آمار توصیفی

قسمتی از روش‌های آماری که تنها به توصیف و تجزیه و تحلیل گروه معینی، بدون تعمیم نتایج حاصله به گروه بزرگتر از آن محدود می‌گردد، آمار توصیفی نامیده می‌شود.

جامعهٔ آماری: تعدادی از عناصر جامعه که حداقل دارای یک صفت مشخصه باشند، جامعه آماری را تشکیل می‌دهند.

صفت مشخصه: صفتی است که بین همه عناصر جامعهٔ آماری مشترک و متمایز کننده آن جامعهٔ آماری از سایر جوامع می‌باشد.

جامعه آماری به دو دسته تقسیم می‌شود:

۱) **محدود:** تعداد افراد جامعه محدود است.

۲) **نامحدود:** تعداد افراد جامعه نامحدود است.

نمونه: به هر بخش از جامعه آماری محدود یا نامحدود یک نمونه گفته می‌شود.

تعریف دقیق نمونه: به تعداد محدودی از اعضای جامعه آماری که بیان کننده تمام ویژگی‌های جامعه اصلی باشند، نمونه گویند.

نظریه نمونه‌ها و توزیع نمونه‌ای:

نمونه تصادفی: نمونه تصادفی به نمونه‌ای گویند که افراد تشکیل‌دهنده آن دارای خصوصیات افراد جامعه اصلی باشند.

نکته: تنها در انتخاب تصادفی، تمام افراد جامعه از شانس مساوی برای انتخاب شدن برخوردار هستند.

نکته مهم: اندازه‌گیری جامعه، برای بدست آوردن برخی از شاخص‌هاست.

۱) اگر اندازه‌گیری از کل جامعه باشد، شاخص بدست آمده پارامتر جامعه است.

۲) اگر اندازه‌گیری از نمونه‌هائی از جامعه باشد، به شاخص بدست آمده آماره می‌گویند.

آماره و پارامتر:

آماره: اصطلاحی است که در مورد نمونه استفاده می‌شود و خصوصیتی از آن را بررسی می‌کند. مانند: میانگین نمونه $(\overline{x})$، واریانس نمونه (S^2)،نسبت نمونه $(\overline{p})$.

نکته مهم: هر آماره یک متغیر تصادفی است، چرا که از یک نمونه به نمونه دیگر تغییر می‌کند.

پارامتر: عددی است که خصوصیتی از یک جامعه را بیان می‌کند مانند: میانگین جامعه (μ)، واریانس جامعه (δ^2) ، میانه (Md).

نکته مهم: پارامترها در جامعه ثابت هستند ولی مجهول و باید آن‌ها را از طریق آماره‌ها در نمونه‌گیری تخمین بزنیم.

مثال: $\overline{x}$ به عنوان برآوردکننده‌ای از μ:

۱) یک متغیر تصادفی است اگر که μ نیز تصادفی باشد

۲) یک متغیر تصادفی است در حالی‌که μ یک ثابت است

۳) یک کمیت ثابت است در حالی‌که μ متغیر باشد

۴) یک کمیت ثابت است و μ نیز ثابت است

حل:

گزینه (۲) صحیح می‌باشد.

پارامترها و آماره‌های مهم:

شاخص	گروه	نماد کلی	میانگین	واریانس	نسبت
آماره	نمونه	$\hat{\theta}$	$\bar{x}$	S^2	P
پارامتر	جامعه	θ	μ	δ^2	π

تفاوت در نوع و همچنین کاربردهای شاخص آماره و پارامتر، موجب تقسیم‌بندی آمار به آمار توصیفی و آمار استنباطی (استنتاجی) شده است.

مهم: سیر تحول آمار دارای ۳ بخش می‌باشد:

۱) آمار توصیفی: این آمار به توصیف جامعه می‌پردازد و هدف آن محاسبه پارامترهای جامعه است. چنانچه محاسبه مقادیر و شاخص‌های آماری برای جامعه از طریق سرشماری تمامی عناصر انجام گیرد، به آن آمار توصیفی گفته می‌شود.

۲) آمار استنتاجی (استنباطی): به قسمتی از آمار که می‌تواند نتایج حاصل از تجزیه و تحلیل نمونه را به جامعه تعمیم دهد، آمار استنتاجی گفته می‌شود. به عبارت دیگر آماره‌ها از طریق نمونه‌گیری بدست می‌آیند، سپس به کمک تخمین (برآورد) و آزمون فرض، به پارامترهای جامعه تعمیم داده می‌شوند.

۳) آمار ناپارامتریک: در مقابل آمار پارامتریک قرار دارد.

ـ در آمار پارامتریک فرض اساسی برخوردار بودن مشاهدات از توزیع نرمال است.

ـ در آمار ناپارامتریک، فرض فوق وجود نداشته و بیشتر متغیرها با مقیاس کیفی سنجیده می‌شوند و آزاد از توزیع هستند. در حقیقت در این آمار به هیچ توزیعی وابستگی وجود ندارد.

مثال: کدام تعریف برای صفت مشخصه صحیح است؟

۱) صفتی است که اندازه‌گیری آن از فردی به فرد دیگر تغییر می‌کند
۲) صفتی است که قابل اندازه‌گیری است
۳) صفت مشترک برای افراد جامعه است
۴) صفتی که قابل شمارش باشد

حل:

گزینه (۳) صحیح می‌باشد.

مثال: کدام دسته از فنون آماری زیر بر فرض آزاد از توزیع بنا شده‌اند ؟

۱) پارامتریک ۲) ناپارامتریک ۳) توصیفی ۴) استنباطی

حل:

گزینه (۲) صحیح می‌باشد.

صفت: کمیت یا کیفیتی است که متعلق به عناصر جامعه آماری بوده و همواره به دو بخش تقسیم می‌شود:

۱) صفت مشترک (ثابت): صفتی است که میان افراد جامعه به صورت مشترک وجود دارد. مانند صفت دانش‌آموز بودن برای دانش‌آموزان یک کشور.

۲) صفت متغیر: خاصیتی است که افراد یک جامعه را از یکدیگر متفاوت، جدا و مشخص می‌سازد. مانند صفات: قد، سن، وزن، ... .

نکته: صفات متغیر به دو دسته کمی و کیفی تقسیم می‌شوند.

صفات متغیر کمی
۱) پیوسته: قد ـ وزن
۲) گسسته: تعداد اعضاء خانواده

صفات متغیر کیفی | رشته تحصیلی ـ رنگ پوست ـ کیفیت و مرغوبیت کالا

نکته ۱: در صفات کمی امکان اندازه‌گیری و بیان یک عدد واحددار مانند: کیلومتر ـ کیلو و ... وجود دارد.

نکته ۲: در صفات کیفی امکان اندازه‌گیری با ابزارهای اندازه‌گیری وجود نداشته و نمی‌توان آن را به صورت عددی واحددار بیان نمود.

نمونه‌گیری:

در بسیاری از موارد پژوهشگران به دنبال تعیین پارامترهای جامعه (مانند: میانگین جامعه (μ) ، واریانس جامعه (δ^2)) هستند. البته به‌طور طبیعی این عمل امکان‌پذیر نیست. به همین دلیل با استفاده از نمونه‌گیری به استنباط پارامترهای جامعه آماری می‌پردازند.

نکته مهم: اگر پارامتر را **شاخص** بدست آمده از طریق نمونه‌گیری بنامیم، به این شاخص در یک نمونه n تائی **آماره** می گوئیم. گفتنی است آماره از یک نمونه به نمونه دیگر تغییر می کند. به همین علت برای رسیدن به یک پایائی و reliability باید به یک تقریب برای **توزیع نمونه‌گیری** آماره برسیم.

توزیع آماره:

تابع احتمالی است که از نمونه‌گیری مکرر حاصل می‌شود. در شکل کامل‌تر به آن، توزیع نمونه‌گیری آماره (Statistic sampling distribution : (SSD)) می گوئیم.

دلایل نمونه‌گیری:

۱ـ هزینه
۲ـ به روز بودن
۳ـ درستی و صحت
۴ـ صرفه‌جوئی در زمان
۵ـ آزمون تخریب کننده

نکته مهم: انواع روش‌های نمونه‌گیری:

۱ـ تصادفی ساده
الف) قرعه‌کشی
ب) جدول اعداد تصادفی

۲ـ نمونه‌گیری منظم (Systematic)

۳ـ نمونه‌گیری گروهی (Stratified Sampling)

۴ـ نمونه‌گیری خوشه‌ای (Cluster sampling)

۵ـ نمونه‌گیری مرحله‌ای (Stage Sampling)
دقت کنید، صرف‌نظر از این‌که چه روش آماری برای استنباط آماری موردنظر است، قدرت آن به روش بکار رفته برای انتخاب نمونه بستگی دارد. در صورتی که نمونه موردنظر نماینده واقعی جامعه نباشد، نمونه موردنظر دارای **اریب** (bias) است. در این حالت پیش‌بینی صحیح و دقیق درباره پارامترهای جامعه امکان نخواهد داشت.

۱ـ نمونه‌گیری تصادفی ساده:
در این حالت هر یک از عناصر جامعه برای انتخاب شدن شانس مساوی دارند (هم تراز هستند). در این حالت افراد یا اشیاء به طور تصادفی از لیست تهیه شده از جامعه انتخاب می‌شوند و باید دارای ویژگی‌هائی همانند ویژگی‌های همان جامعه‌ای که از آن انتخاب می‌شوند، باشند. که این راه به دو روش : قرعه‌کشی و جدول اعداد تصادفی انجام می‌شود.

۲ـ نمونه‌گیری منظم:
در این روش، شکل تغییر یافته حالت تصادفی ساده به کار گرفته می‌شود. در این روش یک نقطه از فهرست افراد یا اشیاء جامعه را به طور تصادفی انتخاب می‌کنیم و بعد از آن نمونه موردنظر را به صورت منظم پشت سر هم انتخاب می‌کنیم. این روش برای آن دسته از جوامع آماری که کد از پیش تعیین شده و مرتبی دارند، کاربرد فراوان دارد (شماره کارمندی ـ شماره دانشجوئی) با مشخص شدن اولین عضو بقیه اعضا نیز مشخص می‌شوند این خاصیت از یک سو یک حسن است اما چون شانس را از بقیه اعضا می‌گیرد عیب محسوب می‌شود.

۳ـ نمونه‌گیری گروهی:
در این روش جامعه را به گروه‌های متجانس تقسیم و هر گروه دارای ویژگی‌های مشابهی هستند. پس از تقسیم جامعه به گروه‌های متجانس از هر گروه نمونه موردنظر به روش تصادفی ساده و منظم گرفته می‌شود. (نکته مهم این است که در جوامعی مورد استفاده قرار می‌گیرد که از نظر صفت موردنظر ناهمگون است). مانند بررسی عملکرد واحدهای مختلف یک سازمان.

۴ـ نمونه‌گیری خوشه‌ای:
هر گاه جامعه موردنظر خیلی وسیع و گسترده باشد مانند وضعیت معاش یا تحصیل یک شهر بزرگ یا یک کشور برای کارمندان برای نمونه‌گیری ابتدا سازمان‌ها یا اداراتی را به روش تصادفی ساده یا سیستماتیک (منظم) انتخاب می‌کنیم سپس کارمندان موردنیاز را با استفاده از همین روش به دست می آوریم (در این‌جا واحد نمونه‌گیری خوشه‌ای سازمان بوده است).

۵ـ نمونه‌گیری مرحله‌ای:
شکل گسترش یافته نمونه‌گیری خوشه‌ای است . در این حالت نمونه‌گیری از جامعه طی چند مرحله انجام می‌شود. یعنی انتخاب نمونه از نمونه دیگر، به طور مثال چند سازمان از شهر انتخاب می‌کنیم سپس از بین هر سازمان چند واحد را معین می‌کنیم سپس عناصر نمونه را به صورت تصادفی بدست می‌آوریم.

مقیاس:
با توجه به نوع صفات کیفی و کمی، مقیاس‌های متفاوتی برای اندازه‌گیری متغیرها وجود دارد. انواع مقیاس‌ها به شرح زیر است:

۱ـ اسمی ۲ـ رتبه‌ای (ترتیبی) ۳ـ فاصله‌ای ۴ـ نسبی (نسبتی)

۱ـ مقیاس اسمی: ضعیف‌ترین شکل اندازه‌گیری است که در آن از اعداد و علائم برای طبقه‌بندی اشیاء، اشخاص یا خصوصیت استفاده می‌شود. مانند مشخص کردن سازمان‌ها با اسم‌های A و B و C و ... .

این نوع مقیاس به علت ضعف در اندازه‌گیری، در صفات کیفی استفاده می‌شود.

۲ـ مقیاس رتبه‌ای (ترتیبی): در مواردی صرف‌نظر از محتویات یک طبقه یا گروه با طبقه یا گروه دیگر نوعی ارتباط بین آن‌ها برقرار است، این روابط با توجه به نوع مقیاس نشان‌دهندهٔ حالت ترتیبی است. برای مثال طبقه‌بندی افراد جامعه به صورت پردرآمد ـ متوسط و کم درآمد ـ ترتیبی است. مثال‌هایی مانند قوی ـ متوسط ـ ضعیف ، بالا ـ وسط ـ پایین، بزرگتر ـ مساوی ـ کوچکتر نیز نشان‌دهنده حالت ترتیبی است.

این نوع مقیاس نیز به علت ضعف در اندازه‌گیری، در صفات کیفی مورد استفاده قرار می‌گیرد.

۳ـ مقیاس فاصله‌ای: وقتی یک مقیاس همهٔ خصوصیات مقیاس ترتیبی را داشته و به علاوه فاصله بین هر دو عدد نیز در آن مشخص باشد، به یک مقیاس قوی‌تر رسیده‌ایم که می‌تواند برای صفات کمی مورد استفاده قرار گیرد.

در این مقیاس، صفر به صورت قراردادی و اختیاری است. مانند: سانتی‌گراد و فارنهایت که دارای صفرهای قراردادی مختلفی هستند.

در این نوع مقیاس نسبت هر دو فاصله مستقل از واحد اندازه‌گیری و مستقل از نقطه صفراست.

۴ـ مقیاس نسبتی (نسبی): دقیق‌ترین مقیاس برای صفات کمی است که علاوه بر داشتن تمام خصوصیات مقیاس فاصله‌ای ، دارای صفر واقعی نیز می‌باشد. مقیاس‌هایی مثل: پوند، گرم، متر، اینچ دارای صفر واقعی هستند. همچنین نسبت هر دونقطه دلخواه مستقل از واحد اندازه‌گیری است (مانند مقیاس فاصله‌ای). برای مثال وزن دوشی را می‌توان هم با گرم و هم با پوند اندازه‌گیری کرد. اما نسبت آن‌ها با هم فرقی نمی‌کند و به واحد اندازه‌گیری ربطی ندارد.

مقیاس	ترتیب	فواصل	صفر قراردادی	صفر مطلق (واقعی)
اسمی	ندارد	ندارد	ندارد	ندارد
رتبه‌ای	دارد	ندارد	ندارد	ندارد
فاصله‌ای	دارد	دارد	دارد	ندارد
نسبی	دارد	دارد	دارد	دارد

طبقه‌بندی صفات:

از آن‌جا که اطلاعات آماری به صورت اعداد و ارقام بیان می‌شوند، اگر بتوان آن‌ها را به صورت طبقه‌بندی شده بیان کرد، به راحتی می‌توان به خصوصیات مهم آن‌ها پی برد:

ـ به اعدادی که طبقات یک جدول توزیع فراوانی را مشخص می‌سازند، حدود طبقات می‌گویند.

ـ مرکز یک طبقه برابر نصف مجموع حد پائین و حد بالای آن طبقه است.

ـ طول دسته تفاوت بین حدود بالا یا پائین دو طبقه متوالی است.

طبقات	$0-5$	$5-10$	$10-15$
فراوانی	3	4	13

ـ به طور مثال در طبقه اول 5 حد بالا و 0 حد پائین را تشکیل می‌دهد. از طرفی $\frac{0+5}{2}=2.5$ مرکز طبقه اول، $\frac{5+10}{2}=7.5$ مرکز طبقه دوم، $\frac{10+15}{2}=12.5$ مرکز طبقه سوم و 5 طول طبقات است؛ چرا که $5-0=5$ ، $10-5=5$ یا $15-10=5$ می‌باشد.

فراوانی‌ها:

۱ـ فراوانی مطلق: تعداد دفعات تکرار هر داده i را فراوانی مطلق دادهٔ i گوئیم که با F_i نشان داده می‌شود.

مجموع فراوانی‌های مطلق داده‌ها برابر با حجم جامعه می‌شود:

در صورتی که K داده مختلف داشته باشیم : $\sum_{i=1}^{K} F_i = N$

۲ـ فراوانی نسبی (f_i): نسبت فراوانی مطلق هر داده i به حجم جامعه را فراوانی نسبی می‌نامند.

نکته: مجموع فراوانی‌های نسبی باید ۱ شود. (یعنی $f_i = P_i$ است)

۳ـ درصد فراوانی نسبی: به حاصلضرب فراوانی نسبی هر داده i در ۱۰۰، درصد فراوانی نسبی گفته می‌شود.

۴ـ فراوانی تجمعی (F_{C_i}): فراوانی تجمعی هر داده (طبقه) برابر است با جمع فراوانی مطلق همان طبقه به علاوه فراوانی‌های مطلق طبقات ماقبل آن :

$$F_{C_i} = F_1 + F_2 + \cdots + F_i$$

نکته: فراوانی تجمعی طبقه آخر برابر با حجم کل جامعه است.

مثال:

C-L	$0-5$	$5-10$	$10-15$	$15-20$
F_i	10	20	30	40
F_{C_i}	10	30	60	100
f_i	0.1	0.2	0.3	0.4

$N = \sum_{i=1}^{K} F_i = 100$

حجم کل جامعه = فراوانی تجمعی طبق آخر ⟵

$$f_i = \frac{F_i}{N} = \frac{F_i}{\Sigma F_i} \quad ; \quad F_{C_i} = F_1 + F_2 + \cdots + F_i$$

مشخص کننده‌های عددی:

پارامترهایی هستند که برای مقایسه بین چند جامعه به کار می‌روند و به سه بخش تقسیم می‌شوند:

۱) **پارامترهای مرکزی:** میانگین ـ میانه ـ چندک‌ها (چارک‌ها ـ دهک‌ها ـ صدک‌ها) ـ مد

۱) میانگین:

اصلی‌ترین شاخص مرکزی میانگین است. در حقیقت اگر داده‌ها روی یک محور به صورت منظم ردیف شوند، مقدار میانگین دقیقاً نقطه تعادل و مرکز ثقل توزیع است.

۲) پارامترهای پراکندگی

۳) پارامترهای نسبی پراکندگی

انواع میانگین:

۱ـ میانگین حسابی $(\mu = \bar{x})$

۲ـ میانگین هندسی $(\overline{X}_G)$

۳ـ میانگین هارمونیک $(\overline{X}_H)$

۴ـ میانگین وزنی $(\overline{X}_\omega)$

نکته مهم: همیشه $\overline{X} > \overline{X}_G > \overline{X}_H$ است و فقط زمانی که داده‌ها با یکدیگر برابر باشند $\overline{X} = \overline{X}_G = \overline{X}_H$ خواهد بود.

