ نفر در یک صف ویا تعداد طریق انتخاب ۲ تفر از بین ۶ نفر را به دست آوریم

برای این منظور می توان از اصول شمارش و مفهوم جایگشت استفاده کرد.ترتیبی را که می توان اشیا یک مجموعه را در کنار یکدیگر قرار داد یک جایگشت گویند. در حالت کلی داریم

اگر n عنصر متمایز را بخواهیم در یک صف کنار یکدیگر قرار دهیم جایگشتهای مختلف این عناصر برابراست با

$n=n(n-1)(n-2)...(2)(1)!$

و اگر تکرار مجاز باشد جایگشتهای این عناصر برابراست با

$$\frac{n!}{n1!n2!...nr!}$$

مثال:

چهار پزشک و پنج مهندس می خواهند در یک صف کنار یکدیگر قرار گیرند

الف احتمال اینکه مهندس ها در یک صف و پزشک ها در صف دیگر قرار گیرند را بیابید.

ب احتمال اینکه دو مهندس بخصوص کنار هم قرار گیرند را بیابید.

حل:

در این حالت $n(s)=9!$یعنی تعداد کل حالات قرار گرفتن این ۹ نفر در صف می باشد.

الف) اگر A پیشامد قرار گرفتن پزشک ها در یک صف و مهندس ها در طرف دیگ صف باشد چون شروع صف می تواند با پزشک ها یا مهندس ها باشد پس $n(A)= 2 ۵! ۴!$و در نتیجه

$$P(A) = \frac{n(A)}{n(S)} = \frac{2!4!5!}{9!}$$

ب) اگر B پیشامد قرار گرفتن پزشک ها و مهندس ها یک در میان در صف باشد آن گاه مهندس ها به ۵! طریق در صف قرار می گیرند وپزشک ها به ۴! طریق در بین مهندس ها قرار می گیرند. پس$P(B)=5!4!$ و در نتیجه

اگر از بین n عنصر متمایز بخواهیم r عنصر را

انتخاب کرده و در یک صف قرار دهیم در این صورت

الف – اگر تکرار عناصر مجاز نباشد آنگاه تعداد راه های ممکن برابراست با:

ب اگر تکرار عناصر مجاز باشد آنگاه تعداد راههای ممکن برابراست با

$$P_r^n = \frac{n!}{(n-r)!}$$

اگر از بین n عنصر متمایز بخواهیم r عنصر را انتخاب کنیم به طوری که ترتیب انتخاب مهم نباشد در این صورت

الف – اگر تکرار عناصر مجاز نباشد آنگاه تعداد راه های ممکن برابراست با

$$\left[\begin{matrix} n \\ r \end{matrix}\right] = \frac{n!}{r!\,(n-r)!}$$

ب – اگر تکرار عناصر مجاز باشد آنگاه تعداد راه های ممکن برابراست با

$$\left[\begin{matrix} n+r-1 \\ r \end{matrix}\right]$$

مثال:

از بین ۴ پزشک و ۵ پرستار می خواهیم یک کمیته ۴ نفری تشکیل دهیم.

احتمال اینکه اعضای کمیته شامل ۲ پزشک و ۲ پرستار باشد را بیابید.

حل:

چون در انتخاب افراد ترتیب مهم نیست بنابراین تعداد انتخاب ۴ نفر از این ۷ نفر برابراست با

$$n(S) = \left[\begin{matrix} 7 \\ 4 \end{matrix}\right]$$

انتخاب ۲ پزشک به و انتخاب دو پرستار به طریق انجام می شود بنابراین اگر A پیشامد انتخاب ۲ پزشک و ۲ پرستار باشد آنگاه

$$P(A) = \frac{\left[\begin{matrix} 3 \\ 2 \end{matrix}\right]\left[\begin{matrix} 4 \\ 2 \end{matrix}\right]}{\left[\begin{matrix} 7 \\ 4 \end{matrix}\right]}$$

➢ نظریه احتمال

نظریه احتمال مکانیسمی به دست می‌دهد که می توان به کمک آن احتمال وقوع این رویداد را تعیین کرد . اندازه احتمال مقیاسی بین صفر و یک دارد. صفر برای احتمال رویداد غیر ممکن و یک برای احتمال رویداد حتمی به کار می رود . نقاط واقع بین صفر و یک مبین احتمال است و هر اندازه نزدیک به یک باشد احتمال وقوع آن بیشتر و هرچه به صفر نزدیک باشد احتمال رویداد کمتر است. اگر احتمال حادثه ۰/۴ و دیگری ۰/۲ باشد انتظار وقوع اولی بیشتر از دومی است.

نماد احتمال: هنگامی که احتمال در محاسبات نوشته یا به کار می رود، احتمال وقوع یک پیشامد را با حرف **p** نشان می دهند در نتیجه احتمال پیشامد تصادف اتومبیل را با P(A) نشان میدهند این طریقه نمایش از نوشتن مطالب زیاد جلوگیری کرده و به راحتی مفهوم را می رساند.

چگونگی محاسبه احتمال : در این قسمت به سه اصل اساسی که در محاسبه احتمال به کار میرود اشاره می کنیم.

۱. احتمال پیشین

روش احتمال پیشین برای حالاتی به کار می رود که تمامی پیشامد های ممکن از قبل شناخته شده و در ضمن احتمال وقوع هر یک از آنها مساوی باشد. مانند پرتاب سکه و و دو پیشامد با احتمال مساوی روی دهد. یا شیر می آید یا خط . شانس آمدن شیر ۵۰ ٪ است . با تقسیم حالات مساعد بر حالات ممکن مقدار احتمال به صورت زیر تعیین می شود

۰/۵ = ۱/۲ = حالات ممکن / حالات مساعد

غالباً مقدار احتمال را به صورت اعشاری نشان می‌دهند و کمتر از درصد استفاده می کنند همانند این مسئله را در مورد احتمال انتخاب دو نفر از میان ده نفر به صورت زیر نشان داد :

۰/۵ = ۲/۱۰ = حالات ممکن / حالات مساعد

۲. احتمال تجربی(فراوانی نسبی)

بزرگترین مانع در کاربرد روش احتمال پیشین عدم اطلاع از تمامی پیشامد های ممکن است . در عمل حالاتی که بتوان کلیه پیشامد های ممکن را از قبل شناخت کم بوده از اینرو روش های دیگری نیز به کار می رود .

در شروع سال یک شرکت بیمه نمی‌داند چه تعداد خسارت آتش سوزی خواهد داشت، چند حادثه تصادف پیش خواهد آمد ، چه تعداد بیمه‌نامه صادره خواهد کرد وغیره .

در پایان سال این شرکت میداند چه تعداد خسارت آتش سوزی دارد، چند حادثه تصادف پیش آمده و چه تعداد بیمه نامه جدید صادر کرده است . ولی قادر نیست که معین کند چه تعداد آتش سوزی اتفاق نیفتاده ، چه تعداد حادثه تصادف پیش نیامده و یا چه تعداد بیمه نامه صادر نشده است. این ارقام ناشناخته مانده و در نتیجه تمامی حالات ممکن ناشناخته باقی می مانند .

این شرکت بیمه چه می تواند انجام دهد؟ شرکت می تواند کارنامه سال گذشته را بررسی کند ، از تعداد آتش سوزی ها ، از تعداد حوادث اتومبیل ، از تعداد بیمه‌نامه‌های صادر شده مطلع گردد و به همین ترتیب از بررسی سال های بیشتر نیز اطلاعاتی به دست آورد .

شرکت بیمه با نگرشی بر گذشته ، فراوانی نسبی حوادث را تعیین می‌کند ، با فرض اینکه در ماهیت رویدادها تغییری حاصل نشده باشد ، منطقی است که فرض کنیم روند گذشته که موسوم به احتمال تجربی است برای آینده ، تخمین مناسبی است . به عبارت دیگر در این حالت احتمال برابر است با فراوانی نسبی.