۱ـ میانگین حسابی: تعدادی از مشاهدات که بیشتر داده‌ها حول آن متمرکز شده‌اند، میانگین حسابی نامیده می‌شود.

مثال: کدام‌یک از روابط زیر بین میانگین حسابی $(\overline{x})$ ، میانگین هندسی $(\overline{x}_G)$ و میانگین هارمونیک $(\overline{x}_H)$ برقرار است؟

۱) $\overline{X}_H < \overline{X}_G < \overline{X}$ ۲) $\overline{X}_G < \overline{X} < \overline{X}_H$ ۳) $\overline{X}_G < \overline{X}_H < \overline{X}$ ۴) $\overline{X} < \overline{X}_G < \overline{X}_H$

حل:

با توجه به تعاریف بالا گزینه (۱) صحیح می‌باشد.

محاسبه میانگین حسابی:

۱) داده‌های طبقه‌بندی نشده:

$$\mu = \overline{X} = \frac{\sum_{i=1}^{n} x_i}{N} = \frac{\sum_{i=1}^{n} F_i x_i}{N} = \sum_{i=1}^{n} f_i x_i$$

مثال: اگر میانگین $x_1, x_2, \ldots, x_{10}$ برابر ۵ باشد و عدد ۱۰ به آن‌ها اضافه شود، میانگین $x_1, x_2, \ldots, x_{10}$ و ۱۰ چیست؟

$$5 = \frac{\sum_{i=1}^{10} x_i}{10} \Rightarrow \sum_{i=1}^{10} x_i = 50$$

$$\overline{X} = \frac{\sum_{i=1}^{10} x_i + 10}{11} = \frac{50+10}{11} = \frac{60}{11}$$

نکته مهم: میانگین اعداد طبیعی ۱ تا N برابر $\frac{N+1}{2}$ است.

$$\overline{X} = \frac{1+2+3+\cdots+N}{N} = \frac{\frac{N\ (N+1)}{2}}{N} = \frac{N+1}{2}$$

مثال: میانگین داده‌های زیر چیست؟

$x_i =$ داده	3	1	2
F_i فراوانی مطلق	2	4	5

$$\overline{X} = \frac{3\times 2 + 1\times 4 + 2\times 5}{2+4+5} = \frac{20}{11}$$

نکته: میانگین داده‌های x_1 تا x_n که تشکیل تصاعد حسابی را می‌دهند به صورت روبرو محاسبه می‌شود:

$$\overline{X} = \frac{x_1 + x_N}{2}$$

مثال: جدول توزیع فراوانی زیر را در نظر بگیرید. اگر $\mu = 2$ و $N = 28$ باشد، مقادیر a و b عبارتند از:

x_i	0	1	2	3	4
F_i	3	a	10	b	3

۱) $a = b = ۶$ ۲) $b = ۷$ و $a = ۵$ ۳) $a = b = ۷$ ۴) $b = ۸$ و $a = ۴$

حل:

$$\mu = \frac{\sum F_i x_i}{N} \to 2 = \frac{(0\times 3)+(1\times a)+(2\times 10)+(3\times b)+(4\times 3)}{28}$$

$$\to 56 = a + 20 + 3b + 12 \to a + 3b = 24$$

از طرفی

$$\sum F_i = N \to 3 + a + 10 + b + 3 = 28 \to a + b = 12$$

حال می‌توان a , b را به صورت زیر بدست آورد:

$$\begin{cases} a+3b=24 \\ a+b=12 \end{cases} \to \begin{cases} a+3b=24 \\ -a-b=-12 \end{cases}$$

$$2b = 12 \to b = 6$$

$$a + 3(6) = 24 \to a = 6$$

۲) داده‌های طبقه‌بندی شده:

ابتدا مرکز طبقات را بدست آورده و سپس به فرم زیر عمل می‌کنیم:

$$\overline{X} = \frac{\sum F_i\, x_i}{N} = \sum f_i\, x_i \qquad (x_i\text{: مرکز طبقه})$$

$$f_i = \frac{F_i}{N}$$

مثال: مطلوبست میانگین جدول زیر:

C−L	0−5	5−10	10−15	
F_i	2	3	4	$\Sigma F_i = N = 9$

ابتدا مرکز طبقات را بدست می‌آوریم:

$$\text{مرکز هر طبقه} = \frac{\text{حد بالا} + \text{حد پائین}}{۲}$$

x_i	$\underbrace{\frac{5+0}{2}}_{2.5}$	$\underbrace{\frac{5+10}{2}}_{7.5}$	$\underbrace{\frac{10+15}{2}}_{12.5}$
F_i	2	3	4

$$\Rightarrow \frac{\sum F_i x_i}{N} = \frac{(2\times 2.5)+(3\times 7.5)+(4\times 12.5)}{9} = \frac{5+22.5+50}{9} = \frac{77.5}{9}$$

۲ـ میانگین هندسی $(\bar{X}_G)$: اگر داده‌های بدست آمده نسبت،درصد، شاخص نرخ رشد و ... باشد، برای بدست آوردن مقدار متوسط از میانگین هندسی استفاده می‌کنیم:

$$\bar{X}_G = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n}$$

نکته: باید توجه داشت که اگر در بین داده‌ها عدد صفر یا منفی وجود داشته باشد، نمی‌توان از این روش استفاده کرد.

مثال: میزان سود سهامی شرکت بتا در ۵ سال گذشته بر حسب درصد فروش به ترتیب ۳،۴،۴،۳،۲ می‌باشد. کدام یک از کمیت‌های زیر به عنوان شاخص مرکزی وضع سودآوری شرکت را بهتر نشان می‌دهد؟

۱) ۲/۵   ۲) ۳/۲   ۳) ۴   ۴) ۳/۱

حل:

از آن‌جا که کمیت‌های تصادفی x بر حسب درصد فروش بیان شده‌اند، میانگین هندسی به عنوان شاخص مرکزی وضع سودآوری شرکت را بهتر نشان می‌دهد.

$$\bar{x}_G = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots \times x_n} = \sqrt[3]{2 \times 3 \times 4 \times 4 \times 3} = \sqrt{288} = 3.104$$

مثال: فرض کنید شاخص قیمت خرده فروشی از ۲۰۰ در سال ۷۸ به ۴۵۰ در سال ۸۰ رسیده باشد، متوسط نرخ تورم در این فاصله زمانی چه بوده است؟

۱) ٪۵۰   ۲) ٪۶۲/۵   ۳) ٪۱۲۵   ۴) ٪۱۵۰

حل:

چون $\bar{X}_G$، 1.5 برابر شده بنابراین 50% تورم افزایش یافته است.

$$\bar{X}_G = \sqrt{\frac{\text{سال } 79}{\text{سال } 78} \times \frac{\text{سال } 80}{\text{سال } 79}} = \sqrt{\frac{450}{200}} = 1.5$$

$$1.5 - 1 = 0.5 = \%50$$

مثال: تعداد کارکنان کارخانه‌ای طی چهار سال متوالی ۱۹۰، ۱۸۰، ۱۶۰، ۱۵۰ بوده است. متوسط رشد سالانه تعداد کارکنان چند درصد است؟

۱) ٪۵/۵   ۲) ٪۶/۰۸   ۳) ٪۷/۴   ۴) ٪۸/۲

حل:

چون $\bar{X}_G$، 1.082 برابر شده بنابراین 8.2% رشد داشته است.

$$\bar{X}_G = \sqrt[3]{\frac{160}{150} \times \frac{180}{160} \times \frac{190}{180}} = 1.082$$

$$1.082 - 1.000 = 0.082 = \%8.2$$

مثال: با تغییر مدیریت در یک فروشگاه بزرگ، فروش در سال اول دو برابر سال قبل، در سال دوم سه برابر سال اول و در سال سوم چهار برابر سال دوم شده است.به طور متوسط، فروش از آغاز مدیریت چند برابر شده است؟

۱) بیش از سه برابر   ۲) سه برابر   ۳) قدری کمتر از سه برابر   ۴) ۲/۵ برابر

حل:

$\bar{X}_G$ 2.88 برابر شده است

$$\bar{x}_G = \sqrt[n]{x_1 \times x_2 \times \ldots \times x_n} = \sqrt[3]{2 \times 3 \times 4} = \sqrt[3]{24} = 2.88$$

مثال: سرمایه شرکتی در سال ۱۳۶۴، ۲ میلیون تومان، در سال ۱۳۶۵، ۴ میلیون تومان و در سال ۱۳۶۶، ۳۲ میلیون تومان بوده است. به طور متوسط این شرکت هر سال نسبت به سال قبل چند برابر سود داشته است؟

۱) ۳/۸۴ ۲) ۴ ۳) $\frac{۳۸}{۳}$ ۴) $(۲۵۶)^{\frac{۱}{۲}}$

حل:

$$\overline{X}_G = \sqrt[n]{x_1 \times x_2 \times \dots\dots \times x_n} = \sqrt[2]{\frac{4}{2} \times \frac{32}{4}} = \sqrt{16} = 4$$

۳ـ میانگین هارمونیک (توافقی) معکوس (همساز) $\overline{X}_h$: اگر مقیاس داده‌ها به صورت ترکیبی باشد از این میانگین استفاده می‌کنیم. مانند: متر در ثانیه، کیلومتر بر ساعت و.....

$$\overline{X}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

نکته: ممکن است داده‌ها دارای وزن باشند، آنگاه میانگین هارمونیک به فرم زیر محاسبه می‌شود:

$$\overline{X}_h = \frac{\sum_{i=1}^{n} \omega_i}{\frac{\omega_1}{x_1} + \dots + \frac{\omega_n}{x_n}}$$

مثال: اتومبیلی مسیری را با سرعت ۱۰۰ کیلومتر رفته، و $\frac{1}{3}$ مسیر را با سرعت ۸۰ کیلومتر و باقیمانده را با سرعت ۱۲۰ کیلومتر برگشته، سرعت متوسط این اتومبیل چقدر بوده است؟

۱) ۹۰ ۲) ۱۰۰ ۳) ۱۰۱/۴ ۴) ۱۰۲/۸

حل:

$$\overline{X}_h = \frac{\sum_{i=1}^{n} \omega_i}{\frac{\omega_1}{x_1} + \dots + \frac{\omega_n}{x_n}} = \frac{1 + \frac{1}{3} + \frac{2}{3}}{\frac{1}{100} + \frac{1/3}{80} + \frac{2/3}{120}} = 101.4$$

مثال: اگر ۳ اتومبیل مسیر ۶۰ کیلومتری بین دو منطقه را به ترتیب با سرعت ۱۲۰ و ۶۰ و ۹۰ کیلومتر در ساعت طی نمایند. میانگین سرعت این سه اتومبیل برابر با چند کیلومتر در ساعت است؟

۱) تقریباً ۸۳ ۲) تقریباً ۸۶ ۳) تقریباً ۹۰ ۴) ۹۰

حل:

$$\overline{X}_h = \frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}} = \frac{3}{\frac{1}{20} + \frac{1}{60} + \frac{1}{90}} = 83.076$$

مثال: سه ماشین به تولید یک کالا مشغول‌اند، اولی یک کالا را در ۲، دومی در ۳ و سومی در ۶ دقیقه تولید می‌کنند. اگر این سه ماشین با هم کار کنند به طور متوسط یک کالا در چند دقیقه تولید می‌شود؟

۱) ۳/۶۷ ۲) ۳/۳ ۳) ۳ ۴) ۳/۱

حل:

$$\overline{x}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{3}{\frac{1}{2} + \frac{1}{3} + \frac{1}{6}} = 3$$

٤ـ میانگین وزنی $(\overline{X}_\omega)$: در صورتی که داده‌ها دارای وزن‌های متفاوتی باشند، از میانگین وزنی استفاده می‌کنیم:

$$\overline{X}_\omega = \frac{\sum_{i=1}^{n} x_i \omega_i}{\sum_{i=1}^{n} \omega_i}$$

مثال: در صورتی که نمرات یک دانشجو در یک ترم به فرم زیر باشدمطلوبست محاسبه معدل یا میانگین معدل او در این ترم.

	ω_i	x_i
آمار	۳ واحد	۱۰
ریاضی	۲ واحد	۱۵
فیزیک	۳ واحد	۱۰

$$\overline{X}_\omega = \frac{\sum_{i=1}^{n} x_i \omega_i}{\sum_{i=1}^{n} \omega_i} = \frac{3\times10+2\times15+3\times10}{3+2+3} = \frac{90}{8} = 11.25$$

خواص مهم میانگین:

۱) $\sum(x_i-\mu)=0$ ، به عبارت دیگر مجموع انحرافات از میانگین همیشه صفر است.

۲) $\sum (x_i-\mu)^2 < \sum (x_i-a)^2$ ، مجموع مجذور انحرافات از میانگین همیشه می‌نیمم است. «a عدد دلخواه است»

۳) اگر a و b اعداد ثابتی باشند:

الف) $\mu(x\pm a)=\mu_x \pm a$ ب) $\mu(a)=a$ ج) $\mu(x\pm y)=\mu_x \pm \mu_y$

د) $\mu\left(\frac{x}{a}\right)=\frac{1}{a}\mu_x$ هـ) $\mu_{(ax)} = a\mu_{(x)}$

۴) در جامعه آماری فقط یک میانگین داریم.

۵) مقادیر بزرگ و کوچک به سهم خود در میانگین سهم دارند.

۶) میانگین تنها پارامتری است که اگر به جای کلیه داده‌ها قرار گیرد، مجموع آن‌ها تغییری نمی‌کند:

$$\underbrace{1,2,3}_{\text{مجموع}=6} \rightarrow \overline{x} = \frac{1+2+3}{3} = 2 \Rightarrow \underbrace{2,2,2}_{\text{مجموع}=6}$$

مثال: اگر کمیت‌های $x_1, x_2, \dots, x_n$ با حجم n به دست آمده باشند، کدام یک از روابط زیر صادق است؟

۱) $\sum X_i = n\overline{X}^2$ ۲) $\sum(X_i-\overline{X})=0$ ۳) $\sum(X_i-\overline{X})^2=0$ ۴) $\sum(X_i-me)=0$

حل:

مجموع انحرافات از میانگین همواره صفر است. به عبارت دیگر $\sum(x_i-\overline{x})=0$ می‌باشد.

۲ـ میانه (Md)

عددی است که ۵۰٪ داده‌ها قبل و ۵۰٪ داده‌ها بعد آن قرار دارند. به عبارت دیگر تعداد داده‌های قبل و بعد از میانه با هم برابر هستند.

نکته: برای محاسبه میانه همیشه در دو حالت زیر می‌توان عمل نمود:

۱) داده‌های طبقه‌بندی نشده

۲) داده‌های طبقه‌بندی شده

۱ـ داده‌های طبقه‌بندی نشده:

برای این داده‌ها به طریق زیر عمل می‌کنیم:

الف) ابتدا کل داده‌ها را به صورت صعودی مرتب می‌کنیم.

ب) محل میانه توسط $\frac{N}{2}+\frac{1}{2}$ مشخص می‌شود.

مثال: میانه داده‌های زیر را محاسبه نمایید.

۹ و ۷ و ۵ و ۰ و ۴ و ۱ـ

حل:

مرتب‌سازی: ۹ و ۷ و ۵ و ۴ و ۰ و ۱ـ

$$\text{محل میانه}=\frac{n}{۲}+\frac{۱}{۲}=\frac{۶}{۲}+\frac{۱}{۲}=\frac{۷}{۲}=۳/۵$$

$$\text{مقدار میانه}=۴+۰/۵\left(۵-۴\right)=۴/۵$$

نکته مهم: اگر تعداد داده‌ها فرد باشد، میانه بعد از مرتب‌سازی دادهٔ وسط است. اما در صورتی که تعداد داده‌ها زوج باشد، میانه بعد از مرتب‌سازی میانگین دو عدد وسط خواهد بود.

۶ = میانه ← ۹ و ۷ و ۶ و ۰ و ۱ـ

$\frac{۶+۷}{۲}=۶/۵$ میانه ← ۱۰ و ۹ و ۷ و ۶ و ۰ و ۱ـ

مثال مهم: جمعیت خانواده‌های یک روستا به صورت زیر است، میانه جمعیت خانواده چقدر است؟

جمعیت خانواده x_i	۱	۲	۳	۴	۵	۶	جمع
تعداد F_i	۵	۱۰	۴۰	۲۵	۱۵	۵	$۱۰۰=\sum F_i=N$
F_{C_i}	۵	۱۵	۵۵	۸۰	۹۵	۱۰۰	

۱) ۳   ۲) ۳/۵   ۳) ۳/۲۵   ۴) ۴

حل:

از روی جدول، فراوانی تجمعی را بدست می‌آوریم. محل میانه با توجه به ترتیب صعودی که در جدول حفظ شده به صورت زیر بدست می‌آید:

$$\text{محل میانه}=\frac{n}{2}+\frac{1}{2}=\frac{100}{2}+\frac{1}{2}=50.5$$

داده پنجاهم

$$\text{مقدار میانه} = 3 + 0.5\left(3-3\right) = 3$$

داده پنجاهم   داده پنجاه و یکم

البته چون ۵۰/۵ در داخل جمعیت خانواده‌ای ۳ نفری است، به راحتی می‌توانیم بدون محاسبه ۳ را به عنوان میانه انتخاب کنیم.

۲ـ داده‌های طبقه‌بندی شده:

برای محاسبه میانه در داده‌های طبقه‌بندی شده به طریق زیر عمل می‌کنیم:

۱ـ ابتدا $\frac{n}{2}$ را بدست آورده، سپس در جدول اولین طبقه‌ای که مقدار فراوانی تجمعی $\left(F_{C_i}\right)$ آن بیشتر یا مساوی $\frac{n}{2}$ است را به عنوان طبقه میانه‌دار انتخاب می‌کنیم.

۲ـ با استفاده از رابطه زیر مقدار میانه را بدست می‌آوریم:

$$\text{(میانه) } md = \text{حد پائین طبقه میانه‌دار} + \left(\frac{\frac{N}{2} - \text{فراوانی تجمعی طبقه ماقبل میانه‌دار}}{\text{فراوانی مطلق طبقه میانه‌دار}}\right) \times \text{طول هر طبقه}$$

به جای فراوانی مطلق طبقه میانه‌دار در مخرج می توانیم (فراوانی تجمعی طبقه ماقبل میانه ـ فراوانی تجمعی طبقه میانه‌دار) را انتخاب کنیم، اما تفاوتی ندارد.

مثال: میانه داده‌های زیر کدام است؟

حد پایین طبقه میانه‌دار

C−L	10-20	20-30	30-40	40-50	جمع
F_i	10	20	30	40	$N = \sum F_i = 100$
F_{C_i}	10	30	60	100	

فراوانی تجمعی طبقه ماقبل

فراوانی مطلق طبقه میانه‌دار

حل:

۱) طبقه ۴۰ ـ ۳۰ اولین طبقه‌ای است که فراوانی تجمعی‌اش بیشتر یا مساوی ۵۰ می‌باشد.

$$\frac{n}{2} = \frac{100}{2} = 50$$

۲) مقدار میانه با توجه به فرمول به صورت زیر محاسبه می‌شود:

$$md = 30 + \frac{\left(\frac{100}{2} - 30\right)}{\underbrace{(60-30)}_{30}} \times 10 = 30 + \frac{20}{30} \times 10 = 36.66$$

نکته بسیار مهم: در صورتیکه جدول دارای طبقات گسسته باشد، به صورت زیر عمل می‌کنیم:

مثال: میانه جدول زیررا بدست آورید.

C – L	3-5	6-8	9-11	جمع
F_i	4	20	12	$\sum F_i = N = 36$

طبقات را به صورت پیوسته ظاهر می‌کنیم:

C – L	2.5-5.5	5.5-8.5	8.5-11.5	جمع
F_i	4	20	12	$\sum F_i = N = ۳۶$
F_{C_i}	4	24	36	

۱) **محل میانه:** اولین فراوانی تجمعی بیشتر یا مساوی $18 = \frac{36}{2} = \frac{n}{2}$ در طبقه ۸/۵ ـ ۵/۵ قرار دارد.

۲) **مقدار میانه:**

$$\text{مقدار میانه} = 5.5 + \frac{\left(\frac{36}{2} - 4\right)}{\underbrace{(24-4)}_{20}} \times \overbrace{(8.5-5.5)}^{\text{طول طبقه}} = 5.5 + \frac{14}{20} \times 3 = 7.6$$

نتیجه‌گیری:

به طور کلی در جداول با طبقات گسسته کافیست $\frac{\text{حد پایین طبقه بعدی + حد بالا طبقه}}{۲}$ شده و از حد پایین طبقه اول ۰/۵ واحد کم کنیم و به حد بالای طبقه آخر ۰/۵ واحد اضافه کنیم.

خواص مهم میانه:

۱) در هر جامعه آماری فقط یک میانه وجود دارد.

۲) برخلاف میانگین، میانه از اعداد بسیار بزرگ یا بسیار کوچک متأثر نمی‌شود.

۳) مهم‌ترین خاصیت میانه این است که:

$\sum |x_i - Md| < \sum |x_i - a|$ (مجموع قدرمطلق انحرافات از میانه همیشه می‌نیمم است)

«a عدد دلخواه است»

۴) از لحاظ هندسی میانه طول خط عمودی است بر محور xها در نمودار بافت‌نگار که آن را به دو قسمت مساوی تقسیم می‌کند.

۵) چنانچه اندازه افرادی که در ابتدا و انتهای توزیع واقع شده‌اند، از سایر اندازه‌ها به طور قابل ملاحظه‌ای تفاوت داشته باشند، بهتر است از میانه به عنوان پارامتر حد وسط استفاده کنیم. زیرا میانگین از حد وسط واقعی دور است.

۶) در نمودار اجایو محل برخورد دو نمودار، فراوانی تجمعی کمتر از میانه و فراوانی تجمعی بیشتر از میانه را به دست می‌دهد. می‌توان گفت میانه طول نقطه‌ای است که عرض آن ۵۰٪ است.

مثال: کدام یک از عبارت‌های زیر می‌تواند یکی از خواص مهم میانه را در توزیع جامعه بیان کند؟

۱) $\sum(X_i - me)^2$ ۲) $\sum(X_i - me)$ ۳) $\sum|X_i - me|$ ۴) $\sum(\overline{X} - me)$

حل:

مجموع قدرمطلق انحرافات از میانه همواره می‌نیمم است بنابراین گزینه ۳ صحیح می‌باشد.

مثال: در صورتی که به بزرگ‌ترین عدد یک سری داده مقدار ثابتی اضافه گردد، این افزایش بر کدام معیار تأثیر نمی‌گذارد؟

۱) ضریب پراکندگی ۲) میانه ۳) میانگین ۴) واریانس

حل:

۵۰ درصد داده‌ها قبل و ۵۰ درصد داده‌ها بعد از میانه قرار دارند. از این رو بزرگ یا کوچک بودن متغیرها تأثیری بر مقدار میانه نخواهد داشت.

۳ـ چندک‌ها (چارک‌ها ـ دهک‌ها ـ صدک‌ها):

چندک‌ها مقادیری از مشاهدات هستند که دامنهٔ تغییرات را در فواصل چندکی تقسیم می‌کنند. به طوری که فراوانی‌ها در هر یک از این فواصل، درصد معینی از فراوانی کل را تشکیل می‌دهد.

چارک‌ها: دامنهٔ تغییرات را به چهار قسمت مساوی تقسیم می‌کنند.

دهک‌ها: دامنهٔ تغییرات را به ده قسمت مساوی تقسیم می‌کنند.

صدک‌ها: دامنهٔ تغییرات را به صد قسمت مساوی تقسیم می‌کنند.

محاسبه چندک‌ها:

۱ـ داده‌های طبقه‌بندی نشده :

الف) ابتدا داده‌ها را به ترتیب صعودی مرتب می‌کنیم.

ب) سپس با توجه به نوع چندک محل آن را با استفاده از:

چارک: $\frac{aN}{4}+\frac{1}{2}$; $a=1,2,3$

دهک: $\frac{aN}{10}+\frac{1}{2}$; $a=1,2,\ldots,9$ پیدا می‌کنیم.

صدک: $\frac{aN}{100}+\frac{1}{2}$; $a=1,2,\ldots,99$

مثال: مطلوبست دهک ۴ ام از داده‌های زیر:

$-1, 4, 0, 2, 6, 9 \rightarrow$ داده‌ها به ترتیب صعودی :- $-1, 0, 2, 4, 6, 9$

محل دهک ۴ ام $= \frac{4\times 6}{10}+\frac{1}{2}=\frac{24}{10}+\frac{1}{2}=\frac{29}{10}=2.9$

=0 دومین داده

=2 سومین داده

$0+0.9(2-0)=1.8$ = مقدار دهک ۴ام

۲ـ داده‌های طبقه‌بندی شده:

الف) ابتدا با استفاده از $\frac{aN}{4}\,(a=1,2,3)$ یا $\frac{aN}{10}\,(1,2,\dots,9)$ یا $\frac{aN}{100}\,(1,2,\dots,99)$ اولین طبقه‌ای که فراوانی تجمعی‌اش بیشتر یا مساوی یکی از مقادیر فوق باشد را با توجه به چارک، دهک یا صدک پیدا می‌کنیم.

ب) سپس با استفاده از فرمول زیر آن را محاسبه می‌نماییم:

$$\text{چارک} = \text{مقدار چندک} = \text{حد پائین طبقه چندک‌دار} + \frac{\left(\frac{aN}{4} - \text{فراوانی تجمعی طبقه ماقبل}\right)}{\text{فراوانی مطلق طبقه چندک‌دار}} \times \text{طول طبقه}$$

$$\text{دهک} = \text{مقدار چندک} = \text{حد پائین طبقه چندک‌دار} + \frac{\left(\frac{aN}{10} - \text{فراوانی تجمعی طبقه ماقبل}\right)}{\text{فراوانی مطلق طبقه چندک‌دار}} \times \text{طول طبقه}$$

$$\text{صدک} = \text{مقدار چندک} = \text{حد پائین طبقه چندک‌دار} + \frac{\left(\frac{aN}{100} - \text{فراوانی تجمعی طبقه ماقبل}\right)}{\text{فراوانی مطلق طبقه چندک‌دار}} \times \text{طول طبقه}$$

نکته: دهک پنجم = چارک دوم = صدک پنجاهم = میانه است.

مثال: مطلوبست دهک دوم جدول زیر:

C−L	40-50	50-60	60-70	
F_i	5	18	7	$\sum F_i = N = 30$
F_{C_i}	5	23	30	

۱) ۴۸/۲ ۲) ۵۱/۵ ۳) ۵۰/۵۵ ۴) ۶۲/۳۸

حل:

مقدار دهک دوم =

الف) $\frac{\overset{a}{2}\times\overset{N}{30}}{10}=6$ ۶ در طبقه ۶۰ـ۵۰ قرار دارد.

ب) مقدار $=50+\frac{(6-5)}{18}\times 10=50+\frac{10}{18}=50.55$

مثال: چارک سوم جدول زیر کدام است؟

C-L	2-5	6-9	10-13
F_i	10	30	20

۱) ۶/۸ ۲) ۹/۵ ۳) ۱۰ ۴) ۱۰/۵

حل:

C-L	1.5-5.5	5.5-9.5	9.5-13.5
F_i	10	30	20

F_{ci}	10	40	60

$\sum F_i = N = 60$

برای محاسبه چارک داده‌های گسسته، باید حدود طبقات به صورت پیوسته باشند. چارک در طبقه ۱۳/۵ ـ ۹/۵ قرار دارد.

$$\text{محل چارک} = \frac{aN}{4} = \frac{3\times 60}{4} = 45$$

$$\text{مقدار چارک} = \text{حد پایین طبقه چارک‌دار} + \frac{\frac{aN}{4} - \text{فراوانی تجمعی طبقه ماقبل}}{\text{فراوانی مطلق طبقه چارک‌دار}} \times \text{طول طبقه}$$

$$= 9.5 + \frac{45-40}{20}\times(13.5-9.5) = 10.5$$

مثال: مطلوب است محاسبه چارک اول در جدول زیر:

۱) ۵ ۲) ۶/۱۷ ۳) ۹/۵ ۴) ۱۰

C-L	۲ـ۵	۶ـ۹	۱۰ـ۱۳
Fi	۱۰	۳۰	۲۰

حل:

طبقات گسسته جدول را به پیوسته تبدیل می‌کنیم.

C-L	1.5-5.5	5.5-9.5	9.5-13.5
Fi	10	30	20
Fci	10	40	60

$N = \sum_{i=1}^{n} F_i = 60$

طبقه ۹/۵ ـ ۵/۵ طبقه موردنظر است؛ چرا که اولین طبقه‌ای است که Fci آن بیشتر یا مساوی ۱۵ می‌باشد. $۱۵ = \frac{۱\times ۶۰}{۴}$ (الف

ب) مقدار چارک اول $= 5.5 + \frac{(15-10)}{30}\times 4 = 6.17$

۴ـ مُد یا نما (MO):

متغیری که دارای بیشترین فراوانی است مد (نما) نامیده می‌شود. گفتنی است نما کم اهمیت‌ترین پارامتر مرکزی نیز به حساب می‌آید.

محاسبه مد:

۱ـ محاسبه مد در داده‌های طبقه‌بندی نشده:

مد داده‌ای است که فراوانی مطلق آن از سایر داده‌ها بیشتر است.

مثال:

۱ و ۱ و ۳ و ۴ و ۵ MO = ۱

۱ و ۱ و ۳ و ۲ و ۲ MO = ۱ و ۲

۲ و ۲ و ۱ و ۱ $MO =$ نداریم

۲ـ محاسبه مد در داده‌های طبقه‌بندی شده:

الف) ابتدا در ستون فراوانی مطلق طبقه‌ای را پیدا می کنیم که بیشترین فراوانی مطلق را دارد.

ب) سپس مقدار مد را از رابطه زیر بدست می‌آوریم:

$$MO = \text{حد پائین طبقه مددار} + \frac{d_1}{d_1 + d_2} \times \text{طول طبقه}$$

فراوانی مطلق طبقه ماقبل طبقه مددار ـ فراوانی مطلق طبقه مددار = d_1

فراوانی مطلق طبقه بعداز طبقه مددار ـ فراوانی مطلق طبقه مددار = d_2

مثال: نظر گروهی از سوادآموزان راجع به زمان پخش برنامهٔ نهضت سوادآموزی از سیمای جمهوری اسلامی جمع‌آوری شده است. کدام شاخص مرکزی برای آن داده‌ها مناسب‌تر است؟

۱) میانگین ۲) میانه ۳) نما ۴) چارک اول

حل:

با توجه به آن‌که بیشترین فراوانی سنجیده می‌شود، مد (نما) مناسب‌ترین شاخص مرکزی برای داده‌هاست.

مثال : در جدول زیر مد کدام مقدار است؟

۱) ۶/۳۳ ۲) ۶/۵۳ ۳) ۷/۵ ۴) ۲۰

C-L	3-5	6-8	9-11	جمع
F_i	4	20	12	$\underbrace{36}_{N}$

$\Rightarrow$

C-L	2.5-5.5	5.5-8.5	8.5-11.5
F_i	4	20	12

الف) طبقه مددار = (5.5-8.5)

$$\text{ب) مقدار مد} = 5.5 + \frac{(20-4)}{(20-4)+(20-12)} \times 3 = 5.5 + \frac{16}{16+8} \times 3 = 7.5$$

۱) ۶/۳۳ ۲) ۶/۵۳ ۳) ۷/۵ ۴) ۲۰

نکته: برای تعیین مد در هیستوگرام، در ستونی که بیشترین ارتفاع را دارد دو خط مورب رسم می کنیم. محل تقاطع خطوط، نشاندهنده مقدار مد است.

مثال: در جدول توزیع فراوانی زیر، مقدار مد و میانه به ترتیب (از چپ به راست) کدام است؟

x_i	-1	0	1	2	3
فراوانی	10	30	10	25	25

۱) (۰ ، ۰/۵۰) ۲) (۰ ، ۱/۵) ۳) (۲۵ ، ۰/۵۰) ۴) (۳۰ ، ۲۵)

حل:

x_i	-1	0	1	2	3
فراوانی	10	30	10	25	25

Fci	10	40	50	75	100

$\frac{N}{2}+\frac{1}{2}=\frac{50}{2}+\frac{1}{2}=50.5$ = محل میانه

$1+0.5(2-1)=1.5$ = مقدار میانه

مد متغیری است که دارای بیشترین فراوانی است، بنابراین صفر مد جدول است.

نکته مهم: تفاوت‌های اساسی بین میانگین، میانه و مد:

۱) میانگین برحسب مقیاس داده‌ها است و در محاسبه آن فراوانی و کمیت داده در نظر گرفته می‌شود. اما میانه و مد تابع ترتیب و فراوانی داده‌ها هستند.

۲) میانگین از ترکیب داده‌ها حاصل نشده و هر افزایش یا کاهش داده‌ها مقدار میانگین را عوض می‌کند، اما اگر افزایش یا کاهش ترتیب داده‌ها را عوض نکند در مقدار مد و میانه تأثیری ندارد.

پارامترهای پراکندگی:

به‌طور کلی پارامترهای پراکندگی معیارهایی برای تعیین میزان پراکندگی داده از یکدیگر یا میزان پراکندگی آن‌ها نسبت به میانگین است، که از پارامترهای پراکندگی می‌توان به دامنه تغییرات ـ دامنه میان چارکی ـ انحراف چارکی ـ واریانس ـ انحراف معیار ـ نیمه واریانس ـ انحراف متوسط از میانگین ـ ضریب پراکندگی ،ضریب تغییرات، (CV) اشاره نمود.

الف) دامنه تغییرات: R = max داده ـ داده min داده

کم اهمیت‌ترین پارامتر پراکندگی دامنه تغییرات است. زیرا تنها تابع تغییرات دو اندازه است و وضعیت اعداد وسط را مشخص نمی‌کند. از طرفی نمی‌تواند چگونگی توزیع اعداد را مشخص کند. حتی ممکن است در مواردی پراکندگی را بیش از حد بزرگ نشان دهد.

مثال:

۱۰ = (۱ـ) ـ ۹ = R ⇐ ۹ و ۰ و ۱ و ۷ و ۴ و ۱ـ

ب) دامنه میان چارکی (IQR) :

دامنه تغییرات ۵۰٪ از مشاهدات است. در این تعریف، میانه را که ۵۰٪ داده‌هاست ملاک قرار داده و از پائین تا ۲۵٪ و از بالا تا ۷۵٪ گسترش می‌دهیم.به عبارت بهتر IQR چارک اول و سوم را شامل می‌شود و برابر است با:

$IQR = Q_3 - Q_1$

نکته مهم: در بعضی از متون از انحراف چارکی استفاده شده که برابر است با $SIQR=\frac{IQR}{2}$ و از دامنه تغییرات باثبات‌تر می‌باشد.

نکته تستی ۱: در توزیع‌هائی که دارای تعداد اندکی مقدار در ابتدا و انتها هستند از انحراف چارکی به عنوان شاخص پراکندگی استفاده می‌شود.

نکته تستی ۲: در توزیع‌های نامتقارن اغلب از میانه به عنوان شاخص مرکزی و از انحراف چارکی (SIQR) به عنوان شاخص پراکندگی استفاده می‌شود.

مثال: کدام پارامتر پراکندگی برای توزیع فراوانی زیر مناسب‌تر است؟

حدود طبقاتی	۰ـ۲۰	۲۰ـ۴۰	۴۰ـ۶۰	۶۰ و بیش از آن
فراوانی	۲۵	۳۵	۳۰	۱۰

۱) انحراف چارکی ۲) انحراف معیار ۳) ضریب تغییرات ۴) انحراف متوسط از میانگین

حل:

* چون دارای تعداد کمی اطلاعات از ۶۰ وبیش از آن است انحراف چارکی مناسب‌ترین پارامتر پراکندگی برای توزیع فراوانی می‌باشد.

ج) انحراف متوسط از میانگین

هیچ‌کدام از شاخص‌هائی که تا به حال در مورد آن‌ها صحبت شده قادر به بیان تمامی تغییرات نیستند. تغییر زمانی مفهوم پیدا می‌کند که هر یک از داده نسبت به مبدأ مقایسه شوند و بهترین مبدأ یا مرکز داده‌ها میانگین است. اما از آن‌جا که مجموع انحرافات از میانگین همیشه صفر است، برای محاسبه انحراف متوسط از میانگین از رابطه زیر استفاده می‌کنیم:

$$\boxed{A.D_{\mu} = \frac{\sum_{i=1}^{n} |x_i - \mu|}{N}}$$

نکته مهم: این شاخص دارای ۲ مشکل اساسی است.

۱) این فرمول از علامت جبری داده‌ها صرف‌نظر می‌کند.

۲) تأثیر انحرافات بزرگ را در شرایطی که تعداد زیادی انحرافات کوچک وجود داشته باشد، نشان نمی‌دهد.

ویژگی‌های انحراف متوسط:

۱) در صورتی که تمام داده‌ها با هم برابر باشند انحراف متوسط از میانگین، صفر است.

۲) اگر به متغیرها عدد ثابتی اضافه یا کم شود، انحراف متوسط تغییری نمی‌کند.

۳) اگر متغیرها را در عدد ثابتی ضرب یا تقسیم کنیم، انحراف متوسط در قدر مطلق آن عدد ضرب یا بر قدرمطلق آن عدد تقسیم می‌شود.

۴) انحراف متوسط اعداد ثابت صفر است.

د) واریانس و انحراف معیار:

یکی از شاخص‌های پراکندگی داده نسبت به میانگین،واریانس است. هنگام محاسبه انحراف متوسط از میانگین برای جلوگیری از خنثی شدن انحرافات منفی از قدرمطلق استفاده می‌کردیم، اما در واریانس از مجذور انحرافات استفاده می‌کنیم که شاخص بهتری نسبت به انحراف متوسط از میانگین است.

نکته مهم:

واریانس و انحراف معیار می‌توانند تأثیرات انحرافات بزرگ را به راحتی نشان دهند، بنابراین بهترین شاخص برای نشان دادن انحرافات بزرگ هستند.

خواص واریانس:

۱) واریانس اعداد مساوی، صفر است.