مثلاً فرض کنیم از میان ۱۰۰۰۰ بیمه گذار آتش‌سوزی ۱۰۰ خسارت در سال گذشته پرداخت شده باشد . می‌توان گفت احتمال وقوع آتش سوزی برای سال آینده برابر است با :

$$p = ۱۰۰ / ۱۰۰۰۰ = ۰/۰۱$$

اگر چه روش های مختلفی برای محاسبه احتمال با استفاده از احتمال پیشین و احتمال تجربی به کار می رود ولی پس از محاسبه، تفسیر یکسانی برای هر دو احتمال وجود دارد.

اگرچه در پرتاب سکه، احتمال آمدن شیر ۱/۲ است ولی نمی توان گفت که اگر شش سکه را پرتاب کنیم حتماً سه شیر می آید . یا نمی توان گفت که از میان یکصد بیمه نامه آتش سوزی حتماً یک آتش‌سوزی به وقوع خواهد پیوست و شرکت مجبور با پرداخت خسارت خواهد بود . در پرتاب سکه ممکن است شش شیر یا پنج شیر ، .... و یا ترکیبات دیگری از شیر و خط بیاید. در آتش‌سوزی نیز ممکن است هر تعدادی آتش سوزی اتفاق افتاده و شرکت مجبور به پرداخت خسارت باشد.

مقدار به دست آمده برای احتمال یک راهنما است و یک پیش بینی کامل نیست ولی برای دوره های بعد و تعداد زیاد و در دراز مدت قابل اعتماد است.

۳. احتمال ذهنی

دانستیم که مانعی برای محاسبه احتمال پیشین وجود دارد، برای محاسبه احتمال به روش فراوانی نسبی نیز ، مانعی وجود دارد. همانطوری که می دانیم در محاسبه احتمال به روش فراوانی نسبی ، محاسبه متکی بر اطلاعات قبلی است. مانع اصلی در محاسبه این روش هنگامی پیش می آید که اطلاعاتی از گذشته نداشته باشیم .

اگر بررسی در مورد حادثه منحصر به فرد و یا رویداد هایی که تنها چند بار رخ داده و درباره آنها اطلاعات کمی وجود دارد باشد ، در این صورت از احتمال ذهنی استفاده می کنند . فرض کنید یک شرکت بیمه، بیمه نامه می خواهد بفروشد ، روشن است که نمی تواند حالات ممکن و صور مناسب را تعیین کند، لذا از احتمال پیشین استفاده نمی شود در ضمن اطلاعات آماری از گذشته نیز وجود ندارد که از احتمال فراوانی نسبی یا تجربی استفاده کند در این صورت تنها راه نسبت دادن احتمالی معین بر اساس ترکیبی از عقیده شخصی و تجربیات گذشته و تحلیل منطقی و موسوم به احتمال ذهنی است . گاهی محاسبه احتمال به روش پیشین و تجربی، موسوم به احتمال عینی یا شهودی است. امروز از احتمال ذهنی در بسیاری از تصمیم گیری های تجاری استفاده می کند

➢ کاربرد احتمال

با داشتن اندازه احتمال وقوع یک حادثه گاهی ضروری به نظر می رسد که از این احتمال در برخی محاسبات استفاده کنیم. اکنون دو راه اصلی که می توان احتمالات مختلفه را باهم ترکیب کرد بررسی می کنیم :

۱. رویداد متناوب

فرض کنیم شما دارای سهامی شرکت بیمه بوده و علاقمند به تخمین احتمال افزایش قیمت سهام خود باشید، دو امکان اصلی وجود دارد : یکی افزایش قیمت سهام دیگری کاهش قیمت آنها. فرض کنید A رویداد افزایش قیمت و B احتمال کاهش یا افزایش قیمت سهام باشد. در این صورت P (A یا B) احتمال افزایش یا کاهش قیمت سهام است. فرض کنیم به طریقی می دانیم P(A) = ۰/۵ و P(B) = ۰/۲۵ در این صورت P (A یا B) چه مقدار است؟ در این حالت از قانون جمع استفاده می کنیم . یعنی

$$P(A \text{ یا } B) = P(A) + P(B)$$

$$P(\text{A یا B}) = ۰/۵ + ۰/۲۵ = ۰/۷۵$$

اگر رویدادی در ۵۰٪ زمان و حادثه دیگری در ۲۵ ٪ زمان پیش آید، منطقی است احتمال پیش آمدن یکی از این دو را ۷۵ ٪ یعنی مجموع اعمال این دو در نظر بگیریم .

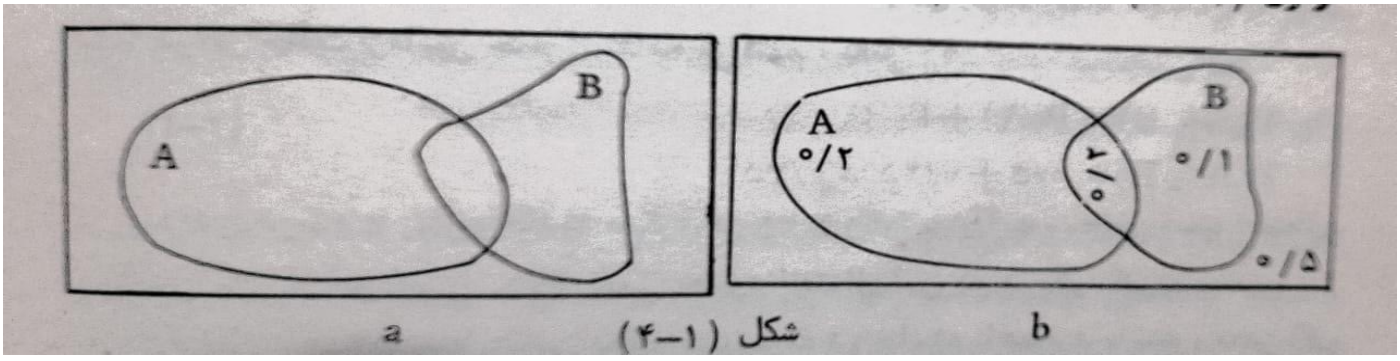
فرمول فوق تنها زمانی به کار می رود که دو حادثه نا همآیند باشند. نا همآیند یعنی اگر یک رویداد واقع شد حتما" رویداد دیگر واقع نگردد . قیمت سهام نمی تواند همزمان افزایش و کاهش یابد ، یک سکه نمی‌تواند در یک زمان هم شیر بیاید و هم خط . اینها رویدادهای ناهمآیند هستند. فرض کنیم میزان خسارتی که برای تصادف منجر به جرح در بیمه اتومبیل برای دیه پرداخت می شود ۱۰۰ هزار تومان و برای فوت ۵۰۰ هزار تومان باشد . در این صورت در یک تصادف امکان جرح و فوت برای یک نفر با هم وجود ندارد. این نیز مثالی از رویدادهای ناهمایند است، لذا بیمه دیه برای یک نفر و ۱۰۰ هزار تومان است و یا ۵۰۰ هزار تومان.

۲. رویدادهای همآیند

رویداد های مختلفی وجود دارد که می توان با هم رخ دهند آنها را رویدادهای همآیند گویند، برای مثال وضعیت ۱۰۰ متقاضی کار را بررسی می کنیم ۴۰ نفر دارای تجربه و ۳۰ نفر دارای مدرک لازم و ۲۰ نفر هم دارای تجربه و مدرک می باشد .

بهترین روش برای نمایش این رویداد ها استفاده از نمودار ون است (شکل زیر a) . تمامی حالات ممکن را در داخل مستطیلی در نظر گرفته و هر مجموعه از رویدادهایی که مساعد و یا دارای خاصیتی مشترک اند در داخل منحنی های بسته در داخل مستطیل رسم می کنند. اگر در مثال بالا کلی متقاضیان یعنی ۱۰۰ نفر را در داخل مستطیل به شکلی U نشان دهیم و آن را فضای ممکن بنامیم، مجموعه A را افرادی که تجربه داشته و در مجموعهB را افرادی که مدرک لازم را دارند در نظر بگیریم.