مثال: مطلوبست واریانس اعداد 1667 و 1667 و 1667 و ...؟ چون داده‌ها برابرند بنابراین، واریانس مساوی صفر است.

۲) فرمول محاسبه:

$$\delta_x^2 = \frac{\sum (x_i - \bar{x})^2}{N} = \frac{\sum x_i^2}{N} - (\bar{x})^2$$

$$=\frac{\sum F_i(x_i-\bar{x})^2}{N}=\frac{\sum F_i x_i^2}{N}-\left(\frac{\sum F_i x_i}{N}\right)^2$$

$$=\sum f_i(x_i-\bar{x})^2=\boxed{\sum f_i x_i^2-\left(\sum f_i - x_i\right)^2}$$

۸) نکته ۲ در مورد واریانس دادههای یک جامعه است در صورتیکه، واریانس یک **نمونه** nتائی برخلاف نکته۲ از رابطه زیر بدست می‌آید:

$$\boxed{\delta_x^2=\frac{\sum(x_i-\bar{x})^2}{N-1}}$$

۳)

$$\text{واریانس}\begin{cases}\delta^2(x\pm a)=V(x\pm a)=V(x)=\delta_x^2\\ \delta^2(bx\pm a)=b^2\delta_x^2\end{cases}$$

۴)

$$\text{انحراف معیار}\begin{cases}\delta(x\pm a)=\delta_x\\ \delta(bx\pm a)=b\delta_x\end{cases}$$

۵) اگر همه متغیرها با هم برابر باشند یا اعداد ثابت داشته باشیم، واریانس و انحراف معیار صفر است.

۶) اگر در دادههایی که می‌خواهیم واریانس آن را محاسبه کنیم قسمت مشترکی وجود داشته باشد. برای محاسبه واریانس کافی است آن اعداد را به غیر از قسمت مشترک در نظر بگیریم.

مثال: مطلوبست واریانس اعداد 25474 , 25473 , 25472 , 25471 , 25470 ؟ در این دادهها قسمت 2547 مشترک است، بنابراین کافی است واریانس اعداد 0 و 1 و 2 و 3 و 4 را حساب کنیم.

$$\sigma^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}=\frac{(0-2)^2+(1-2)^2+(2-2)^2+(3-2)^2+(4-2)^2}{5}=2$$

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{0+1+2+3+4}{5}=2$$

۷) واریانس اعداد 1، 2، 3، ....، N برابر است با $\boxed{\frac{N^2-1}{12}}$

مثال: مطلوبست واریانس اعداد 1 و 2 و 3 و 4؟ $\frac{N^2-1}{12}=\frac{4^2-1}{12}=\frac{15}{12}=\frac{3}{4}$

مثال: واریانس دادهها با جدول فراوانی کدام است؟

x_i	-1	0	1	2
F_i	2	3	4	1

۱) ۰/۷۶ ۲) ۰/۷۸ ۳) ۰/۸۲ ۴) ۰/۸۴

حل:

$$\delta^2=\frac{\sum F_i x_i^2}{N}-\left(\frac{\sum F_i x_i}{N}\right)^2$$

$$\delta^2 = \frac{2(1)^2+3(0)^2+4(1)^2+1(2)^2}{2+3+4+1} - \left(\frac{2(-1)+3(0)+4(1)+1(2)}{2+3+4+1}\right)^2 = \frac{10}{10} - \left(\frac{4}{10}\right)^2 = \frac{10}{10} - \frac{16}{100} = 0.84$$

مثال: حقوق پرداختی به کارمندان شرکت آلفا به طور متوسط ۱۵ هزار تومان با انحراف معیار ۳ هزار تومان است. اگر ۲۰٪ میانگین به حقوق کارمندان اضافه شود، به ترتیب میانگین و انحراف معیار حقوق پرداختی چقدر خواهد شد؟

۱) ۱۵/۳، ۳ ۲) ۱۵/۳، ۳/۶ ۳) ۱۸، ۳ ۴) ۱۸، ۳/۶

حل:

واریانس آن تغییری نمی‌کند اما به میانگین حقوق کارمندان اضافه می‌شود. زمانی که ۲۰٪ میانگین به حقوق کارمندان اضافه شود داریم: $3 = \frac{20}{100} \times 15$؛ لذا به میانگین ۳ هزار تومان اضافه شده و میانگین حقوق به ۱۸ هزار تومان می‌رسد و انحراف معیار بدون تغییر باقی خواهد ماند.

$$E\overline{(x+3)} = E(\overline{x}) + E(3) = E(\overline{x}) + 3 = 15+3 = 18$$

$$\delta_{x+3} = \delta_x = 3000$$

مثال: اگر μ_x میانگین x_1، x_2، ...، x_N باشد. واریانس مشاهدات $\left(-\frac{x_1}{2}+3\right)$، $\left(-\frac{x_2}{2}+3\right)$، ...، $\left(-\frac{x_N}{2}+3\right)$ کدام است؟

۱) $-\frac{1}{2}\sigma_x^2$ ۲) $-\frac{1}{2}\sigma_x^2+3$ ۳) $\frac{1}{4}\sigma_x^2+3$ ۴) $\frac{1}{4}\sigma_x^2$

حل:

علامت ضریب، منفی است که وقتی بیرون از واریانس می‌آید، مثبت می‌شود.

$$\mathrm{var}\left(\frac{-x}{2}+3\right) = \frac{1}{4}\mathrm{var}(x) = \frac{1}{4}\delta^2 x$$

مثال: جدول مقابل توزیع فراوانی فروش یک شرکت را نشان می‌دهد. میانگین و انحراف معیار فروش به ترتیب چقدر است؟

۱) ۳۵، ۸/۶
۲) ۳۶، ۵/۷
۳) ۳۶، ۷
۴) ۳۵، ۹

فروش به هزار تومان	تعداد روزها
۲۰ تا کم‌تر از ۳۰	۱۰
۳۰ تا کم‌تر از ۴۰	۲۵
۴۰ تا کم‌تر از ۵۰	۱۵

حل:

C-L	F_i	x_i	$F_i x_i$	$F_i x_i^2$
20-30	10	25	250	6250
30-40	25	35	875	30625
40-50	15	45	675	30375

x_i (مرکز طبقات) $= \frac{\text{حد بالا} + \text{حد پایین}}{۲}$

$$\overline{x} = \frac{\Sigma F_i x_i}{N} = \frac{(10\times25)+(25\times35)+(15\times45)}{10+25+15} = \frac{250+875+675}{50} = \frac{1800}{50} = 36$$

$$var(x)=\delta_x^2=\frac{\Sigma F_i x_i^2}{N}-\left(\frac{\Sigma F_i x_i}{N}\right)^2=\frac{67250}{50}-\left(\frac{1800}{5}\right)^2=1345-1296=49$$

$$\delta_x^2=49\rightarrow\delta_x=7$$

مثال: با توجه به جدول زیر، سرمایه‌گذاری در کدام شرکت مناسب‌تر است؟

شاخص	الف	ب	ج
میانگین سود	۷	۵	۷
انحراف معیار سود	۲	۱	۲
ضریب چولگی پیرسون	۴	۰	-۴

۱) الف

۲) ب

۳) ج

۴) تفاوتی ندارد

حل:

سرمایه‌گذاری در شرکت ب مناسب‌تر است؛ چرا که دارای کمترین انحراف معیار بوده و چولگی آن صفر می‌باشد.

مثال: در صورتی که انحراف معیار ۱۲ عدد مساوی ۲/۴ باشد و به هر یک از اعداد این توزیع، عدد ۴ را اضافه کنیم، انحراف معیار جدید چقدر خواهد شد؟

۱) $\sqrt{۴}$ ۲/۴ ۲) ۹/۶ ۳) ۶/۴ ۴) ۲/۴

حل:

$$var(x+4)=var(x)$$

$$\delta(x+4)=\delta_x=2.4$$

مثال: دستگاه A در اندازه‌گیری مکرر از شیء واحدی دارای واریانس $\sigma^2=9$ بوده است. دستگاه B در اندازه‌گیری مکرر از همان شیء دارای واریانس $\sigma^2=25$ بوده است.

۱) دستگاه A دقیق‌تر است.

۲) دستگاه B دقیق‌تر است.

۳) دستگاه A اندازه‌گیری‌های بزرگ‌تری از دستگاه B به دست می‌دهد.

۴) دستگاه B اندازه‌گیری‌های بزرگ‌تری از دستگاه A به دست می‌دهد.

حل:

در بین دو برآوردکننده، برآوردکننده‌ای که واریانس کمتری داشته باشد کاراتر و بهتر است. $\delta_A^2=9<\delta_B^2=25$

مثال: کدام یک از پارامترهای زیر بیش‌تر تحت تأثیرات انحرافات بزرگ است؟

۱) واریانس ۲) نیم‌دامنه ۳) انحرافات چارکی ۴) انحراف متوسط از میانگین

حل:

واریانس بیش از سایر پارامترها تحت تأثیر انحرافات بزرگ قرار دارد.

مثال: میانگین قد دانش‌آموزان مدرسه‌ای ۱۲۰ سانتی‌متر با واریانس ۱۰۰ است. اگر هر فرد ۱۴٪ قدش در سال آینده بلند شود، میانگین و واریانس قد آن‌ها در سال آینده (از راست به چپ) چقدر خواهد بود؟

۱) ۱۲۰ و ۱۰۰
۲) ۱۱۴ و ۱۲۰
۳) ۱۱۴ و ۱۳۶/۸
۴) ۱۳۶/۸ و ۱۲۹/۹۶

حل:

$$E(x+\%14x)=E(1.14x)=1.14E(x)$$

$$=1.14\times 120=136.8$$

$$\operatorname{var}(x+0.14x)=\operatorname{var}(1.14x)=(1.14)^2\operatorname{var}(x)=1.2996\times 100=129.96$$

میانگین و واریانس کل چند جامعه آماری:

اگر K جامعه با تعداد مشاهدات N_1 و $N_2,\ldots,N_k$ و با میانگین‌های $\mu_1,\mu_2,\ldots,\mu_k$ باشد و دارای واریانس‌های $\delta_1^2,\delta_2^2,\ldots,\delta_k^2$ باشد آنگاه میانگین و واریانس کل داده‌های آماری:

$$\boxed{\mu=\frac{N_1\mu_1+N_2\mu_2+\ldots+N_k\mu_k}{N_1+N_2+\ldots+N_k}}$$

و) ضریب پراکندگی (ضریب تغییرات) (CV):

در شرایط زیر از معیار پراکندگی ضریب پراکندگی استفاده می‌کنیم که میانگین و واریانس فاقد آن هستند:

مهم
۱) دو یا چند جامعه در مقایسه با هم دارای مشاهدات ناهمگون از نظر واحد اندازه‌گیری می‌باشند (مانند: یک جامعه بر حسب متر و یک جامعه بر حسب اینچ)
گاهی اوقات صفات یکسان است ولی بزرگی مشاهدات به طور قابل ملاحظه‌ای تفاوت دارد.
۲) زمانی که دو یا چند جامعه دارای میانگین‌های متفاوتی باشند.

مثال: برای تعیین آن‌که در ۳۰ روز گذشته، به نسبت، قیمت دلار از ثبات بیش‌تری برخوردار بوده است یا یورو، استفاده از کدام شاخص آماری مناسب‌تر است؟

۱) انحراف متوسط
۲) ضریب پراکندگی
۳) ضریب چولگی
۴) واریانس

حل:

زمانی که دو یا چند جامعه در مقایسه با هم دارای مشاهدات ناهمگون از نظر واحد اندازه‌گیری باشند، یا زمانی که دو یا چند جامعه دارای میانگین‌های متفاوتی باشند از ضریب تغییرات (ضریب پراکندگی) استفاده می‌کنیم که میانگین و واریانس فاقد آن هستند.

نکته مهم: مقدار ضریب پراکندگی از فرمول زیر بدست می‌آید:

$$CV=\frac{\delta}{\bar{x}}$$

$$\text{درصد } CV = \frac{\delta}{\bar{x}} \times 100$$

مثال: میانگین ۲۰ داده آماری ۱۵ و واریانس آن‌ها برابر ۲/۲۵ است. درصد ضریب تغییرات آن‌ها چقدر است؟

۱) ۱۰ ۲) ۱۲ ۳) ۱۵ ۴) ۲۰

حل:

$$CV = \frac{\delta}{\bar{x}} = \frac{\sqrt{S^2}}{\bar{x}} = \frac{\sqrt{2.25}}{15} = \frac{1.5}{15} = 0.1$$

$$\text{درصد ضریب تغییرات} = CV \times 100 = 0.1 \times 100 = 10$$

نکته مهم:

$$CV_{ax} = CV_x$$

$$CV_{x+a} = \frac{\delta_{(x+a)}}{E(x+a)} = \frac{\delta_x}{\mu_x + a}$$

مثال: ضریب تغییرات (Coefficient of Variation) عدد ۵ برابر است با:

۱) ۵ ۲) ۱ ۳) ∞ ۴) ۰

حل:

واریانس عدد ثابت صفر است پس انحراف معیار آن نیز صفر خواهد شد. همچنین میانگین یا امید ریاضی هر عدد نیز مساوی خود عدد می‌باشد، بنابراین:

$$CV = \frac{\delta}{\bar{x}} = \frac{0}{5} = 0$$

مثال: با فرض این که داشته باشیم $\sum_{i=1}^{3} X_i = 3$ و $\sum_{i=1}^{3} X_i^2 = 6$ ، ضریب تغییرات کدام است؟

۱) ۰/۵ ۲) ۱ ۳) ۱/۵ ۴) ۲

حل:

$$\bar{x} = \frac{\sum x_i}{N} = \frac{3}{3} = 1$$

$$\delta^2 = \frac{\sum x_i^2}{N} - \bar{x}^2 = \frac{6}{3} - 1 = 1$$

$$CV = \frac{\delta_x}{\bar{x}} = \frac{\sqrt{\delta_x^2}}{\bar{x}} = \frac{1}{1} = 1$$

مثال: میانگین سن یک گروه ۱۲ سال و ضریب تغییرات سن آنان ۲۰ درصد است. انحراف معیار سن آنان چقدر است؟

۱) ۰/۶ ۲) ۲/۴ ۳) ۶۰ ۴) ۲۴۰

حل:

$$CV = \frac{\delta}{\overline{x}} \to \delta = CV \times \overline{x} = \frac{20}{100} \times 12 = 2.4$$

مثال: اگر میانگین و انحراف معیار متغیر تصادفی x به ترتیب ۱ و ۲ باشد، درصد ضریب تغییرات y=x+3 چقدر است؟

۱) ٪۲۵ ۲) ٪۵۰ ۳) ٪۷۵ ۴) ٪۱۰۰

حل:

$$\delta_{x+3} = \delta_x$$

$$\mu_{x+3} = \mu_x + 3$$

$$CV = \frac{\delta}{\overline{x}} = \frac{\delta_x}{\mu_x + 3} = \frac{2}{1+3} = \frac{2}{4} = \frac{1}{2} = 50\%$$

مثال: اگر در جامعه‌ای که مقدار ضریب تغییرات آن ۷۳/۱ درصد بدست آمده است، کلیه متغیرها را در عدد ثابت ۱۰ ضرب کنیم ضریب تغییرات چقدر خواهد بود؟

۱) ٪۷/۳۱ ۲) ٪۷۳/۱ ۳) ٪۷۳۱ ۴) نمی‌توان تعیین کرد

حل:

گزینه (۲) صحیح است. زیرا ضریب ، CV را تغییر نمی‌دهد.

مثال: اگر $N = 100$ و $\sum x_i = 200$ و $\sum x_i^2 = 500$ باشد، مقدار ضریب تغییرات کدام است؟

۱) ٪۵۰ ۲) ٪۹۰ ۳) ٪۱۰ ۴) ٪۲۰

حل:

$$\overline{x} = \frac{\sum x_i}{N} = \frac{200}{100} = 2$$

$$var(x) = \frac{\sum x_i^2}{N} - \overline{x}^2 = \frac{500}{100} - 4 = 1$$

$$CV = \frac{\delta}{\overline{x}} = \frac{1}{2} \times 100 = \%50$$

مثال: متوسط درآمد ماهانه کارگران کارخانه A ، ۱۷ هزار تومان با واریانس ۴ می‌باشد. در کارخانه B متوسط درآمد ماهانه ۲۵۰ هزار ریال با واریانس ۹۰۰ می‌باشد.

۱) اختلاف درآمد در کارخانه A بیش از کارخانه B است.

۲) اختلاف درآمد در کارخانه B بیش از کارخانه A است.

۳) درآمدهای اکثر افراد کارخانه A کم‌تر از اکثر افراد کارخانه B است.

۴) کم‌ترین درآمد در کارخانه A بیش از کارخانه B است.

حل:

واحدهای اندازه‌گیری و میانگین‌ها برابر نیستند بنابراین برای مقایسه بین دو جامعه از ضریب تغییرات استفاده می‌کنیم.

$$CV_A = \frac{\delta_A}{\overline{x}_A} = \frac{2}{17} \times 100 = 11.76$$

$$CV_B = \frac{\delta_B}{\bar{x}_B} = \frac{30}{250} \times 100 = 12$$

و چون $CV_A < CV_B$ است یعنی پراکندگی در کارخانه B بیشتر است و اختلاف درآمد در کارخانه B بیش از کارخانه A است.

مثال: در جامعه‌ای به حجم $N = 50$ برای صفت متغیر X کمیت‌های زیر به دست آمده است:

$$\sum x_i = 250 \quad , \quad \sum x_i^2 = 2500$$

ضریب تغییرات صفت X کدام است؟

۱) CV= ٪۲۵ ۲) CV= ٪۵۰ ۳) CV= ٪۷۵ ۴) CV = ٪۱۰۰

حل:

$$CV = \frac{\delta}{\bar{x}}$$

$$\delta^2 = \frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right) = \frac{2500}{50} - \left(\frac{250}{50}\right)^2 = 50 - (5)^2 = 50 - 25 = 25$$

$$\delta^2 = 25 \rightarrow \delta = 5$$

$$\bar{x} = \frac{\sum x_i}{n} \rightarrow \bar{x} = \frac{250}{50} = 5$$

در نتیجه:

$$CV = \frac{5}{\bar{x}} \times 100 = \frac{5}{5} \times 100 = \%100$$

مثال: میانگین و انحراف معیار حقوق در یک سازمان به ترتیب ۵۰ هزار تومان و ۲۰ هزار تومان است اگر حقوق‌ها در این سازمان ۲۵ درصد افزایش یابند، ضریب تغییرات حقوق چه خواهد شد؟

۱) نصف خواهد شد ۲) تغییر نخواهد کرد ۳) چهار برابر خواهد شد ۴) ۲۵ درصد افزایش خواهد یافت

حل:

$$E(x + 0.25x) = E(1.25x) = 1.25E(x)$$

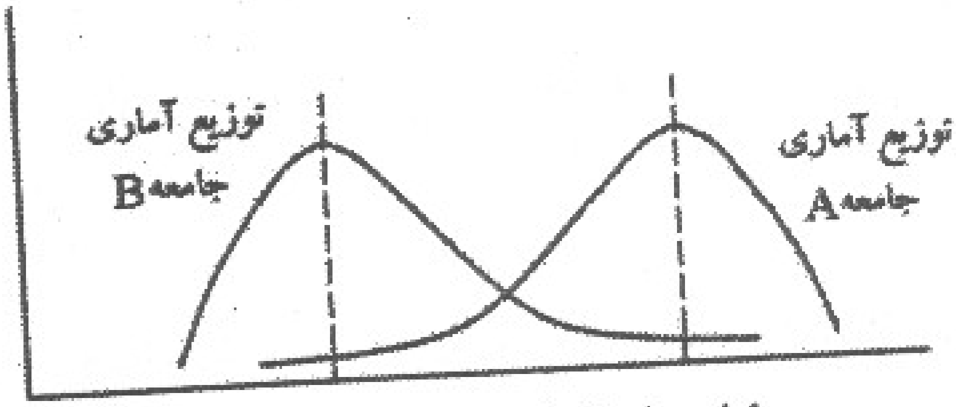
$$\delta(x + 0.25x) = \delta(1.25x) = 1.25\delta(x)$$

بنابراین تغییری نمی‌کند.