در شکل زیر b احتمال متناظر هر یک نوشته شده است ، مطابق شکل احتمال آن که فرد انتخاب شده دارای تجربه باشد ۰/۴ و دارای مدرک باشد ۰/۳ و دارای مدرک و تجربه باشد ۰/۲ و فاقد مدرک و تجربه ۰/۵ است . حال اگر می خواستیم احتمال آن که فرد انتخاب شده هم مدرک و تجربه داشته باشد را محاسبه کنیم می بایست از فرمول $P(\text{A یا B}) = P(A) + P(B)$ استفاده کنیم.



شکل (۱-۴)

$$P(\text{A یا B}) = P(A) + P(B) \quad = ۰/۴ + ۰/۳ = ۰/۷$$

ولی از روی نمودار ون به راحتی مشاهده می‌شود که این احتمال برابر ۰/۵ است. حال این اشکال را چگونه باید برطرف ساخت . پاسخ به این سوال آنستکه وقتی $P(A)$ را بر $P(B)$ می افزائیم. قسمت مشترک دو مجموعه A و

www.asia26090.ir

B را دو بار به حساب می آوریم، لذا می بایست این قسمت را یک بار به حساب آوریم در نتیجه فرمول زیر به دست می آید :

$$P(\text{A یا B}) = P(A) + P(B) - P(\text{A و B})$$

$$P(\text{A یا B}) = ۰/۴ + ۰/۳ - ۰/۲ = ۰/۵$$

رابطه فوق را می توان برای محاسبه احتمال دو حادثه نا همآیند با مانعه الجمع نیز به کار برد در حالت اخیر روشن است که $P(\text{A و B}) = ۰$ است رابطه فوق برای محاسبه احتمال رویداد های نا همآیند همآیند به کار می رود.

۳. احتمال رویداد های مستقل

حال احتمال دو رویداد را که همزمان با هم و مستقل از یکدیگر رخ می دهند بررسی می کنیم.

فرض میکنیم دو کارخانه داشته باشیم ، A در تهران و B در اصفهان. احتمال آتش سوزی برای A ، $P(A)$ و برای B ، $P(B)$ می باشد حال احتمال آنکه این دو کارخانه آتش سوزی اتفاق افتد برابر با :

$$P(\text{A و B}) = P(A) . P(B)$$

اگر احتمال آتش سوزی هر یک ۰/۲ باشد، احتمال آتش سوزی دو کارخانه برابر است با :

$$P(\text{A یا B}) = ۰/۲ \times ۰/۲ = ۰/۰۴$$

به بیان دیگر احتمال رویداد مستقل ،کمتر از احتمال هر یک از رویدادها است و این موضوع منطقی به نظر می رسد . برای مثال اگر ۱۰۰ کارخانه را در نظر بگیریم، احتمال آتش سوزی همه با هم عددی فوق العاده کوچک است ، توجه شود در این حالت احتمال وقوع هر یک اثری بر دیگری ندارد و بدین لحاظ آنها را رویدادهای مستقل میدان می نامند.

۴. احتمال رویدادهای وابسته

فرمول $P(\text{A و B}) = P(A) . P(B)$ فرمول بسیار خوبی برای محاسبه احتمال و رویدادهای مستقل می باشد . اگر دو رویداد مستقل باشند بدین معنی که وقوع یکی هیچگونه اثری بر وقوع دیگری ندارد . در مثال آتش سوزی دو کارخانه این استقلال کاملا مشهود است.

اگر دو کارخانه تنها یک متر از هم فاصله داشته باشند . در این صورت آتش سوزی آنها از یکدیگر مستقل نبود. یعنی اگر آتش سوزی در یکی به وقوع می پیوست در این صورت برای محاسبه احتمال و آتش‌سوزی این دو کارخانه با هم از رابطه زیر استفاده می گردید :

$$P(\text{A و B}) = P(A) . P(B/A)$$

که در آن $P(B/A)$ احتمال وقوع B است اگر A اتفاق افتاده باشد . فرمول $P(B/A)$ را احتمال شرطی B زمانی که A اتفاق افتاده باشد گویند . فرض می کنیم احتمال وقوع آتش‌سوزی در کارخانه اول $P(A) = ۰/۲$ باشد و در این کارخانه نیز آتش سوزی اتفاق افتاده باشد ، احتمال آتش سوزی در این صورت در کارخانه B افزایش یافته و $P(B/A)= 0/8$ می شود در این صورت احتمال آتش سوزی در هر دو کارخانه برابر است با:

$$P(\text{A و B}) = P(A) . P(B/A)= 0/2\times0/8=0/16$$

همان طوری که انتظار داشتیم احتمال آتش سوزی مشترک دو کارخانه که نزدیک هم هستند که یکی بر دیگری اثر می گذارد، از احتمال آتش سوزی دو کارخانه که در دو شهر مختلف قرار دارد بیشتر است . از این احتمال شرطی می توان برای مثال در محاسبه اثر جلوگیری از آتش سوزی و یا مکانیسمی برای کاهش آتش سوزی ، استفاده کرد . می‌دانیم که احتمال آتش سوزی در هر دو کارخانه ۱۶ / ۰ است ، در اینصورت می‌بایست دنبال مکانیسم هایی مانند ماشین های آبپاش و یا کپسول های آتش خاموش کن باشیم تا باعث کاهش احتمال آتش سوزی شده و در نتیجه در مورد هزینه آنها بتوان تصمیم گرفت.

گفتیم که به شانس وقوع حادثه احتمال گفته می‌شود احتمال پیش آمد قطعی، برابر با یک و احتمال پیشامد نا ممکن برابر صفر است. میزان احتمال وقوع یا عدم وقوع هر حادثه ای ما بین صفر تا یک در نوسان است . احتمال وقوع ۱ بیانگرحتمیت وقوع حادثه است و در این نقطه ریسک / عدم اطمینان برابر با صفر خواهند بود .

زمانیکه شانس وقوع واقعه ای ۵۰ درصد تخمین زده شود، با بیشترین ریسک / عدم اطمینان مواجه خواهیم بود. با افزایش یا کاهش احتمال از ۵۰ درصد، ریسک / عدم اطمینان رو به کاهش می گذارد .

بعبارت دیگر احتمال وقوع هر رویداد ممکن است از ۱۰۰ درصد شروع شده و به صفر درصد می رسد. در دو نقطه یک و صفر قطعیت وجود داشته و نامعلومی یا عدم اطمینان و ریسک وجود ندارد . بالاترین نقطه نامعلومی یا ریسک موقعی است که شانس احتمال وقوع یا عدم وقوع حوادث مساوی یا ۵۰/۵۰ باشد . به این مثال توجه کنید :

فرض کنید از شما سوال شود " آیا فردا باران خواهد آمد؟ " . نتایج حاصل از این سوال تشکیل یک فضای نمونه را خواهد داد :

آیا فردا باران خواهد آمد ؟

۱) خیر (قطعیت عدم بارش) = احتمال صفر درصد

۲) به احتمال ۲۰٪ باران می آید

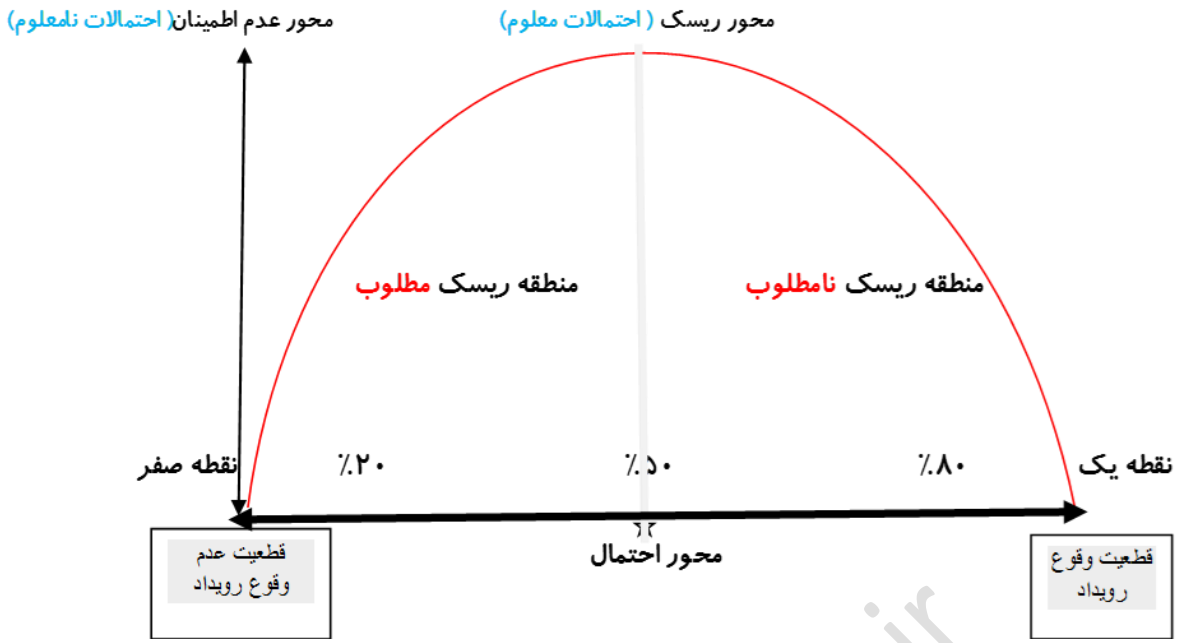
۳) به احتمال ۵۰٪ باران می آید

۴) به احتمال ۸۰٪ باران می آید

۵) بله (قطعیت بارش) = احتمال ۱۰۰٪

www.asia26090.ir

نمودار احتمال / ریسک / عدم اطمینان



ریسک دلیل وجود بیمه است، ریسک نتیجه نامطلوب و ناخوشایند و زیان آور حادثه آینده است. حادثه احتمالی و عمدتاً ناخوشایند نسبت به جان و یا مال بیمه گذار. خطرات مختلف که جان و مال انسان را تهدید می کند موجب گرایش انسان به سوی بیمه می شود.

– یکی از ویژگی‌های خطرهای بیمه‌پذیر این است که تناوب داشته باشند. تناوب خطرها به چه مفهوم است؟ **خطر بیمه‌پذیر باید احتمالی و اتفاقی و درجه احتمال آن بین صفر و یک باشد. تناوب خطر به مفهوم آن است که خطر با شدت و ضعف مختلف در گذشته اتفاق افتاده و احتمال وقوع آن در آینده نیز وجود دارد. بیمه‌گر با استفاده از تجربه خسارتی گذشته می‌تواند بر اساس احتمالات، خطر را ارزیابی و حق بیمه را محاسبه کند. احتمالی بودن خطر مربوط به بیمه‌گذار است. بیمه‌گر با تکیه بر حساب احتمالات، نه تنها می‌تواند تعداد خسارت‌ها بلکه میزان خسارتی را که در یک دوره باید بپردازد برآورده کند. تنها چیزی که برای بیمه‌گر قابل پیش‌بینی نیست این است که اگر از میان جامعه بیمه‌گذاران قرار است به ۵۰ مورد خسارت پرداخت شود، مشخص نیست که این ۵۰ مورد چه کسانی هستند.**

➢ قانون اعداد بزرگ

پایه و اساس بیمه ، بر قانون اعداد بزرگ یا قانون متوسط ها استوار است و این قانون از تئوری احتمالات مشتق می گردد . اکثر مردم با تئوری احتمالات یا مفهوم احتمال آشنایی دارند و در محاوره اصطلاحات و عباراتی مانند اغلب، شاید ، ممکن است ، شانس، احتمال به کار می برند ، که مفاهیم معین و ساخته شده دارند. مانند شاید فردا باران ببارد ، احتمال دارد فلان اسب در مسابقه برنده شود و امثال اینها . هر یک از این عبارتها از پیشامدی خبر می‌دهند که وقوع آن برای گوینده محقق نیست.

اطمینان یا عدم اطمینان به وقوع امری ، خود درجات مختلف دارد، در روزی که آسمان صاف و آفتابی و آرام است عبارت احتمال بارندگی وجود دارد با بیان همین عبارت در روزی که آسمان ابری و هوا مرطوب و طوفانی هست به کلی فرق دارد . در هر دو روز بیان این مطلب که اطمینان قطعی به نزول باران را نشان نمی‌دهد اما روشن است که در روز اول انتظار رو به باران کمتر و در روز دوم بیشتر است . بیان درست برای روز اول این است که احتمال بارندگی وجود ندارد و برای روز دوم احتمال بارندگی وجود دارد .

مثال : احتمال وقوع شیر در پرتاب سکه ۱/۲ و احتمال آمدن ۵ در پرتاب تاس ۱/۶ است. این مثال هایی در مورد محاسبه احتمال وقوع قبل از مشاهده و تجربه است که به **احتمال تئوری یا احتمال نظری** معروف است، که در آن صورت کسر تعداد حالات مساعد یا موافق برای رخداد آن و در مخرج کسر تعداد حالات ممکن درج میگردد .

نتایج به دست آمده از پرتاب سکه و احتمال آمدن شیر

کل	شیر	درصد
۵۰	۲۹	۰/۵۸۰
۱۰۰	۵۵	۰/۵۵۰
۲۰۰	۱۰۹	۰/۵۴۵
۱۰۰۰	۴۹۲	۰/۴۹۲
۱۵۰۰	۷۴۹	۰/۴۱۹
۲۵۰۰	۱۲۴۷	۰/۴۹۹
۴۰۰۰	۱۹۹۹	۰/۴۹۹
۵۰۰۰	۲۴۸۵	۰/۴۹۸
۶۰۰۰	۲۹۸۸	۰/۴۹۸

از مجموعه گفتار فوق نتیجه می‌گیریم که هرگاه بخواهیم احتمال تجربی (اندازه‌گیری **و آزمایش**) وقوع یک حادثه تصادفی را حساب کنیم باید به مشاهده و آزمایش بپردازیم و تا می توانیم تعداد آزمایش را زیاد کنیم تا هر چه ممکن است به حقیقت نزدیکتر شویم . مفهوم و فوق در علم احتمالات به نام **قانون اعداد و بزرگ** نامیده می‌شود ، یعنی فراوانی نسبی هر نقطه نمونه‌ای به احتمال اختصاص داده شده به این نقطه میل می کند و پیش بینی ما برای آینده بیشتر به حقیقت مقرون می گردد .

نتایج مهم: ۱- تعداد دفعات آزمایش را هر چه بیشتر کنیم احتمال تجربی به احتمال تئوری نزدیکتر و در تعداد های بسیار زیاد آزمایش این دو احتمال برابر می‌شود ۲- تعداد دفعات هرچه بیشتر باشد پراکندگی نسبی رو به کاهش گذاشته تا این که عملاً صفر می‌شود ولی پراکندگی مطلق افزایش می‌یابد.

همانطور که گفته شد در بیمه اختلاف بین خسارت های احتمالی که بر اساس **احتمال ریاضی نظری** برآورد می‌شود با خسارت وارد شده واقعی **انحراف** نامیده می‌شود به عبارت دیگر تفاوت بین برآورد قبلی و خسارات واقعی بعدی انحراف آماری نام دارد .

انحراف با هم شکلی و تشابه بین خطرهای بیمه شده بستگی تام دارد . تشابه وقتی کامل است که از نظر میزان تعهد (جنبه کمی) و نوع خطر (جنبه کیفی) کلیه خطرهای بیمه شده با هم برابر بوده و تفاوتی نداشته باشد که کدام یک از خطرهای بیمه شده به وقوع بپیوندد.

بطوری که ملاحظه شد، نقش اصلی بیمه گر آن است که با استفاده از اصول و قواعد علم آمار و قانون اعداد بزرگ بر اساس روش تجربی، خطر های مشابه را در یک گروه قرار دهد .