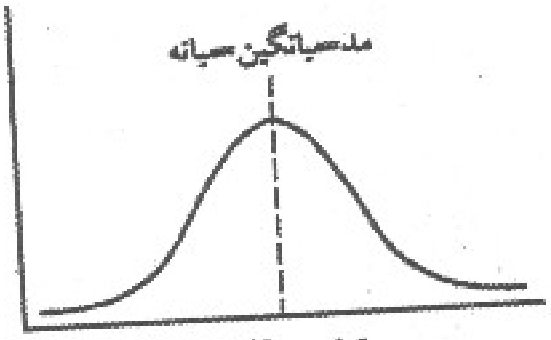
$$CV = \frac{\delta}{\bar{x}} = \frac{1.25\delta}{1.25E(x)} = \frac{\delta}{\bar{x}}$$

چولگی (عدم قرینگی)

در مقایسه دو یا چند جامعه با یکدیگر ابتدا از پارامترهای مرکزی استفاده می‌شود. اما در صورت تساوی برخی از پارامترهای مرکزی مانند میانگین، اختلاف جوامع آماری به کمک شاخص‌های پراکندگی مانند انحراف معیار مشخص می گردد. گاهی پارامترهای پراکندگی نیز به علت مساوی بودن جوابگو نیستند. برای مثال در شکل روبرو دو توزیع دارای میانگین و واریانس مساوی هستند، اما توزیع یکسانی ندارند. توزیع جامعه A دارای تراکم در حول وحوش مبدأ مختصات است، در حالی که مد جامعه B در نقطه مقابل آن قرار دارد. این تفاوت را چولگی (skewness) یا انحراف از قرینگی می گویند. چولگی توزیع‌ها در مقایسه با توزیع متقارن معین می‌شود. لازم به ذکر است که پارامترهای مرکزی (مد، میانه و میانگین) در توزیع متقارن با یکدیگر برابرند. هر چه یک توزیع با توزیع متقارن تفاوت بیشتری داشته باشد، انحراف از قرینگی آن بیشتر خواهد بود.



توزیع های نامتقارن جوامع B,A

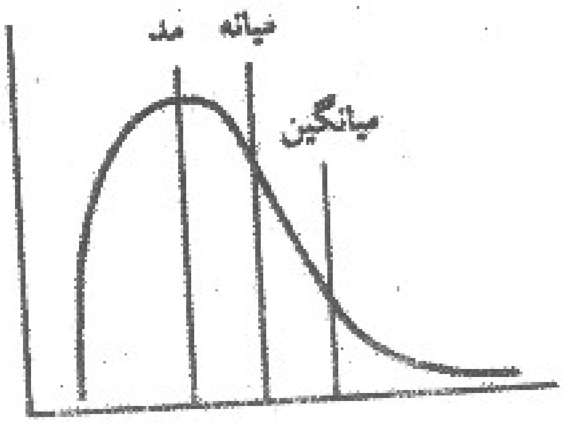


توزیع متقارن

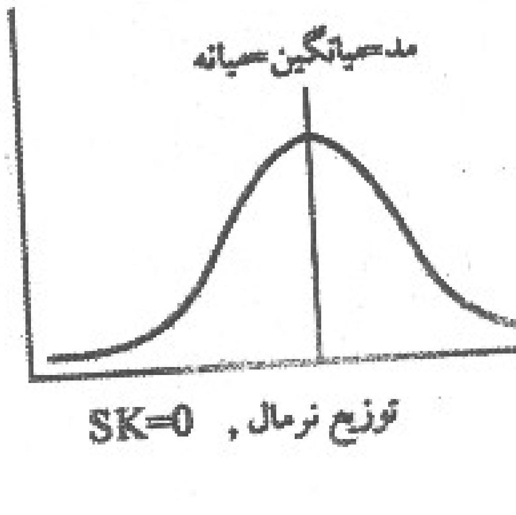
در صورتی که بین نمودارهای مختلف به دنبال میزان عدم تقارن یا عدم قرینگی باشیم، می‌توانیم از معیاری به نام چولگی استفاده کنیم. این معیار بدون واحد است و از فرمول زیر بدست می‌آید:

$$\text{(ضریب چولگی)}\ SK=\frac{\mu_3}{\delta^3}=\frac{\dfrac{\sum_{i=1}^{N}(x_i-\mu)^3}{N}}{\left(\dfrac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}\right)^{\frac{3}{2}}}=\frac{\dfrac{\sum_{i=1}^{N}F_i(x_i-\mu)^3}{N}}{\left(\dfrac{\sum_{i=1}^{N}F_i(x_i-\mu)^2}{N}\right)^{\frac{3}{2}}}=\frac{\sum_{i=1}^{N}f_i(x_i-\mu)^3}{\left(\sum_{i=1}^{N}f_i(x_i-\mu)^2\right)^{\frac{3}{2}}}$$

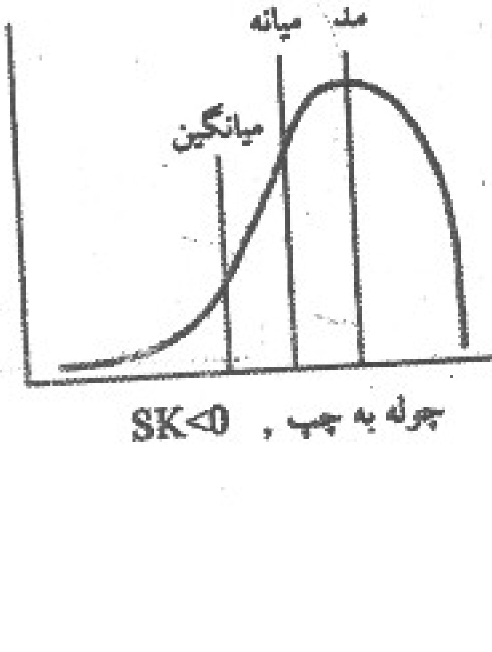
حال برای بررسی میزان چولگی با توجه به عدد بدست آمده شکل‌های زیر وجود دارند:



چوله به راست , SK>0



توزیع نرمال , SK=0



چوله به چپ , SK<0

نکته: دو فرمول مهم، به نام فرمول‌های پیرسون به صورت زیر وجود دارد:

$$SK_1 = \frac{(\bar{x} - MO)}{\delta}$$: فرمول اول پیرسون

$$SK_2 = \frac{3(\bar{x} - Md)}{\delta}$$: فرمول دوم پیرسون

زمانی که چولگی توزیع متعادل یا متناسب باشد، فرمول‌های پیرسون با یکدیگر برابر شده و رابطه زیر را به وجود می‌آورند:

$$\boxed{\bar{x} - MO = 3(\bar{x} - Md)}$$

مثال: در جامعه‌ای به حجم $n = 20$ پس از محاسبات لازم کمیت‌های زیر به دست آمده است:

$$\sum(X_i - \bar{X})^2 = 1620 \quad , \quad \sum(X_i - \bar{X})^3 = -328$$

ضریب چولگی توزیع کدام است؟

۱) ٪۴۴۹
۲) ـ٪۴۴۹
۳) ۰/۰۲۲
۴) ـ۰/۰۲۲

حل:

$$\text{ضریب چولگی} = \frac{\mu_3}{\delta^3}$$

$$\mu_3 = \frac{\sum(x_i - \bar{x})^3}{N} = \frac{-328}{20} = -16/4$$

$$\delta_x^2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{1620}{20} = 81 \qquad \delta_x^2 = var(x) = 81 \to \delta_k = \sqrt{81} = 9$$

$$\text{ضریب چولگی} = \frac{\mu_3}{\delta^3} = \frac{-16.4}{(9)^3} = \frac{-16.4}{729} = -0.022$$ با توجه به منفی بودن ضریب چولگی، چوله به چپ است.

مثال: در صورتی که جامعه‌ای دارای چولگی مثبت باشد:

۱) میانه در وسط، میانگین سمت راست و نما سمت چپ آن قرار دارد.
۲) میانگین در وسط، میانه سمت راست و نما سمت چپ آن قرار دارد.
۳) میانه در وسط، نما سمت چپ و میانگین سمت چپ آن قرار دارد.
۴) میانگین در وسط، نما سمت چپ و میانه سمت چپ آن قرار دارد.

حل:

در جامعه‌ای با چولگی مثبت، همواره رابطه $Mo < Md < \bar{x}$ برقرار است.

مثال: در یک توزیع متمایل به راست، کدام گزینه صحیح است؟

۱) $\mu_x > Md > Mo$
۲) $\mu_x > Mo < Md$
۳) $Mo > Md > \mu_x$
۴) $Mo > Md < \mu_x$

حل:

زمانی که تمایل توزیع به سمت راست است بدین معناست که توزیع چوله به چپ می‌باشد بنابراین گزینه(۳) صحیح می‌باشد.

مثال: در توزیعی با چولگی منفی انتظار می‌رود که .... کم‌ترین مقدار را داشته باشد.

۱) دامنه تغییرات ۲) میانه ۳) میانگین ۴) نما

حل:

در جامعه‌ای با چولگی منفی همواره رابطه $\bar{x} < Md < Mo$ برقرار است. بنابراین گزینه (۳) صحیح می‌باشد.

مثال: در یک توزیع با چولگی خفیف، میانگین حسابی $\bar{X} = ۵۲/۴$ و میانه $Me = ۵۱/۸$ به دست آمده است.مد توزیع کدام است؟

۱) ۵۳/۶ ۲) ۵۰/۶ ۳) ۵۴/۲ ۴) ۵۱/۶

حل:

اگر توزیعی دارای چولگی خفیف باشد، همواره رابطه $\bar{x} - Mo = 3(\bar{x} - Md)$ بین سه مشخص‌کننده مرکزی برقرار است. بنابراین:

$$(52/4 - Mo) = 3(52/4 - 51/8) \rightarrow Mo = 50/6$$

مثال: اگر $N = 10$ ، $\sum (x_i - \mu_x)^2 = 40$ و $\sum (x_i - \mu_x)^3 = 80$ باشد، مقدار ضریب چولگی چقدر است؟ (مدیریت ـ ۷۹)

۱) ۳- ۲) ۳ ۳) ۲ ۴) ۱

حل:

$$SK = \frac{\mu_3}{\delta^3} = \frac{\frac{\sum (x_i - \mu)^3}{N}}{\sqrt{\left(\frac{\sum (x_i - \mu)^2}{N}\right)^3}} = \frac{\frac{80}{10}}{\sqrt{\left(\frac{40}{10}\right)^3}} = \frac{8}{8} = 1$$

تحلیل ضریب چولگی:

قدرمطلق ضریب چولگی نشاندهنده میزان اختلاف جامعه آماری با توزیع نرمال از نظر قرینگی است. بدیهی است هر چه $|SK|$ بزرگتر باشد، تفاوت جامعه از نظر قرینگی با توزیع نرمال بیشتر خواهد بود. طوری که:

تقریباً چولگی وجود ندارد و جامعه از نظر قرینگی تقریباً نرمال است. $|SK| \le 0.1$

چولگی موجود، اما اندک. در حقیقت جامعه از نظر قرینگی دارای تفاوت اندکی با توزیع نرمال است. $0.1 < |SK| \le 0.5$

چولگی زیاد است. به عبارت دیگر جامعه از نظر قرینگی دارای تفاوت فاحشی با توزیع نرمال است. $|sk| > 0.5$

مثال: اگر ضریب چولگی توزیع یک جامعه ۰/۶۶ ـ باشد، آن‌گاه جامعه مورد مطالعه:

۱) نرمال است. ۲) با جامعهٔ نرمال تفاوت مختصری دارد.

۳) با جامعهٔ نرمال تفاوت فاحش دارد. ۴) با اطلاعات داده شده نمی‌توان قضاوت کرد.

حل:

با توجه به تعریف بالا گزینه (۳) صحیح است.

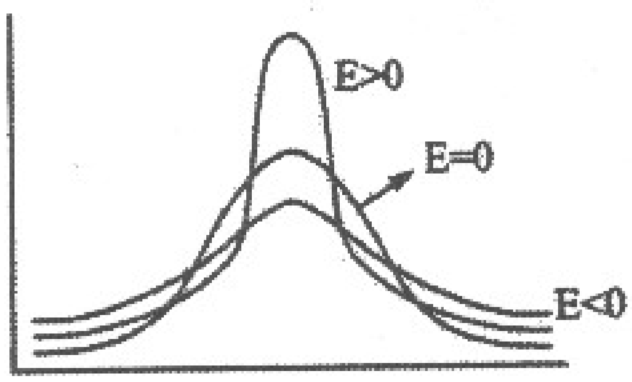
کشیدگی و ضریب کشیدگی

کشیدگی معیاری است بدون واحد که ارتفاع منحنی را مورد بحث قرار می‌دهد و در ارتباط مستقیم با پراکندگی است. برای محاسبه کشیدگی توزیع‌ها از رابطه $\frac{\mu_4}{\delta^4}$ استفاده می‌شود، به فرم زیر:

$$E=\frac{\mu_4}{\delta^4}=\frac{\frac{\sum(x_i-\mu)^4}{N}}{(\text{واریانس})^2}=\frac{\frac{\sum Fi(x_i-\mu)^4}{N}}{(\text{واریانس})^2}$$

نکته مهم: کشیدگی توزیع نرمال همیشه ۳ است و کشیدگی چندکی توزیع نرمال ۰/۲۶۳ می‌باشد. لازم به ذکر است که معیار ضریب کشیدگی اختلاف کشیدگی توزیع‌ها را نسبت به کشیدگی توزیع نرمال بدست می‌آورد.

$$E=\frac{\mu_4}{\sigma^4}-3 \quad (\text{ضریب کشیدگی})$$



چنانچه داده‌ها دارای طبقه‌بندی مشخص و کمی باشند، بهترین روش برای محاسبه کشیدگی استفاده از گشتاورهاست. همان‌طور که گفته شد کشیدگی گشتاوری عبارت است از:

$$E=\frac{\mu_4}{\delta^4}$$

در این رابطه μ_4 گشتاور مرتبه چهارم به مبدأ μ_x و عدد ثابت ۳ نشاندهنده کشیدگی گشتاوری توزیع نرمال است. بنابراین $\frac{\mu_4}{\delta^4}$ نشاندهنده کشیدگی هر توزیع دلخواه است.

تحلیل ضریب کشیدگی:

چنانچه $E=0$ باشد، کشیدگی توزیع هم اندازه و هم ارتفاع توزیع نرمال است.

چنانچه $E>0$ باشد، کشیدگی توزیع از نرمال بلندتر و پراکندگی آن از نرمال کمتر است.

چنانچه $E<0$ باشد، کشیدگی توزیع از نرمال کوتاه‌تر و پراکندگی آن از نرمال بیشتر است.

قدرمطلق ضریب کشیدگی میزان اختلاف ارتفاع را با توزیع نرمال بیان می‌کند:

- اگر $|E|\le 0.1$ باشد، توزیع جامعه از نظر پراکندگی تقریباً نرمال است.
- اگر $0.1<|E|\le 0.5$ باشد، توزیع از نظر کشیدگی دارای تفاوت اندکی با توزیع نرمال است.
- اگر $|E|>0.5$ باشد، تفاوت توزیع با توزیع نرمال به لحاظ پراکندگی فاحش است و اختلاف ارتفاع زیادی با توزیع نرمال دارد.

گشتاورها

گشتاورها به سه دسته تقسیم می‌شوند:

۱ـ گشتاور حول مبدأ

۲ـ گشتاور حول میانگین (مرکزی)

۳ـ گشتاور حول نقطه دلخواه

۱ـ گشتاور حول مبدأ

گشتاور nام حول مبدا با پیش فرض $F_i = 1$ به فرم زیر محاسبه می‌شود:

$$m_n = \frac{\sum_{i=1}^{N} X_i^n}{N} = \sum_{i=1}^{n} \frac{F_i x_i^N}{N} = \sum_{i=1}^{N} f_i x_i^N$$

بنابراین :

$$\text{میانگین} = \bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} = m_1 = \text{گشتاور حول مبدأ}$$

می‌دانیم:

$$N = \sum_{i=1}^{N} F_i$$

$$m_1 = \frac{\sum_{i=1}^{n} x_i}{N} = \frac{\sum_{i=1}^{N} F_i x}{N} = \sum_{i=1}^{N} f_i x_i = \bar{x} = E(x) = \mu$$

در حالت کلی می‌توان به این نتیجه رسید که:

$$m_n = E(x^n)$$

۲ـ گشتاور حول میانگین

$$\mu_n = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^n}{N} = \frac{\sum_{i=1}^{N} F_i (x_i - \bar{x})^n}{N} = \sum_{i=1}^{N} f_i (x_i - \bar{x})^n$$

به خاطر داشته باشید که همواره: $\begin{cases} \mu_1 = 0 \\ \mu_2 = \delta_x^2 = \operatorname{var}(x) \end{cases}$ می‌باشد.

$$\delta_x^2 = \frac{\sum (x_i - \bar{x})^2}{N} = \frac{\sum x_i^2}{N} - 2\sum \frac{x_i \bar{x}}{N} + \frac{\sum_{i=1}^{N} \bar{x}^2}{N} = \frac{\sum x_i^2}{N} - 2\bar{x}^2 + \bar{x}^2 = \boxed{\frac{\sum x_i^2}{N} - \bar{x}^2 = E(x^2) - E(x)^2}$$

بین گشتاور مرکزی و گشتاور حول مبدا همیشه رابطه مهم زیر وجود دارد:

$$\mu_n = (m - m_1)^n$$

در نتیجه:

$$\mu_1 = (m - m_1)^1 = m_1 - m_1 = 0$$

$$\mu_2 = (m - m_1)^2 = m_2 - 2m_1 m_1 + m_1^2 = m_2 - m_1^2 = E(x^2) - E(x)^2$$

مثال: در جامعه‌ای به حجم $n=10$ کمیت‌های زیر محاسبه شده است:

$$\sum F_i x_i = 40, \sum F_i x_i^2 = 400, \sum F_i x_i^3 = 600$$

گشتاوری مرکزی مرتبه سوم (μ_3) کدام است؟

۱) ۲۹۲- ۲) ۲۹۲ ۳) ۵۴۸- ۴) ۴۸۵

حل:

با توجه به آن‌که بین گشتاور مرکزی و گشتاور حول مبدا رابطه $\mu_n = (m - m_1)^n$ برقرار است داریم:

$$\mu_3 = (m - m_1)^3 = m_3 - m_1^3 - 3m_2m_1 + 3m_1m_1^2 = m_3 - m_1^3 - 3m_2m_1 + 3m_1^3 = m_3 + 2m_1^3 - 3m_2m_1$$

$$m_1 = \frac{\sum F_i x_i}{N} = \frac{40}{10} = 4 \qquad m_3 = \frac{\sum F_i x_i^3}{N} = \frac{600}{10} = 60$$

$$m_2 = \frac{\sum F_i x_i^2}{N} = \frac{400}{10} = 40$$

بنابراین می‌توان گفت:

$$\mu_3 = m_3 + 2m_1^3 - 3m_2m_1 = 60 + 2(4)^3 - 3(40)(4) = 60 + 128 - 480 = -292$$

مثال: اگر $\sum F_i (x_i - \mu_x)^4 = 5000, N = 1000$ و انحراف معیار جامعه ۲ باشد، مقدار ضریب کشیدگی کدام است؟

۱) ۲/۶۹- ۲) ۲/۵۳ ۳) ۰/۳۱ ۴) ۲/۵۳

حل:

$$E = \frac{\mu_4}{\delta^4} - 3 = \frac{\frac{\sum F_i (x_i - \mu)^4}{N}}{(\mathrm{var}_x)^2} - 3$$

$$\delta_x = 2 \rightarrow \mathrm{var}(x) = 4$$

$$E = \frac{\frac{5000}{1000}}{(4)^2} - 3 = -2.69$$

نمایش هندسی مشاهدات:

- برای نمایش توزیع‌های فراوانی، اغلب از نمودارها استفاده می‌شود.
- برای نمایش داده‌ها با توجه به نوع مقیاس داده‌ها، روش‌های مختلفی وجود دارد:

۱ـ نمودارهای کمی برای مقیاس‌های نسبی و فاصله‌ای استفاده می‌شود.