به این ترتیب هنر و تخصص بیمه‌گر **جمع کردن خطر های هم شکل دور همدیگر** است تا از نظر احتمال بروز خسارت بین آنها تناوب یا فرکانس وجود داشته باشد و **بعلاوه** چنانچه گفتیم شرط اساسی دیگری که برای خطر های مرتبط بیمه قائل می شوند، متفرق بودن و

پراکنده بودن آنها است به این معنی که برای جلوگیری از آنکه حادثه ای به آسیب رساندن به تعداد زیادی از خطرها منتهی نگردد، بیمه گر باید بکوشد که حدود فعالیت خود را گسترش دهد و هیچگاه فعالیت خود را به یک رشته از فعالیت‌های بیمه‌ای به یک منطقه و به یک شهر محدود نکند و تا حدود امکان خطرهای بیمه شده متفرق و پراکنده باشد و به اصطلاح معروف "تمام تخم مرغها را در یک سبد قرارندهد"

به عبارت کلی بیمه گر باید حداکثر افرادی را که در معرض خطر یا خطر های مشخص می باشند دور هم جمع کند و تا آنجا که ممکن است بر دامنه عملیات خود بیفزاید به نحوی که هر قدر تعداد موارد بیمه شده زیادتر باشد با استفاده از قانون اعداد بزرگ ، خواه از نظر کیفی و خواه از نظر کمی به متوسط آماری در نزدیک شود و از انحراف ها بکاهد .

➢ ریسک Risk:

ریسک یعنی احتمال متحمل شدن زیان(احتمال خسارت)

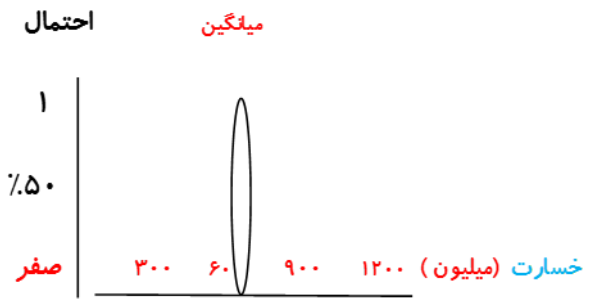
این تعریف شامل دو جنبه اصلی از ریسک است:

۱- مقدار زیان می بایست محتمل باشد؛

۲- عدم اطمینان در رابطه با آن زیان نیز می‌بایست وجود داشته باشد.

هر نوع فعالیتی با یک یا چند مخاطره مواجه پیدا می کند که موجب می شود نتیجه مورد انتظار به دست نیامده ، بنابراین هر عاملی که بتواند راه حصول نتیجه مورد نظر را سد کند ریسک نامیده می شود

از نقطه نظر آماری ریسک چیست؟ به احتمالات معلوم ریسک می گوییم. از نقطه نظر ریاضی ریسک به درجه پخش احتمال تصادف در اطراف نقطه مرکزی تعبیر می شود. پس هرچه درجه پخش احتمال تصادف در نقطه مرکزی بیشتر باشد درجه ریسک نیز بیشتر است



✓ مدیریت ریسک چیست؟ و مدیر ریسک وظایف خود را چگونه انجام می‌دهد؟

مدیریت ریسک مفهومی جز پیش‌گیری و کاهش خطر نیست. وظیفه مدیر ریسک در رسیدن به هدف‌های مؤسسه، مکمل وظایف اصلی مدیریت است تا آن موسسه بتواند با ظرفیت و توانائی‌های کامل فعالیت کند. به مفهوم دیگر، مدیریت ریسک با

ایجاد هماهنگی و فعالیت امر تولید را در زمینه پتانسیل ریسک‌هائی که سازمان با آن روبه‌رو است مورد توجه قرار می‌دهد.

مدیر ریسک وظایف خود را در چهار مرحله به شرح زیر انجام می‌دهد:

الف- شناسائی ریسک؛ ب- ارزیابی ریسک؛ پ- کنترل ریسک؛ و ت- تأمین منبع مالی برای برخورد با ریسک

➢ عدم اطمینان چیست:

به احتمالات نامعلوم عدم اطمینان می گوییم . بنابراین قابلیت حساب کمتری دارند . مانند کارهایی که در حوزه اقتصاد و بورس و سیاست انجام می گیرد . تغییر رفتار مصرف کننده یک عدم اطمینان است.

ریسک و عدم اطمینان دو مفهوم مکمل و اجزای تشکیل دهنده فرآیند تصمیم گیری در سازمان ها هستند و معمولاً عدم اطمینان در تعریف ریسک گنجانده می شود.

عدم اطمینان را می توانیم تابعی از احتمال نیز بدانیم . میزان احتمال وقوع یا عدم وقوع حادثه ای ما بین صفر و یک است زمانی که احتمال صفر است عدم اطمینان هم صفر است و به تدریج با افزایش احتمال عدم اطمینان نیز افزایش یافته و نهایتاً به یک دوم (۵۰٪) خواهد رسید و سپس به تدریج با کا هش از یک دوم ، عدم اطمینان نیز کاهش یافته و در نقطه صفر عدم اطمینان از بین میرود. همچنین احتمال و ریسک نیز دارای مفاهیم یکسان نیستند زیرا زمانیکه احتمال برابر یک است اصولاً ریسک برابر صفر است هر چند در تعریف ریسک از مفهوم توزیع احتمال استفاده می شود.

عدم اطمینان نقطه مقابل اطمینان است و به معنی ابهام و بد بینی ما به قضایا است. عدم اطمینان امری ذهنی توسط میزان باور و عقاید اندازه گیری می‌شود. احتمال سقوط ارزش دلار در برابر نفت . کاهش ارزش دلار در برابر یورو بر اثر یک قرارداد بزرگ نفتی

ولی در مورد زیر احتمالات قابل محاسبه است. مانند ریسک مرگ میر بیمه شدگان در اثر سرطان که مثلا با عدد ۳۰ درصد قابل نشان دادن است.

عدم اطمینان دارای طیف گسترده تصمیم گیری است که ناشی از عوامل ۱- ماهیت و شرایط تصمیم گیری ۲- نوسانات محیط ۳- تعداد اهدافی که باید بین آنها تعادل برقرار شود ۴ - نتایج آتی تصمیم ۵ - کیفیت واحد تصمیم گیرنده .

➢حق بیمه

حق‌بیمه مابه ازای خدمات و پوشش بیمه ای است و باید توسط بیمه گذار پرداخت شود. بیمه عقدی معوض است و حق بیمه عوض تعهدات بیمه گر به شمار می رود. همین امر منشأ ایجاد احکامی در عقد بیمه است.

از جمله چنانچه تعهدات بیمه گر به علتی منتفی شود، تعهد بیمه گذار به پرداخت بیمه از بین خواهد رفت و چنانچه بیمه را پرداخته باشد به نسبت مدت باقی مانده، بیمه باید به او عودت شود. مصداق بارز این امر در موردی است که مال موضوع بیمه بعلت غیر از خطرات تحت پوشش تلف شود. در این حالت به علت منتفی شدن عقد بیمه، بیمه گر مکلف به بازگرداندن مبلغی است که برای پوشش بیمه برای مدت باقی مانده عقد دریافت نموده است. البته چنانچه موضوع بیمه به علت حادثه ای که تحت پوشش بیمه قرار داشته و با وقوع آن تعهد بیمه گر به جبران خسارت ایجاد شده باشد، موجبی برای عودت حق بیمه وجود ندارد.

حق بیمه تولیدی

منظور مبالغ حق‌بیمه تولیدی بیمه‌نامه‌هایی است که طی دوره گزارش توسط مؤسسات بیمه صادر شده‌اند. این مبالغ در بیمه های غیر زندگی بر مبنای بیمه صادره و در بیمه های زندگی بر مبنای حق بیمه های وصولی (دریافتی) لحاظ می گردد.

حق بیمه سرانه

حق‌بیمه سرانه از تقسیم حق‌بیمه تولیدی صنعت بیمه بر جمعیت هر کشور به‌دست می‌آید و بیانگر مقدار بیمه‌ای است که به طور متوسط هر نفر پرداخت کرده است.

حق بیمه مستقیم

حق بیمه مربوط به بیمه‌نامه‌های صادره که مستقیماً توسط شرکت بیمه یا شعب مربوطه دریافت شده باشد.

بیمه غیر مستقیم

حق بیمه تولید شده مربوط به بیمه‌نامه‌هایی است که توسط نمایندگان یا کارگذاران عرضه شده است.

ذخیره حق بیمه

عبارت است از حق بیمه‌های مربوط به فاصله زمانی بین تاریخ ترازنامه تا انقضای مدت قراردادهای بیمه ای.

حق بیمه عاید شده(درآمد حق بیمه)

حق بیمه عاید شده حاصل جمع حق بیمه صادره و تفاضل ذخیره حق بیمه ابتدا و انتهای سال است و بیانگر آن است که چه میزان از حق بیمه‌های صادره به دلیل انقضای مدت بیمه نامه نصیب شرکت بیمه شده است .

بر این اساس عوارض قانونی (سهم پرداختی به نیروی انتظامی، وزارت بهداشت و صندوق تامین خسارت های بدنی) از حق بیمه عاید شده رشته های شخص ثالث و مازاد و حوادث راننده، حسب مورد کسر گردیده است.

➢ فرمول محاسبه حق بیمه



www.asia26090.ir

با توجه به آنچه که در ابتدای فصل درباره عوامل موثر در تعیین حق بیمه گفته شد ، حق بیمه از دو عامل مجزا ترکیب می شود :

۱) حق بیمه خالص خطر ۲) هزینه های عمومی بیمه گر و سود او

۱) حق بیمه خالص خطر : این قسمت حق بیمه بر اساس محاسبات فنی و مبتنی بر احتمالات برآورد می شود و در واقع احتمال به وقوع پیوستن حادثه مورد تعهد بیمه گر است . مجموع حق بیمه خالص خطر به بیمه گر امکان پرداخت خسارت احتمالی را می دهد .

در تنظیم تعرفه و تعیین حق بیمه بایستی فرض شود که جمع خسارتهایی که واقع می شوند(S)برابر جمع حق بیمه ها(P) باشد . p = s **کل خسارت = کل حق بیمه**

اگر فرمول محاسبه حق بیمه خالص خطر بشرح ذیل باشد :

(تعداد کل موارد بیمه / تعداد خسارت) * میانگین خسارت = حق بیمه خالص خطر

حق بیمه‌ای که به این ترتیب تعیین می شود در اصطلاح بیمه گران به **حق بیمه تئوری یا حق بیمه خالص خطر** معروف است یعنی حق بیمه ای که به فرض تمام پیش‌بینی‌ها و بدون در نظر گرفتن هزینه های بیمه گری بایستی از بیمه گذار دریافت شود .

ممکن است تصور شود که اگر احتمال وقوع خطر یک در هزار باشد باید حق بیمه را هم در حدود یک در هزار ارزش مورد بیمه تعیین کرد ولی تصور در صورتی درست است که در هر بار که خطر واقع می شود تمام موضوع بیمه از بین برود . مثلاً اگر موضوع بیمه خانه مسکونی است ، خسارت تمام خانه هائی که دستخوش حریق می‌شوند ۱۰۰٪ و برابر ارزش کامل آنها باشد خسارت کلی واقع شود حال آنکه معمولاً این طور نیست و خسارت کلی به ندرت واقع می شود

فرمول محاسبه حق بیمه فنی یا حق بیمه تئوری یا حق بیمه خالص خطر را با یک مثال دنبال می کنیم :

- فرض کنید آمار نشان می‌دهد که احتمال آتش سوزی از هر ۱۰،۰۰۰ واحد مسکونی ۱۲ مورد آن است، نرخ خطر در این صورت ۱۲/۱۰۰۰۰ و برای خانه ای که ۱،۰۰۰،۰۰۰ ریال ارزش دارد حق بیمه معادل ۱۲۰۰ ریال است

حق بیمه ناخالص ۱۲۰۰ = ۱۲/۱۰۰۰۰ * ۱،۰۰۰،۰۰۰

از طرف دیگر خسارات ناشی از هر حریقی جز در موارد استثنایی کلی نیست ، بلکه اغلب خسارت جزئی و بعضا" کلی می باشد. لذا باید متوسط هر خسارت را نیز با استفاده از آمار و تجربیات گذشته به دست آورد.

فرض کنیم به طور متوسط ۱۲ واحد براساس محاسبات زیر دچار خسارت شده است:

۵۰۰۰۰+۲۰۰۰۰۰+۱۲۰۰۰۰ +۲۵۰۰۰۰+۲۰۰۰۰۰+۱۵۰۰۰۰+۲۰۰۰۰۰+۳۰۰۰۰۰+۱۰۰۰۰۰+۱۰۰۰۰۰۰ +۱۸۰۰۰۰

+۲۵۰۰۰۰= ۳،۰۰۰،۰۰۰ جمع خسارت

متوسط خسارت (مبلغ) ۳،۰۰۰،۰۰۰ / ۱۲ = ۲۵۰۰۰۰

اجزای فرمول حق بیمه خالص خطررا بکار می گیریم:

(تعداد کل موارد بیمه / تعداد خسارت) * میانگین خسارت = حق بیمه خالص خطر

حق بیمه خالص خطر = 300 = 250000 . (12 / 10000) = حق بیمه خالص خطر

❖ روش تعیین نرخ خالص خطر چیست؟ =

متوسط خسارت (درصد) = مبلغ متوسط ارزش کل موارد بیمه / مبلغ متوسط خسارات

متوسط خسارت (درصد) ۲۵٪ = ۱،۰۰۰،۰۰۰ / ۲۵۰،۰۰۰

نرخ خالص خطر ۰/۰۰۰۳ = ۰/۰۰۱۲ * ۲۵٪

مجددا" حق بیمه را بر اساس نرخ خالص خطر محاسبه کنید:

حق بیمه = نرخ * ارزش یک واحد ۳۰۰ = ۰/۰۰۰۳ * ۱،۰۰۰،۰۰۰

مسئله ۱: آمار نشان می‌دهد که از هر ۱۰۰۰ دستگاه اتومبیل تعداد ۲۰ دستگاه دچار تصادف می شوند و به طور میانگین ۲۰ درصد خسارت می‌بینند مطلوب است محاسبه نرخ خالص خطر:

نرخ خطر = میانگین خسارت * (تعداد خسارت / تعداد کل موارد بیمه)

حق بیمه خالص خطر = 20 /100 * (20/1000) = 400/100000=4/1000= 0/004

مسئله ۲ : شرکت بیمه‌ای برای ۱۰۰۰۰ نفر بیمه حوادث شخصی صادر کرده است وآمار خسارت در پنج سال گذشته به صورت زیر بوده است:

سال	اول	دوم	سوم	چهارم	پنجم
تعداد خسارت	۳۰	۶۰	۵۰	۵۵	۴۵

میانگین خسارت مصدومین و فوتی در این سال ها ۵۰۰،۰۰۰،۰۰۰ ریال بوده است.

مطلوب است محاسبه حق بیمه خالص خطر : ۲۴۰/۵ = ۴۸ ۳۰ + ۶۰ + ۵۰+۵۵+۴۵ = ۲۴۰

پاسخ نهائی ۲،۴۰۰،۰۰۰ = (۴۸/ ۱۰،۰۰۰) * ۵۰۰،۰۰۰،۰۰۰

مسئله ۳ : آمار بیمه بدنه موتور سیکلت نشان می دهد که از هر ۵۰،۰۰۰ دستگاه موتورسیکلت تعداد ۵،۰۰۰ دستگاه دچار تصادف می شود و میانگین خسارت هر دستگاه موتور سیکلت ۲۰٪ ارزش آن می باشد .