۲ـ نمودارهای کیفی برای مقیاس‌های اسمی یا رتبه‌ای استفاده می‌شود.

نکته مهم: مهمترین نمودارهای کمی و کیفی به شرح زیر هستند.

نمودارهای کمی:
- ۱) نمودار بافت‌نگار (Histogram chart)
- ۲) نمودار چندضلعی
- ۳) نمودار فراوانی تجمعی (اجایو) (ogive) (cumulative frequency chart)
- ۴) نمودارهای تحلیل اکتشافی داده‌ها
  - a) شاخه و برگ
  - b) جعبه‌ای (box)

نمودارهای کیفی : (وصفی)
۱) نمودار ستونی (میله‌ای) (bar chart)
۲) نمودار دایره‌ای (pie Chart)
۳) نمودار پاتو (pareto chart)

نمودار بافت‌نگار:

پس از خلاصه کردن مجموعهٔ بزرگ داده‌ها به صورت توزیع فراوانی می‌توان آن را به کمک نمودار بافت‌نگار نمایش داد. بافت‌نگار نمایشی مناسبی از الگوی توزیع است. برای رسم این نمودار از فراوانی مطلق یا فراوانی نسبی و حدود واقعی طبقات استفاده می‌شود؛ زیرا حدودی که هنگام طبقه‌بندی برای هر طبقه مشخص می‌شود، ممکن است به علت گرد کردن با حدود واقعی آن تفاوت داشته باشد؛ برای مثال به طبقه‌بندی مشاهدات در جدول زیر، توجه کنید. در این جدول حد پایین طبقه اول ۳۶ میلی‌لیتر در نظر گرفته شده است؛ در حالی که ممکن است حد واقعی ۳۵/۹۵ یا ۳۶/۰۵ بوده باشد. از آن‌جا که طبقه تا ۳۶/۰۵ میلی‌لیتر را در برمی گیرد؛ بنابراین ۳۵/۹۵ خارج از طبقه می‌ماند. به علاوه ممکن است برخی از حجم‌ها تا دو رقم اعشار تعریف شوند؛ مثلاً مقادیر ۳۶/۷۵ و ۳۷/۵۵ میلی‌لیتر در حدود طبقات را تعریف کرد که از این پس این حدود را «حدود کرانه‌ها» می‌نامیم.

جدول ۱ ـ حجم مایع ۷۰ شیشه برحسب میلی‌لیتر

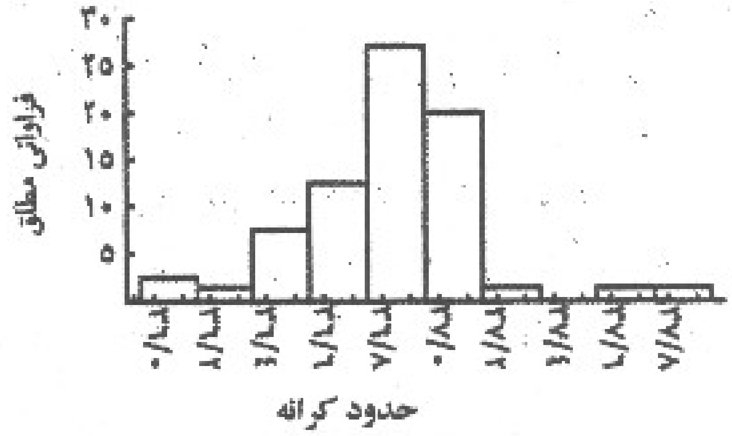
حدود کرانه‌ها	F_i	C−L
۳۵/۹۵ـ۳۶/۱۵	۲	۳۶/۰ـ۳۶/۱
۳۶/۱۵ـ۳۶/۳۵	۱	۳۶/۲ـ۳۶/۳
۳۶/۰۳۵ـ۳۶/۰۵۵	۷	۳۶/۴ـ۳۶/۵
۳۶/۵۵ـ۳۶/۷۵	۱۱	۳۶/۶ـ۳۶/۷
۳۶/۷۵ـ۳۶/۹۵	۲۷	۳۶/۸ـ۳۶/۹
۳۶/۹۵ـ۳۷/۱۵	۱۹	۳۷/۰ـ۳۷/۱
۳۷/۱۵ـ۳۷/۳۵	۱	۳۷/۲ـ۳۷/۳
۳۷/۳۵ـ۳۷/۵۵	۰	۳۷/۴ـ۳۷/۵
۳۷/۵۵ـ۳۷/۷۵	۱	۳۷/۶ـ۳۷/۷
۳۷/۷۵ـ۳۷/۹۵	۱	۳۷/۸ـ۳۷/۹

برای به دست آوردن حدود واقعی توزیع‌های فراوانی به کمک حدود طبقات می‌توان قاعدهٔ سرانگشتی «۵ در مکان بعد از آخرین رقم» را به کار برد. به این ترتیب که ۵ واحد را از محل بعد از آخرین رقم حد پایین طبقه کم می‌کنیم و سپس آن را به مکان بعد از آخرین رقم حد بالای همان طبقه می‌افزاییم؛ بنابراین از روی حدود عنوان شده می‌توانیم به آسانی به حدود کرانه طبقه برسیم؛ برای مثال اگر حدود عنوان شده ۳۶/۱ـ ۳۶/۰ باشد حدود واقعی ۳۶/۱۵ـ ۳۵/۹۵ خواهد بود.

بررسی دقیق‌تر حدود کرانه‌ها نشان می‌دهد که طول وعرض طبقات با یکدیگر مساوی شده است؛ به عبارت دیگر برای به دست آوردن حدود واقعی طبقات به طبقه‌بندی‌ای مشابه طبقه‌بندیِ روش پیوسته رسیده‌ایم؛ بنابراین چنانچه روش طبقه‌بندی از نوع پیوسته باشد، حدود واقعی طبقات همان حدود طبقات خواهند بود.

در نمودار بافت‌نگار، محور افقی دستگاه مختصات با حدود واقعی طبقات و محور عمودی با فراوانی نسبی یا مطلق مدرج می‌گردد. روی کرانه‌های هر طبقه، مستطیلی عمودی رسم می‌کنیم که مساحت آن مساوی با فراوانی نسبی آن طبقه باشد؛ به عبارت ساده‌تر:

$$\text{ارتفاع مستطیل} = \frac{\text{فراوانی نسبی طبقه}}{\text{طول طبقه}}$$



شکل ۱: نمودار بافت نگار فراوانی مطلق حجم مایع بطری‌ها در جدول قبل

بدین ترتیب مساحت هر یک از مستطیل‌های بافت‌نگار نشان‌دهندهٔ نسبت مشاهدات موجود در هر طبقه است که مستطیل بر آن قرار دارد. برای نمایش فراوانی‌های نسبی، استفاده از مساحت مستطیل‌ها به جای ارتفاع آن‌ها فایدهٔ آشکاری دارد. ظاهراً در موقع مقایسه کردن دو قسمت یک بافت‌نگار، یا دو بافت‌نگار مختلف، چشم انسان به طور غریزی مساحت‌ها را با هم مقایسه می‌کند. وقتی دو بافت‌نگار روی طبقات با طول‌های متفاوت بنا شده باشند، این خاصیت که مساحت کل برابر یک است، آن‌ها را قابل مقایسه می‌سازد.

نمودار بافت‌نگار جدول حجم مایع ۷۰ شیشه نشان می‌دهد که بیشترین تراکم در فاصلهٔ ۳۶/۷۵ تا ۳۶/۹۵ است. مقادیر روی مستطیل‌ها نشان‌دهنده فراوانی مطلق هر طبقه است. گفتنی است علت این که در فاصلهٔ ۳۷/۳۵ ـ ۳۷/۵۵ مستطیلی دیده نمی‌شود این است که فراوانی مطلق در آن طبقه صفر است.

نمودار چند ضلعی

نمودار چند ضلعی نموداری است که نقطهٔ میانی هر طبقه روی محور افقی، و فراوانی نسبی یا مطلق هر یک از نقاط میانی روی محور عمودی آن نشان داده می‌شود. متناظر با هر نمایندهٔ طبقه و فراوانی آن، یک نقطه در صفحه مشخص می‌شود که طول آن نماینده طبقه و عرض آن، برابر با فراوانی آن طبقه است. در نتیجه به تعداد طبقات جدول، در صفحه دستگاه نقطه پدید می‌آید. به نقاط مزبور دو نقطهٔ فرضی دیگر اضافه می‌کنیم؛ اولی مرکز طبقهٔ ماقبل اولین طبقه و دیگری نماینده طبقهٔ مابعد آخرین طبقه است. از اتصال متوالی این نقاط به یکدیگر نموداری حاصل می‌شود که آن را چندضلعی می‌نامند.

از نمودارهای چند ضلعی، همانند بافت‌نگارها، برای رسم توزیع‌های فراوانی استفاد می‌شود. تفاوت این دو نمودار آن است که برای نشان دادن رابطهٔ بین سطح‌های متغیر و فراوانی آن متغیر، به جای رسم ستون‌های متصل به هم، از خطوط مستقیم استفاده می‌شود. انتخاب یکی از این دو نمودار اختیاری است. در تحقیقاتی که شامل دو یا چند جامعهٔ آماری هستند و هدف مقایسهٔ آن‌هاست، ترجیح داده می‌شود که توزیع‌های

فراوانی جوامع روی یک نمودار رسم گردد. در چنین حالتی برای سهولت در خواندن و درک اطلاعات، از نمودارهای چندضلعی فراوانی استفاده می‌شود.

نمودار فراوانی تجمعی

توزیع فراوانی تجمعی، توزیعی است که تعداد مشاهدات قبل از یک نقطهٔ معین را در مقیاس مشاهدات نشان می‌دهد. توزیع فراوانی تجمعی پایه لازم برای محاسبه برخی از پارامترهای آماری همانند صدک‌ها، دهک‌ها و چارک‌هاست. نمودار فراوانی تجمعی برای مقایسهٔ هندسی دو یا چند توزیع فراوانی به طور همزمان، به کار می‌رود.

نمودار فراوانی تجمعی بر اساس توزیع فراوانی تجمعی رسم می‌شود. این نمودار به دو روش قابل ترسیم است. درهر دو روش محور عمودی دستگاه مختصات عبارت است از فراوانی‌های تجمعی طبقات، ولی محور افقی در هر روش تعریف خاصی دارد.

در روش اول محور افقی براساس متوسط طبقات مدرج می‌شود. منحنی به دست آمده در این روش را پلی‌گن فراوانی تجمعی گویند.

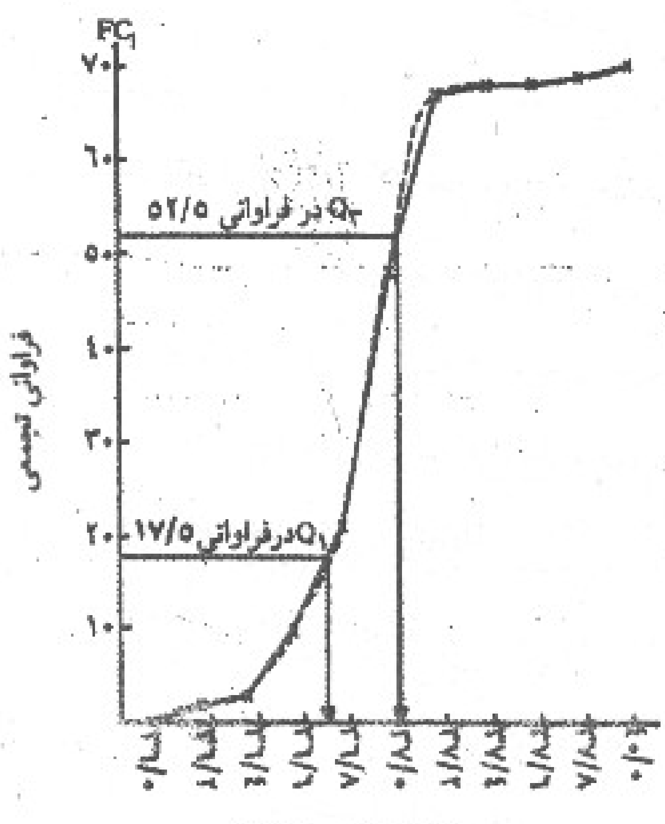
در روش دوم محور افقی عبارت است از کرانه‌های بالای هر طبقه. منحنی این روش را منحنی فراوانی تجمعی خوانند.

بدین ترتیب نقاط متناظر فراوانی تجمعی و محور افقی در صفحه دستگاه مشخص شده، سپس با اتصال آن‌ها به یکدیگر نمودار تجمعی حاصل می‌گردد.

بحث‌های قابل توجهی پیرامون قابلیت نسبی هر یک از روش‌های رسم نمودار فراوانی تجمعی وجود دارد. در پلی‌گن فراوانی تجمعی ارزش‌های همهٔ مقادیر داخل هر طبقه یکسان فرض شده است؛ بنابراین نقطه میانی هر طبقه با فراوانی تجمعی آن طبقه در نظر گرفته می‌شود. سپس آن‌ها را به هم متصل می‌کنند؛ در حالی که در منحنی فراوانی تجمعی فرض بر عدم یکنواخت بودن ارزش مقادیر هر طبقه است. بر همین اساس سعی می‌شود نمودار به گونه‌ای رسم گردد که این خاصیت بخوبی نشان داده شود. ناچار حد واقعی بالای هر طبقه با فراوانی تجمعی آن طبقه در نظر گرفته می‌شود. علاوه بر نکات فوق از نمودار تجمعی برای پاسخ به سؤالاتی همچون «چه تعداد از بطری‌ها حاوی مایعی کمتر از ۳۶/۹ میلی‌لیتر هستند؟» نیز استفاده می‌شود. برای یافتن پاسخ کافی است که خطی عمودی از ۳۶/۹ به نمودار وصل گردد، سپس با خواندن فراوانی تجمعی آن معلوم می‌گردد که جواب ۳۹ است. در صورتی که محور عمودی نمودار با فراوانی تجمعی نسبی مدرج شده باشد عمل نقطه‌یابی مفیدتر خواهد بود.

با نمودار فراوانی تجمعی به راحتی می‌توان هر یک از چندک‌ها را محاسبه کرد. کافی است محل چندک را بر روی محور عمودی منحنی مشخص کرده و از آن‌جا خطی افقی بر نمودار تجمعی وصل کنیم. سپس از آن نقطه خط را به صورت عمودی بر محور افقی رسم نماییم؛ بدین ترتیب مقدار دقیق چندک مشخص می‌گردد.

شکل روبرو چارک اوّل و سوم را به کمک نمودار تجمعی مشخص می‌سازد. چارک اول جایی است که تجمع داده‌ها به ۲۵ درصد N و چارک سوم جایی



شکل ۲: نحوه پیدا کردن چارک اول و سوم

است که تجمع داده‌ها به ۷۵ درصد N برسد. چنانچه مشخص است، چارک اول و سوم به ترتیب مساوی ۳۶/۶۸ و ۳۷ میلی‌لیتر است.

این نمودار جهت مقایسه توزیع‌های فراوانی تجمعی دو یا چند جامعه که از نظر تعداد با هم مساوی هستند، مفید خواهد بود؛ مثلاً برای مقایسه میزان رشد تورم در دو یا چند کشور می‌توان از این نمودار استفاده کرد. هرچه شباهت نمودارها بیشتر باشد، شباهت رشد تورم در آن کشورها بیشتر است.

تحلیل اکتشافی داده‌ها

بافت‌نگار، نمایش هندسی بسیار مفیدی است. تصمیم‌گیرنده می‌تواند به کمک آن درک صحیحی از مشاهدات داشته باشد و نیز در نمایش توزیع، مرکزیت و پراکندگی داده‌ها به کار می‌رود؛ با این حالت بافت‌نگار امکان شناسایی نقاط انفرادی داده‌ها را فراهم نمی‌کند؛ چرا که در این نمودار مشاهدات قرار گرفته در یک طبقه از هم قابل تشخیص نیستند. نمودارهای جدیدی وجود دارد که نسبت به بافت‌نگار اطلاعات بیشتری را در اختیار تصمیم‌گیرنده می‌گذارد. این نمودارها اغلب در مراحل اولیه تحلیل داده‌ها مفید هستند؛ به همین دلیل از آن‌ها با عبارت روش‌های تحلیل اکتشافی داده‌ها نام برده می‌شود. در این بخش دو نمونه از این روش‌ها را معرفی می‌کنیم.

مثال: کدام یک از نمودارهای زیر از نوع تحلیل اکتشافی است؟

۱) میله‌ای　　۲) اجایو　　۳) بافت‌نگار　　۴) ریشه و برگ

حل:

نمودار ریشه و برگ و نمودار جعبه‌ای از انواع تحلیل اکتشافی داده‌ها هستند.

نمودار شاخه و برگ

فرض کنید مشاهدات با $x_1, x_2, \ldots, x_n$ مشخص شده اندو هر مشاهده (X_i) دست کم دو رقمی است. برای تهیه نمودار شاخه و برگ، ارقام مشاهدات را به دو بخش تقسیم می‌کنیم: شاخه شامل یک یا چند رقم اولیه و برگ که شامل ارقام باقیمانده است.