الف)مطلوب است محاسبه نرخ خالص خطر ؟ پاسخ نهائی ۲ ٪ = (۵۰۰۰ / ۵۰،۰۰۰)*۲۰٪

ب) چنانچه ارزش هر دستگاه موتورسیکلت ۲۰ میلیون ریال باشد مبلغ حق بیمه چقدر است ؟

www.asia26090.ir

پاسخ نهائی ۴۰۰،۰۰۰ = ۲ ٪ * ۲۰،۰۰۰،۰۰۰

۲) سایر اجزا حق بیمه تجاری: شامل ضریب احتیاط ، هزینه های عمومی بیمه گر و سود تجاری (E).

الف) ضریب احتیاط : (ذخیره فنی تکمیلی و خطرات طبیعی)

در مبحث حق بیمه دریافتیم که تعرفه حق بیمه بر مبنای آمار گذشته و این فرض تنظیم می شود که واقعیت های آینده مشابه پیش بینی های مبتنی بر آمار موجود خواهد بود این فرض همیشه با واقعیت تطابق ندارد چه بسا یافته خساراتی دوره بسیار بیشتر یا کمتر از آن باشد که بر مبنای آمار و تجربیات گذشته پیش بینی شده است(**خطا = واقعیت – پیش بینی**). به خصوص در بعضی از بیمه نامه ها از قبیل آتش سوزی، باربری و مقاطعه کاری که حوادث طبیعی در ایجاد خسارات نقش عمده ای دارد . در این حوادث تغییرات جوی دوره ای موثرند . طوفان ها ، سیل ها خشکسالی‌ها تگرگ یخبندان ها دوره های تحولی دارند که بسیار بیشتر از یک سال است بنابراین ممکن است یک سال و حتی چند سال متوالی خساراتی بیش از آنچه آمار حکم میکند مواجه شویم از این رو پیش بینی ذخیره تکمیلی و حوادث طبیعی برای مقابله با این احتمالات لازم است.

ب) هزینه های بیمه گری و سود:

ا – هزینه های بیمه گری شامل حقوق و مزایای کارکنان، استهلاک ، اجاره‌بها هزینه آب و برق و تلفن، وسایل نقلیه ، ملزومات اداری ،تجهیزات ، مالیات عوارض و ...

۲- کارمزد واسطه‌ای بیمه یا هزینه تجاری که بیمه‌گر برای فروش بیمه به واسطه های بیمه می‌پردازد . نسبت این کارمزد در دوره‌های مختلف متفاوت است

۳- سود مورد انتظار: قیمت‌گذاری بیمه‌نامه همانند قیمت‌گذاری سایر کالاها است هر مورد فروشنده باید قیمتی وضع کند که علاوه بر پوشش هزینه ها ، از آنها سود متعارف نیز کسب نماید.

حال اگر جمع کارمزد هزینه های اداری و عمومی بیمه گر ضریب احتیاط و سود عملیات بیمه‌گر را حدود ۴۰ درصد نرخ یا حق بیمه خالص خطر در نظر داریم با به دست آمدن نرخ محاسبه حق‌بیمه که مبتنی بر حساب احتمالات و هزینه‌های دیگر بیمه‌گری برآورد شده، دقت در تعیین حق بیمه بیشتر و خطای پیش‌بینی کاهش می‌یابد.

بنابر این با افزودن هزینه های بیمه گری و سود و ضریب احتیاط ، محاسبه حق بیمه با فرمول زیر بدست می آید:

هزینه های بیمه گری + (تعداد کل موارد بیمه / تعداد خسارت) * میانگین خسارت = حق بیمه نهائی

۳۰۰ ٪ ۶۰

X ٪ ۱۰۰

تعدیل حق بیمه ۳۰۰ / ٪۶۰ = ۵۰۰

تعدیل نرخ نهائی ۰/۰۰۰۵ = ٪۶۰ / ۰/۰۰۰۳

و یا سرمایه بیمه * نرخ = حق بیمه خالص ۵۰۰ = ۰/۰۰۰۵ * ۱،۰۰۰،۰۰۰

***

ضمائم : آشنائی با برخی سوالات شاخص های مرکزی و پراکندگی

۱. شاخص های پراکندگی

- برای توصیف توزیع به کار برده می شود
- ویژگی عمومی توزیع نمره هاست

۲. لزوم استفاده از پراکندگی

- ممکن است توزیع هایی وجود داشته باشد که میانگین های مساوی داشته باشند اما پراکندگی آن ها در اطراف میانگین اختلاف داشته باشد

۳. شاخص های پراکندگی چه چیز را نشان می دهند؟

- میزان پراکندگی یا تغییراتی را که در بین نمره ها وجود دارد نشان می دهند

۴. شاخص های پراکندگی را نام ببرید

دامنه تغییرات

انحراف چارکی

واریانس

انحراف استاندارد

۵. دو دسته از روش های آماری که برای توصیف دادهها لازمند:

شاخص های پراکندگی و شاخص های مرکزی

۶. وجود یا نبود پراکندگی خوب است یا بد؟ مثال قضاوت در این مورد باید بر اساس هدف پژوهش انجام بگیرد

مثلا : تعیین میزان معلومات یک گروه: پراکندگی زیاد خوبه

سطح یادگیری پس از تدریس یا آموزش: پراکندگی کم خوبه

۷. دامنه تغییرات / ساده ترین شاخص پراکندگی:

مقدار آن در یک توزیع فراوانی برابر است با تفاضل بین بزرگترین و کوچکترین عدد

۸. چرا دامنه ی تغییرات را به علاوه ۱ می کنند؟

- برای در نظر گرفتن حد پایین واقعی کوچکترین عدد و حد بالای واقعی بزرگترین عدد

۹. اشکالات دامنه تغییرات

- توصیف نمره ها به صورت حقیقی * چون فقط تفاضل بین کوچکترین و بزرگترین رو نشون می ده

یک شاخص پایدار پراکندگی نیست زیرا مقدار آن با تغییر یک نمره (کوچکترین یا بزرگترین) تغییر می کند

۱۰. برای محاسبه ی دامنه تغییرات استفاده از چه مقیاسی لازم است؟

مقیاس فاصله ای اختلاف بین نمره ها دارای تفسیر منطقی باشد

۱۱. انحراف چارکی

میزان پراکندگی را در اطراف مرکز توزیع نمره ها نشان می دهد

پس : از دامنه تغییرات با ثبات تر است

نصف فاصله ی بین چارک های اول و سوم

۱۲. چارک چیه؟

- نقاطی بر روی مقیاس اندازه گیری که کلیه مشاهده ها یا نمره ها را به ۴ قسمت مساوی تقسیم می کنند

۱۳. چارک اول

Q1 نقطه ای که ۲۵ درصد نمره ها را از پایین جدا می کند به آن نقطه ۲۵ درصد هم گفته می شود

۱۴. چارک دوم

Q2

میانه : توزیع را به دو قسمت مساوی تقسیم می کند

۱۵. چارک سوم

Q3

نقطه ای که سه چهارم نمره ها در زیر آن و بقیه در بالای آن واقع شده اند . به آن نقطه ۷۵ درصد هم گفته می شود

۱۶. دامنه تغییرات بین چارک ها

فاصله بین چارک اول و سوم

۱۷. تفاوت انحراف چارکی با دامنه تغییرات،

ثبات بیشتری دارد

خیلی کم تحت تاثیر اختلاف های تصادفی بین نمونه های انتخاب شده قرار می گیرد

۱۸. در چه شرایطی باید از انحراف چارکی استفاده کرد؟

- در شرایطی که تاثیر نمره های خیلی بزرگ یاخیلی کوچک مد نظر نباشد مثل شرایط استفاده از میانه

۱۹. شباهت های میانه و انحراف چارکی

- هر دو یک خانواده آماری را تشکیل می دهند
- اساس محاسبات آن ها نقاط چارکی است
- تحت تاثیر نمره های خیلی کوچک یا خیلی بزرگ قرار نمی گیرند