مثلاً اگر داده‌ها بین ۱۰۰ـ۰ باشند، می‌توان مقداری مثل ۲۵ را به شاخهٔ ۲ و برگ ۵ تقسیم کرد. به طور کلی باید شاخه‌های نسبتاً کمی در مقایسه با تعداد مشاهدات انتخاب شود. معمولاً بهتر است این تعداد بین ۵ تا ۲۰ شاخه باشد. زمانی که مجموعهٔ شاخه‌ها انتخاب شد، این اعداد در ستون چپ صفحه نوشته می‌شوند و در کنار هر شاخه، تمام برگ‌های متناظر با مقادیر داده‌ها به ترتیب مشاهده ثبت می‌شود.

مثال: برای رسم نمودار شاخه و برگ، این داده‌ها را که مربوط به زمان انتظار مردان برای تلفن زدن است در نظر بگیرید (زمان بر حسب دقیقه است):

۱۳	۱۳	۰	۰	۱۵	۵	۰	۰
۵	۱۹	۱۰	۱۹	۰	۲۵	۱۸	۰
۱۴	۷	۱۷	۱۷	۱۷	۱۸	۹	۱۵
۸	۱۸	۲۰	۰	۲۱	۰	۰	۳۵
۲۵	۴	۰	۴	۰	۱۰	۰	۰
۲۵	۱۴	۰	۴	۰	۱۰	۰	۰
			۲۷	۲۳	۴	۰	۰

ابتدا مقادیر، ۰، ۱، ۲ و ۳ را به عنوان شاخه انتخاب کرده، سپس داده‌های مربوط به هر شاخه را در جدول آن می‌آوریم. نمودار به این صورت حاصل می‌شود:

فراوانی	برگ	شاخه
۲۲	۴و۰و۰و۴۰و۰و۰و۸۰و۰و۰و۷و۹و۵و۰و۰و۰و۵و۰و۰	۰
۱۶	۴و۰و۸و۷و۷و۷و۸و۵و۹و۰و۹و۸و۳و۳و۵	۱
۶	۷و۳و۵و۰و۱	۲
۱	۵	۳

شکل ۳ نمودار شاخه و برگ زمان انتظار مردان برای تلفن زدن

فراوانی	برگ	شاخه
۲۲	۹و۸و۷و۵و۵و۴و۴و۰و۰و۰و۰و۰و۰و۰و۰و۰و۰و۰و۰و۰و۰	۰
۱۶	۹و۹و۸و۸و۷و۷و۵۷و۵و۴و۳۴و۳و۰و۰	۱
۶	۷و۵و۵و۳و۱و۰	۲
۱	۵	۳

شکل ۴: نمودار شاخه و برگ مرتب شده زمان انتظار مردان برای تلفن زدن

بررسی اجمالی نمودار شکل ۴ مشخص می‌سازد که اکثر زمان انتظارها بین ۰ و تا ۱۹ دقیقه بوده است. به علاوه تمایل داده‌ها به سمت مقادیر کوچکتر است؛ یعنی این‌که پارامترهای مرکزی بخصوص مد، حول این مقادیر متمرکز شده‌اند. بنابراین نمودار شاخه و برگ همانند بافت‌نگار به ما امکان می‌دهد بعضی از خصوصیات مشاهدات را، که به شکل اولیه قابل شناسایی نیستند، به سرعت تعیین کنیم.

برخلاف بافت‌نگار در این نمودار اعداد اصلی از بین نمی‌روند. به علاوه محاسبه چندک‌ها به راحتی امکان‌پذیر است.

مثال: در این مثال به کمک شکل ۴ نحوه محاسبه چارک اول، میانه و چارک سوم نشان داده می‌شود. از آن‌جا که داده‌ها در نمودار شاخه و برگ به صورت صعودی مرتب شده‌اند، پس به راحتی می‌توان با پیدا کردن محل چارک‌ها، هر یک را تعیین نمود.

مکان چارک اول: $C_{Q_1} = \frac{N}{4} + \frac{1}{2} = \frac{45}{4} + \frac{1}{2} = 11.75$

از آن‌جا که عدد ۱۱/۷۵ در شاخه اول بین یازدهمین و دوازدهمین عدد قرار می‌گیرد، پس مقدار چارک اول صفر خواهد بود.

مکان چارک دوم (میانه): $C_{Md} = \frac{2N}{4} + \frac{1}{2} = \frac{2(45)}{4} + \frac{1}{2} = 23$

پس بیست و سومین عدد که در شاخه دوم قرار گرفته است، میانه مشاهدات خواهد بود؛ یعنی $Md = 10$.

مکان چارک سوم: $C_{Q_3} = \frac{3N}{4} + \frac{1}{2} = \frac{3(45)}{4} + \frac{1}{2} = 34.25$

بنابراین $Q_3 = 18$

برخی مواقع ممکن است برگ یا شاخهٔ بیشتری مورد نیاز باشد. تأمین این نیاز، تعدیل شاخه‌های اصلی است؛ برای مثال تقسیم شاخهٔ 1 به دو شاخهٔ جدید 1° و $1^{\circ\circ}$ که شاخه 1° دارای برگ‌های ۰ و ۰ و ۳ و ۳ و ۴ و ۴ و ۵ و ۵ و شاخهٔ $1^{\circ\circ}$ دارای برگ‌های ۷ و ۷ و ۷ و ۸ و ۸ و ۸ و ۹ و ۹ است. بدین ترتیب تعداد شاخه‌های اولیه دو برابر می‌شود. تعیین تعداد شاخه‌ها به تمایل تصمیم‌گیرنده بستگی دارد که ممکن است بیشتر از دو برابر نیز بشود.

نمودار جعبه‌ای

نمودار جعبه‌ای یکی از مفیدترین نمودارهای اکتشافی برای مقایسهٔ دو یا چند جامعهٔ آماری است. نمودار جعبه‌ای نشان‌دهندهٔ چارک‌ها و حداقل و حداکثر مشاهدات است؛ بدین ترتیب که جعبه شامل اختلاف چارک اول و سوم است. در این نمودار، ابتدای جعبه چارک اول و انتهای آن چارک سوم است (از پایین به بالا). خطی که جعبه را به دو نیم تقسیم می‌کند میانهٔ مشاهدات است. از هر طرفِ جعبه نیز به اندازهٔ نقاط مرزی مشاهدات (حداقل و حداکثر) خطی ادامه می‌یابد که گاهی ریشه نامیده می‌شود. در مورد جوامع آماری بزرگ این خطوط ممکن است تا صدک دهم و صدک نودم یا صدک پنجم و نود و پنجم کشیده شود. مراحل تهیه، نمودار جعبه‌ای به این شرح است:

الف) حداکثر داده‌ها را پیدا کنید ب) حداقل داده‌ها را پیدا کنید ج) میانه را پیدا کنید د) چارک اول را پیدا کنید هـ) چارک سوم را پیدا کنید و) نمودار را رسم کنید.

به منظور تشریح بیشتر کاربرد نمودار جعبه‌ای به ذکر یک مثال می‌پردازیم.

مثال: این جدول نشان‌دهندهٔ تعداد لغات انتخاب شده برای عنوان‌های درشت دو روزنامهٔ کیهان و اطلاعات طی روزهای مختلف است. نمودار جعبه‌ای هر دو را تهیه کرده، مقایسه می‌کنیم.

فراوانی لغات در روزنامه اطلاعات (روز)	فراوانی لغات در روزنامه کیهان (روز)	تعداد لغات عنوان‌های درشت
۱۶۶	۱۰	۳
۲۲۶	۴۶	۴
۲۲۵	۸۶	۵
۱۷۸	۹۴	۶
۷۵	۷۸	۷
۲۱	۵۰	۸
۲۰	۲۰	۹
۲	۱۲	۱۰
۱	۱۰	۱۱
۰	۱۲	۱۲
۳	۱۰	۱۳
$\sum F_i = ۹۱۷$	$\sum F_i = ۴۲۸$	

واضح است که چارک اول، دوم و سوم برای روزنامهٔ کیهان به ترتیب ۵ ، ۶ و ۸ و برای روزنامهٔ اطلاعات ۴ ، ۵ و ۶ است. مقادیر حداقل و حداکثر نیز برای هر دوروزنامه مساوی ۳ و ۱۳ است.

مشخص است که این دو روزنامه فقط در میانه، ارزش نزدیک به هم دارند. لغات استفاده شده در روزنامهٔ کیهان بیشتر از روزنامهٔ اطلاعات است؛ به طوری که در روزنامهٔ اطلاعات بیش از ۲۵ درصد عنوان‌ها از ۳ الی ۴ لغت تشکیل شده‌اند.

نمودارهای وصفی

گفته شد که این دسته از نمودارها برای نمایش هندسی داده‌های کیفی به کار می‌روند. مشاهداتی از نوع خوب، بد، متوسط و گروه A ، B ، C را می‌توان از نوع مشاهدات کیفی دانست. طبیعی است برای این نوع مشاهدات، تعریف فاصله چنانچه در بافت نگار گفته شد، مفهوم ندارد. به عبارت دیگر، برای ساختن توزیع فراوانی به جای آن‌که طبقه را به صورت فاصله در نظر بگیریم، ناچاریم هر مقدار را یک طبقه به شمار آوریم. مهمترین نمودارها برای نمایش داده‌های کیفی به شرح ذیل است.

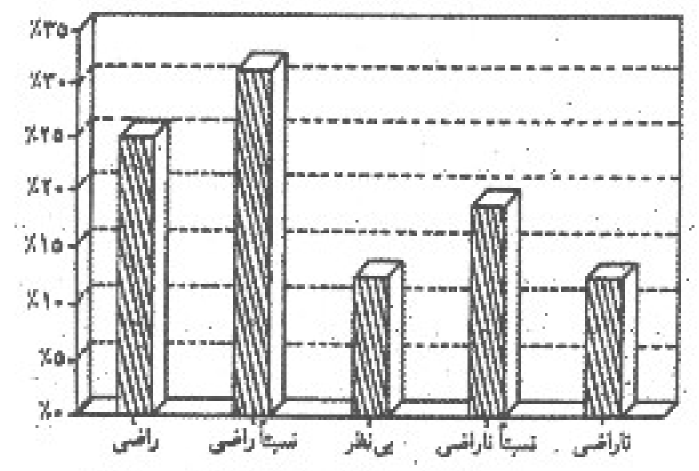
نمودار ستونی

این نمودار را در یک دستگاه مختصات که محور افقی آن نشان‌دهنده کیفیت مشاهدات و محور عمودی آن نشان‌دهنده فراوانی مطلق یا نسبی هر گروه است، رسم می‌کنند. مقادیر متمایز را به صورت نقاطی روی محور افقی مشخص کرده، سپس از نقاط حاصل، خط‌هایی ضخیم بر محور عمود می‌کنیم. ارتفاع هر یک از خطوط بر اساس فراوانی مشخص می‌گردد. در نمودار میله‌ای، خطوط جایگزین مستطیل‌ها می‌شوند تا بر این موضوع تأکید شود که فراوانی‌ها واقعاً روی فاصله‌ها پخش نشده‌اند.

مثال: به منظور اندازه‌گیری رضایت شغلی در سازمانی تحقیقی صورت گرفته و برای این کار از پرسشنامه استفاده شده است. سؤالات پرسشنامه به صورت بسته و با طیف لیکرت (راضی، نسبتاً راضی، بی نظر، نسبتاً ناراضی، ناراضی) تهیه شده‌اند. خلاصهٔ اطلاعات جمع‌آوری شده در این جدول آمده است:

نوع پاسخ (X_i)	تعداد نفرات (F_i)	فراوانی نسبی (f_i)
راضی	۲۰۰	۰/۲۵
نسبتاً راضی	۲۵۰	۰/۳۱۲۵
بی‌نظر	۱۰۰	۰/۱۲۵
نسبتاً ناراضی	۱۵۰	۰/۱۸۷۵
ناراضی	۱۰۰	۰/۱۲۵
	N = ۸۰۰	۱

نمودار ستونی مثال در شکل زیر نشان داده شده است.



نمودار ستونی رضایت شغلی سازمان

نمودار دایره‌ای

نمودار دایره‌ای یکی دیگر از نمودارهایی است که برای نمایش داده‌های کیفی مورد استفاده فراوان قرار می‌گیرد. این نمودار ابزار مناسبی برای تجسّم مشاهدات است و معمولاً برحسب درصد تهیه می‌گردد. برای رسم نمودار دایره‌ای باید این مراحل را در نظر بگیرید:

الف) فراوانی مطلق را به فراوانی نسبی تبدیل کنید.

ب) با استفاده از رابطهٔ $S_i = 360° \times f_i$ مساحت هر قطاع از دایره را پیدا کنید.

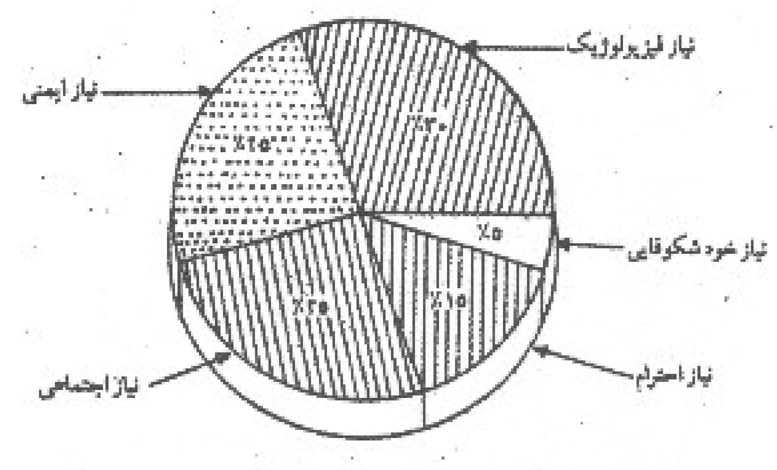
ج) بر اساس S_i مساحت دایره را تقسیم کنید.

د) نوع مشاهدات و درصد آن‌ها را نسبت به کل مشاهدات بر روی دایره بنویسید.

مثال: میزان بلوغ کارکنان یک شرکت براساس سطح نیاز آن‌ها در سلسله مراتب مازلو (Abraham Maslow) اندازه‌گیری شده است. این جدول نشان‌دهندهٔ سطح نیاز کارکنان شرکت است که برای تهیه نمودار دایره‌ای ستون‌های ۳ و ۴ نیز به آن اضافه شده است.

سطح نیاز	تعداد کارکنان (F_i)	فراوانی نسبی (f_i)	$S_i = ۳۶۰° \times f_i$
فیزیولوژیک	۳۰۰	%۳۰	۱۸۰°
ایمنی	۲۵۰	%۲۵	۹۰°
اجتماعی	۲۵۰	%۲۵	۹۰°
احترام	۱۵۰	%۱۵	۵۴°
خودشکوفایی	۵۰	%۰۵	۱۸°
	$\sum F_i = ۱۰۰۰$		$\sum S_i = ۳۶۰°$

چنانچه مشخص است نیاز فیزیولوژیک بیشترین سطح را به خود اختصاص داده و نیاز خود شکوفایی در کمترین سطح قرار دارد؛ به عبارت ساده‌تر در شرکت مورد بحث ۳۰ درصد کارکنان در سطح نیاز فیزیولوژیک، ۳۵ درصد در سطح نیاز ایمنی، ۲۵ درصد در سطح نیاز اجتماعی و ... قرار دارند. این امر نشان می‌دهد که نظریه مازلو درباره سلسله مراتب نیازها در این شرکت صدق می‌کند.



نمودار دایره‌ای سطح نیازهای کارکنان شرکت

مثال: کدام نمودار برای نمایش داده‌ها با مقیاس اسمی به کار می‌روند؟

۱) نمودار مستطیلی ۲) نمودار چند ضلعی یا چند گوش

۳) نمودار دایره‌ای ۴) نمودار تراکمی

حل:

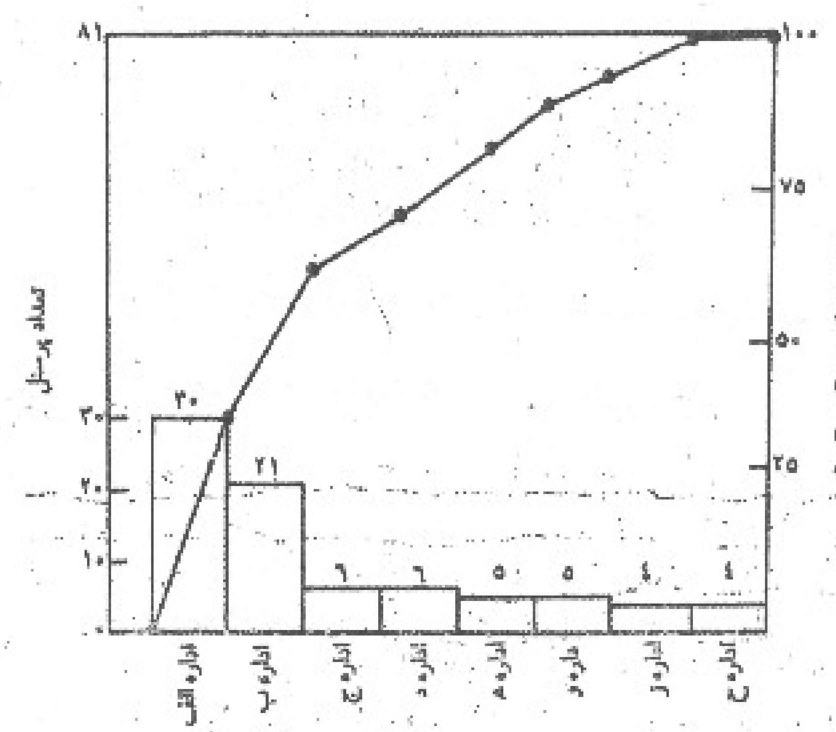
نمودار دایره‌ای برای نمایش داده‌ها با مقیاس اسمی مناسب است.

نمودار پاره‌تو

نموداره پاره‌تو نوعی نمودار ستونی برای داده‌های وصفی است. در این نمودار، فراوانی هر موضوع روی محور عمودی و نوع آن روی محور افقی آورده می‌شود. نمودار پاره‌تو همیشه به ترتیب نزولی فراوانی‌ها ترسیم می‌شود؛ یعنی پروقوع‌ترین موضوع در سمت چپ قرار گرفته و به همین ترتیب فراوانی‌های کمتر در کنار آن جای می‌گیرند.

نام نمودار پاره‌تو از نام یک اقتصاددان ایتالیایی گرفته شده است. طبق نظریهٔ او در بعضی از جوامع بیشتر ثروت در دست عده کمی از مردم است. در داده‌های وصفی معمولاً اصل پاره‌تو به وقوع می‌پیوندد؛ به همین دلیل این نام برای نمودار انتخاب شده است.

این نمودار سه محور دارد که محور سوم آن نشان‌دهنده فراوانی‌های نسبی تجمعی است و به صورت یک محور عمودی دیگر در انتهای محور افقی آورده می‌شود. منحنی روی نمودار پاره‌تو، درصدهای تجمعی K اُمین طبقه را به هم وصل می‌کند. برای رسم نمودار، ستون‌ها به شکل مستطیل در نظر گرفته می‌شود.