۲۰. مقیاس لازم برای محاسبه انحراف چارکی

مقیاس فاصله ای

۲۱. انحراف

فاصله بین هر عدد از یکی از شاخص های مرکزی

۲۲. انحراف متوسط

میانگین قدر مطلق انحرافات از میانگین پراکندگی موجود در توزیع نمره ها را نشان می دهد

۲۳. رابطه انحراف از میانگین و پراکندگی

- اگر پراکندگی نمره ها افزایش یابد، انحراف نمره ها از میانگین نیز زیاد می شود
- مقیاس لازم برای محاسبه انحراف متوسط فاصله ای

۲۴. اشکال انحراف متوسط

- مقدار آن از توزیعی به توزیع دیگر به صورت چشمگیری دستخوش تغییر می شود
- نمی توان عملیات جبری انجام داد

۲۵. واریانس

- استفاده از ارزش های عددی کلیه نمره ها
- میانگین انحراف نمره ها از میانگین
- یا مجموع مجذور انحراف نمره ها از میانگین تقسیم بر تعداد نمره ها

۲۶. در محاسبه انحراف متوسط و انحراف چارکی چه چیز بررسی نمی شود؟

- انحراف متوسط: علائم اعداد
- انحراف چارکی: کلیه ارزش های مقداری تمام اعداد

۲۷. در جدول های طبقه بندی شده میانگین را چگونه می یابیم؟

- معمولا طبقه ای انتخاب می شود که در وسط طبقه ها قرار داشته باشد
- یا دارای بیشترین فراوانی باشد

۲۸. محدودیت واریانس

- برای امار استنباطی کاربرد فراوانی دارد

- اما در آمار توصیفی محدود است
- زیرا با مجذور کردن انحراف نمره ها از میانگین، واحد واریانس یا واحد اندازه گیری تغییر پیدا می کند

۲۹. انحراف استاندارد (معیار) چگونه محاسبه می شود و چه مشکلی را حل می کند

با جذر گرفتن از واریانس – حل مشکل واحد اندازه گیری

۳۰. هدف از گرفتن انحراف استاندارد

- به منظور تعیین تغییرات یا پراکندگی توزیع نمره ها به کار برده میشود
- به تناسب تغییر در پراکندگی نمره ها تغییر می کند

۳۱. انحراف استاندارد با تغییر چه چیزی و به چه صورت تغییر میکند؟

- اگر توزیع نمره ها نسبت به میانگین پراکندگی زیادی داشته باشد، انحراف استاندارد بزرگتر خواهد شد
- اگر همه نمره ها در اطراف میانگین انباشته یا جمع شده باشد انحراف استاندارد کوچکتر می شود

۳۲. در چه شرایطی انحراف استاندارد باید بزرگ یا کوچک باشد؟

- در این زمینه عناوین بزرگ و کوچک وقتی معنا دارند که قصد
- مقایسه ی انحراف استانداردهای مختلف را با یکدیگر داشته باشیم
- بزرگ یا کوچک بودن انحراف استاندارد نسبی است

۳۳. آسان ترین روش محاسبه انحراف استاندارد

- با استفاده از نمره خام

۳۴. چرا در نمونه تعداد را از ۱ کم می کنیم؟

- بتوان از واریانس نمونه برآورد دقیقی از واریانس جامعه به دست آورد

۳۵. تاثیر جمع کردن در انحراف استاندارد

- به همه اعداد یک توزیع یک عدد ثابت اضافه شود انحراف استاندارد تغییر نخاهد کرد

۳۶. تاثیر ضرب کردن در انحراف استاندارد

- همه اعداد یک توزیع در یک عدد ثابت ضرب شود انحراف استاندارد نیز در همان عدد ضرب خواهد شد

۳۷. تاثیر ضرب کردن در واریانس

- همه اعداد یک توزیع در یک عدد ثابت ضرب شود واریانس در مجذور همان عدد ضرب خواهد شد

۳۸. خطای طبقه بندی

- وقتی میانگین نمره ها را از طریق جدول فراوانی تعیین می کنیم با وقتی که میانگین را از روی خود نمره ها تعیین می کنیم، فرق دارد. در میانگین این خطا ناچیز است
- اما در انحراف استاندارد به خاطر مجذورش دن انحراف ها، این خطا بیشتره

۳۹. میزان خطا در انحراف استاندارد به چه چیزهایی بستگی داره

- فاصله طبقه ها – هر چی بیشتر باشه، خطا زیادتره
- تعداد طبقه ها – کمتر از ۱۲ باشه

۴۰. تصحیح شپرد

- راهی برای تصحیح خطای طبقه بندی

۴۱. انحراف استاندارد مرکب

- میانگین یا انحراف استاندارد مرکب دو یا چند انحراف استاندارد را محاسبه می کند
- فرمول مک نیمار

۴۲. ضریب پراکندگی

مقایسه پراکندگی یک ویژگی با پراکندگی ویژگی دیگر از همان گروه

۴۳. نام دیگر ضریب پراکندگی

- ضریب نسبی واریانس

۴۴. مقیاس مورد نیاز برای انحراف استاندارد

- فاصله ای

۴۵. معتبرترین شاخص پراکندگی؟چرا؟

- انحراف استاندارد
- چون معرف جامعه (مخصوصا اگر تصادفی باشد) همه ارزش های مقداری کلیه اعداد استفاده می شود

۴۶. انحراف استاندارد در کدام آمار بیشتر کاربرد دارد؟

- استنباطی

چون امکان انجام عملیات ریاضی دارد

۴۷. شاخص های پراکندگی به ترتیب ثبات

- انحراف استاندارد
- انحراف چارکی
- دامنه تغییرات

۴۸. اشکالات دامنه تغییرات

- از نمونه ای به نمونه دیگر فرق می کند – به درد ارزیابی جامعه نمی خورد
- به حجم نمونه بستگی دارد

۴۹. بیشترین کاربرد انحراف چارکی در کجا. مثال

- در جاهایی که میانه مناسب ترین شاخص مرکزی توزیع نمره ها باشد می توان از انحراف چارکی استفاده کرد

مثال : نمودار نمره ها دارای کجی (مثبت یا منفی) باشد

۵۰. بیشترین کاربرد انحراف استاندارد در کجا

جایی که میانگین به عنوان شاخص مرکزی استافده می شود

- در مورد نمودار با کجی خیلی زیاد باید احتیاط شود چون نمره های خیلی بزرگ یا خیلی کوچک در مقدار انحراف استاندارد تاثیر دارند

۵۱. گشتاور

- دسته ای از شاخص های آمار توصیفی
- میانگین وانحراف استاندارد به آن تعلق دارد
- اگر متغیری را چندین بار اندازه گیری کنیم، انحراف میانگین هر اندازه گیری را با میانگین گشتاور می نامند

۵۲. کاربرد گشتاور های ۳ و ۴

رتبه سوم محاسبه اندازه کجی

رتبه چهارم برا ی محاسبه کشیدگی

۵۳. مورد استفاده از چارک ها

تعیین کجی توزیع نمره ها

۵۴. منحنی متقارن

- کجی صفر
- اختلاف چارک سوم و چارک دوم برابر است با اختلاف چارک دوم و اول
- مکعب مجذور انحراف نمره ها از میانگین صفر است

۵۵. کجی مثبت

- اختلاف چارک سوم و چارک دوم بزرگتر است از اختلاف چارک دو م و اول
- مکعب مجذور انحراف نمره ها از میانگین مثبت است است

۵۶. کجی منفی

- اختلاف چارک سوم و چارک دوم کوچکتر است از اختلاف چارک دو م و اول

- **مکعب مجذور انحراف نمره ها از میانگین منفی است**

۵۷. منحنی طبیعی

ضریب کشیدگی صفر است

۵۸. منحنی با برآمدگی در اوج

ضریب کشیدگی مثبت است

۵۹. منحنی با خوابیدگی در اوج

ضریب کشیدگی منفی است

**** موفق باشید ****

www.asia26090.ir