نمودار پاره‌تو در تحلیل موجود انبار کالاها، نواقص سیستم‌ها، توزیع درآمد و توزیع کارمندان سازمان‌ها کاربرد فراوانی دارد. در شکل روبرو نمودار پاره‌تو نشان‌دهندهٔ فراوانی کارمندان ادارات مختلف یک سازمان است. این سازمان دارای ۸۱ کارمند و ۸ اداره است که چگونگی توزیع آن‌ها در ادارات بر روی مستطیل‌ها نشان داده شده است. اهمیت اداراتی که بیشترین سهم را در توزیع کارمندان دارند، بر روی نمودار مشهود است.

نمودار پاره‌تو یک قسمت مهم از برنامهٔ کنترل کیفیت محسوب

می‌شود؛ چرا که مدیران و مهندسان به کمک آن می‌توانند توجه خود رابه بحرانی‌ترین نواقص محصول یا فرایند معطوف کنند. زمانی که این نواقص بحرانی شناسایی شد، باید اقدامات اصلاحی برای کاهش یا حذف آن‌ها صورت گیرد.

مجموعه سوالات نمونه کنکور

۱ـ کدام‌یک از این نمودارها برای تحلیل اکتشافی داده‌ها استفاده می‌شود؟

۱) بافت‌نگار ۲) دایره‌ای ۳) پاره‌تو ۴) جعبه‌ای و شاخه و برگ

۲ـ کدام یک از نمودارها برای تحلیل مشاهدات کمی استفاده می‌شود؟

۱) شاخه و برگ ۲) دایره‌ای ۳) پاره‌تو ۴) ستونی

۳ـ کدام‌یک از نمودارها برای نمایش مشاهدات با مقیاس رتبه‌ای مناسب است؟

۱) دایره‌ای ۲) چند ضلعی ۳) بافت‌نگار ۴) جعبه‌ای

۴ـ کدام‌یک از نمودارها برای نمایش مشاهدات با مقیاس نسبی مناسب است؟

۱) بافت‌نگار ۲) چند ضعلی ۳) جعبه‌ای ۴) هر سه

۵ـ در رسم نمودارهای بافت‌نگار، محور x را بر اساس کدام اندازه مدرج می‌کنند؟

۱) فراوانی‌های نسبی ۲) کرانه‌های طبقات ۳) متوسط طبقات ۴) فراوانی‌های تجمعی

۶ـ در رسم نمودار تجمعی، محور x را بر اساس کدام اندازه مدرج می‌کنند؟

۱) متوسط طبقات ۲) کرانه‌های طبقات ۳) حد پائین طبقات ۴) گزینه‌های ۱ و ۲

۷ـ در کدام‌یک از این نمودارها ارزش مشاهدات هر طبقه یکسان تلقی می‌شود؟

۱) منحنی فراوانی تجمعی ۲) بافت‌نگار ۳) پلی‌گن فراوانی تجمعی ۴) هر سه

۸ـ اگر حداکثر و حداقل مشاهدات به ترتیب ۴۰۰ و ۲۰۰ و فاصله طبقات ۲۵ باشد، تعداد طبقات جدول طبقه‌بندی داده‌ها کدام است؟

۱) ۱۶ ۲) ۸ ۳) ۵ ۴) ۲۵

۹ـ اگر ۸۹ـ ۸۰ و ۹۹ـ ۹۰ دو طبقه از یک جدول طبقه‌بندی شده باشند فاصله طبقات کدام است؟

۱) ۹ ۲) ۱۰ ۳) ۹/۵ ۴) هیچکدام

۱۰ـ برای مقایسه دو توزیع فراوانی مربوط به حقوق پرداختی به کارگران مرد و زن در یک کارخانه، کدام یک از نمودارهای زیر مناسب‌تر است؟

۱) پلی‌گون (چند گوش) فراوانی نسبی ۲) نمودار فراوانی میله‌ای مطلق

۳) نمودار تجمعی فراوانی مطلق ۴) هیستوگرام (بافت نگار) فراوانی نسبی

۱۱ـ بهترین نمایش تصویری برای مقایسه دو مجموعهٔ دادهٔ اسمی، کدام است؟

۱) جعبه‌ای ۲) بافت‌نگار ۳) نمودار میله‌ای ۴) چند ضلعی

۱۲ـ کدام یک از نمودارهای آماری زیر برای توصیف داده‌ها با مقیاس اسمی مناسب‌تر است؟

۱) جعبه‌ای ۲) ریشه و برگ ۳) پاره‌تو ۴) چند ضلعی

۱۳ـ کدام نمودار برای نمایش مشاهدات کمی طبقه‌بندی نشده به کار می‌رود؟

۱) پاره‌تو ۲) چند ضلعی ۳) ریشه و برگ ۴) بافت نگار

۱۴ـ برای رسم هیستوگرام (نمودار مستطیلی) محورهای x و y بر اساس کدام اندازه‌ها مدرج می‌شوند؟

۱) کرانه‌های طبقات و فراوانی طبقات ۲) حدود طبقات و چگالی

۳) حد وسط طبقات و فراوانی مطلق ۴) مقادیر متغیر x و فراوانی‌های تجمعی

پاسخنامه:

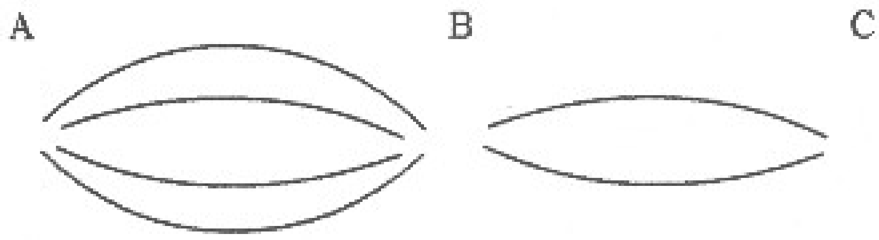
	۱	۲	۳	۴
۱				■
۲	■			
۳	■			
۴				■
۵		■		
۶				■
۷			■	
۸		■		
۹		■		
۱۰	■			
۱۱			■	
۱۲			■	
۱۳			■	
۱۴	■			

آنالیز ترکیبی

اصل ضرب (اصل شمارش)

فرض کنید عملی در k مرحله انجام گیرد و مرحله i ام به n_i طریق ممکن باشد در این صورت عمل مذکور می‌تواند به $n_1 . n_2 ... n_k$ طریقه انجام یابد.

بعنوان مثال ساده فرض کنید شخصی می‌خواهد از شهر A به شهر C مسافرت کند و بایستی حتماً از شهر B عبور کند. یعنی عمل مسافرت این شخص در دو مرحله انجام می‌گیرد. مرحله اول مسافرت از شهر A به شهر B که مثلاً به ٤ طریق ممکن است و مرحله دوم مسافرت از شهر B به شهر C که مثلاً به دو طریق ممکن است. پس این عمل بنا به اصل ضرب به $4 \times 2 = 8$ طریقه ممکن می‌باشد.



مثال:

(a) چند عدد 4 رقمی داریم.

با توجه به آنکه رقم صفر نمی‌تواند در هزارگان باشد.

یکان دهگان صدگان هزارگان

9 10 10 10 → 9000

(b) چند عدد 4 رقمی با ارقام زوج داریم :

با توجه به انکه انتخاب‌ها فقط از میان ارقام زوج 8 , 6 , 4 , 2 , 0 می‌باشد پس

4 5 5 5 → 500

(c) چند عدد 4 رقمی با ارقام یکان و هزارگان یکسان داریم.

رقم یکان پس از انتخاب هزارگان حق انتخاب ندارد.

9 10 10 1 → 900

(d) چند عدد 4 رقمی داریم که فقط ارقام یکان و هزارگان یکسان دارند.

در واقع یکان و هزارگان یکسانند و دهگان و صدگان نیز نه با آنها یکسانند و نه با خودشان.

9 9 8 1 → 648

(e) چند عدد 4 رقمی بدون تکرار ارقام داریم :

9 9 8 7 → 4536

(f) چند عدد 4 رقمی زوج بدون تکرار ارقام داریم :

حل این مساله بصورت انجام یک عمل ضرب امکان‌پذیر نیست لذا بایستی این مساله را به دو نوع مجزا تقسیم نموده و هر نوع را با اصل ضرب حل کرد، سپس دو مقدار بدست آمده را جمع کنیم.

9 8 7 1 → 504

اعدادی که با رقم صفر زوج می‌شوند

8 8 7 4 → 1792

اعدادی که با رقم غیرصفر زوج می‌شوند

دقت شود در اعدادی که با رقم غیرصفر زوج می‌شوند 4 انتخاب برای یکان داریم و 8 انتخاب برای هزارگان چون صفر و رقمی که در یکان بکار برده شده حق انتخاب ندارند بعد از ان که دو رقم استفاده شده است به ترتیب 8 , 7 انتخاب برای صدگان و دهگان داریم.

پس در نتیجه تعداد مطلوب برابر است با :

$504+1792=2299$

(g) چند عدد 4 رقمی داریم که فقط سه رقم یکان و دهگان و صدگان یکسان دارند.

9 9 1 1 → 81

جایگشت: هر ترتیبی را که میتوان اشیاء یک مجموعه را در کناریکدیگر قرار داد یک جایگشت می گویند .

قضیه ۱: تعداد ترتیب یا جایگشت n شیء متمایز در یک صف کنار یکدیگر برابراست با :

$n\times(n-1)\times........\times2\times1=n!$

مثال: 6 نفر و 6 صندلی در یک ردیف داریم. اگر یک نفر از بین این 6 نفر (نفر مشخصی) روی یکی از صندلی‌ها بنشیند، بقیه به چند حالت می‌توانند بر روی صندلی‌های باقی مانده بنشینند؟ !6

ترتیب مدور : نحوه قرار گرفتن افراد یا اشیاء را روی محیط یک منحنی بسته ، ترتیب مدور می گویند.

نکته ۱: هر ترتیب مدورn شیء متمایز برابراست با $(n-1)!$ که چرخش هر ترتیب حول یک محور میانی ترتیب جدیدی محسوب نخواهد شد.

مثال: برای 5 زن و 5 مرد مطلوبست محاسبه تعداد حالات نشستن آن‌ها روی 10 صندلی در یک ردیف و حول یک میزگرد با در نظر گرفتن شرایط زیر:

الف) کل حالات نشستن

!9 → میزگرد

!10 → ردیف

ب) مردها کنار هم بنشینند.

تمامی مردها را یک نفر در نظر گرفته در نتیجه خواهیم داشت: 1 نفر + 5 زن. حال تعداد حالات نشستن این 6 نفر در یک ردیف را حساب می کنیم: !6 پس جواب کلی ما !5 !6 خواهد بود.

میزگرد: !5 تعداد حالات نشسن مردها کنار هم در یک ردیف است. باز همانند فوق 6 نفر خواهیم داشت که باید بر سر یک میزگرد بنشینند. بنابراین تعداد کل حالت !5!5 خواهد بود.

ج) مردها کنار هم و زن‌ها نیز کنار هم بنشینند.

ردیف: !5 تعداد حالات نشستن مردها در کنار هم در یک ردیف و نیز همچنین !5 تعداد حالات نشستن زن‌ها در کنار هم در یک ردیف است. حال دو مجموعه را دو نفر در نظر می گیریم که قرار است در یک ردیف در کنار هم بنشینند، بنابراین جواب !5!5!2 خواهد بود.

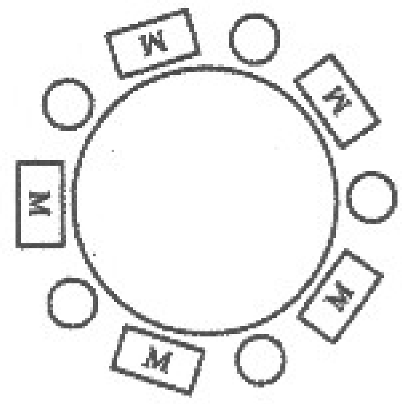
میزگرد: !5 تعداد حالات نشستن مردها در کنار هم در یک ردیف و نیز همچنین !5 تعداد حالات نشستن زن‌ها در کنار هم در یک ردیف است. حال دو مجموعه را دو نفر در نظر می گیریم که قرار است یک میزگرد بنشینند. بنابراین جواب !5!5!1 خواهد بود.

* د) مردها و زن‌ها متناوباً کنار هم قرار بگیرند (حداکثر اختلاف بین تعداد زن‌ها و مردها در شرایطی که می‌خواهیم یکی در میان در یک ردیف بنشینند، 1 باید باشد)

	M		M		M		M		M	

تعداد حالات ممکن نشستن زن‌ها (مردها) !5 خواهد بود. بنابراین جواب !5!5!2 است.

میزگرد: تعداد حالات ممکن نشستن مردها و زن‌ها، حول میزگرد !4 است. حال 5 جا برای نشستن 5 زن (مرد) وجود دارد، یعنی !5 حالت. بنابراین جواب !5 !4 خواهد بود.



نکته: در صورتی که تعداد مردها 5 و تعداد زن‌ها 4 باشد نمی‌توان آن‌ها را یکی در میان دور میز قرار داد. اما می‌توان در یک ردیف به صورت !5!4 قرار داد.

نکته: در صورتی که تعداد مردها 5 و تعداد زن‌ها 3 باشد به هیچ شکل نمی‌توان آن‌ها را یکی در میان قرار داد.(دور میز یا در یک دریف)

نکته ۲: در جا سوئیچیها و تسبیح‌های n تایی،چون میتوان آنها را برگرداند. تعداد جایگشتها برابراست با : $\frac{(n-1)!}{2}$

نکته ۳: تعداد ترتیب یا جایگشت n شی که n_1 تای آن از نوع اول و n_2 تای آن از نوع دوم و .... n_k تای آن از نوع k ام باشد برابراست با :

$$\frac{n!}{n_1!, n_2! .....n_k!} \quad , \quad n_1+n_2+.....+n_k=n$$

نکته ۴: تعداد تقسیمات n شی ء در k سلول به طوریکه n_1 تای آنها در سلول اول و n_2 تای آنها در سلول دوم و ...... n_k تای آنها در سلول k ام قرار گیرد . برابراست با :

$$\binom{n}{n_1\ n_2 ....n_k}=\frac{n!}{n_1!n_2!n_k!}$$

مثال: به چند طریق می‌توان 9 نفر کارمند را در یک اتاق 4 نفره یک اتاق 3 نفره و یک اتاق 2 نفره چیدمان کرد؟

۱) 1400 ۲) 1260 ۳) 72 ٤) 24

حل:

از فرمول جایگشت با تکرار استفاده می کنیم.

$$P_9=\frac{9!}{4!3!2!}=1260$$

گزینه ۲ صحیح می‌باشد.

مثال: به چند طریق می‌توان 9 اسباب‌بازی را بین 4 بچه تقسیم کرد به شرط آن‌که به کوچک‌ترین بچه 3 اسباب‌بازی و به هر کدام از بچه‌های دیگر 2 اسباب‌بازی برسد؟

۱) 27 ۲) 108 ۳) 5674 ٤) 7560

گزینه ٤ صحیح می‌باشد.

$$\frac{9!}{3!2!2!2!}=7560$$

مثال: مطلوبست تعداد کلمات یازده حرفی متفاوتی که می‌توان با حروف کلمه statistical ساخت

$$\frac{11!}{2!3!2!2!}$$

ترکیب: انتخاب r شی از n شی بدون ترتیب و بدون جایگذاری برابر است با :

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

که آن را ترکیب r شی از n شی میگویند.

☑ در صورتیکه در ترکیب با جایگذاری انتخاب انجام شود تعداد حالات n^r می‌شود.

مثال: تعداد نمونه‌های سه‌تایی با جای‌گذاری و بدون جای‌گذاری از جامعه‌ای که دارای 5 عنصر است به ترتیب کدام است؟

۱) 125 و 5 ۲) 125 و 10 ۳) 243 و 10 ٤) 243 و 20

حل:

با جایگذاری : $C_r^n=C_3^3=\frac{5!}{3!2!}=10$

بدون جایگذاری : $n^r=5^3=125$

گزینه ۲ صحیح می‌باشد

☑ **تبدیل:** اگر در ترکیب هنگام انتخاب r شی از n شی ترتیب اهمیت داشته باشد، در این حالت تعداد انتخابها به صورت زیر نوشته می‌شود:

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

مثال: اگر 9 نفر در یک مسابقه شرکت کنند به چند طریق ممکن است جوایز اول و دوم و سوم را دریافت کنند؟

۱) 84 ۲) 504 ۳) 635 ٤) 3024

حل:

چون ترتیب جوایز مهم است، بنابراین:

$$P_9^3=\frac{9!}{6!}=9\times8\times7=504$$

گزینه ۲ صحیح می‌باشد.

مثال : انتخاب ۳ دانشجو از بین ۱۰ دانشجو :

الف) به عنوان شاگردان ممتاز (ترتیب اهمیتی ندارد)

$$\binom{10}{3} = \frac{10!}{3!\,7!} = 120$$

ب) به عنوان شاگرد اول،دوم ، سوم (ترتیب اهمیتی دارد)

$$P_{10}^{3} = \frac{10\ !}{(10-3)!} = 720$$

نکتهٔ ۶: حالتهای خاص در ترکیبات:

$$a)\binom{n}{0} = \binom{n}{n} = 1 \qquad b)\binom{n}{1} = \binom{n}{n-1} = n$$

$$c)\binom{n}{2} = \frac{n(n-1)}{2} \qquad d)\binom{n}{r} = \binom{n}{n-r}$$ رابطه مهم

$$e)\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$$

مثال: عبارت روبرو را ساده کنید:

$$C_r^{n+1} + 2C_{r-1}^{n+1} + C_{r-2}^{n+1}$$
$$C_r^{n+1} + C_{r-1}^{n+1} + C_{r-1}^{n+1} + C_{r-2}^{n-1} = C_r^{n+2} + C_{r-1}^{n+2} = C_r^{n+3}$$

قضیه ۲: تعداد تقسیمات n شی مشابه در k سلول برابر است با:

$$\binom{n+k-1}{k-1} \quad \text{یا} \quad \binom{n+k-1}{n}$$

نتیجه۱: تعداد تقسیمات n شی مشابه در k سلول به طوریکه در هر سلول حداقل r شی قرار گیرد .

$$\binom{n-k(r-1)-1}{k-1}$$

نتیجه۲ : در نتیجه ۱ اگر r=1 در نظر گرفته شود. یعنی تعداد تقسیمات n شی مشابه در k سلول به طوریکه در هر سلول حداقل ۱ شی قرار گیرد، برابر است با :

$$\binom{n-1}{k-1}$$

نتیجه ۳: در بسط چند ، چند جمله‌ای $(x_1 + \dots\dots + x_k)^n$ که به صورت :

$$(x_1 + \dots\dots\dots + x_k)^n = \sum \binom{n}{n_1 \dots\dots n_k} x_1^{n_1}\ x_2^{n_2} \dots\dots x_k^{n_k}$$

است تعداد جملات برابر است با $\binom{n+k-1}{k-1}$ یا $\binom{n+k-1}{n}$

ضریب هر جمله مانند $x_1^{n_1}\ x_2^{n_2} \dots\dots\dots x_k^{n_k}$ برابراست با : $\dfrac{n!}{n_1!\,n_2!\dots\dots n_k!}$

قضیه ۳: تعداد جوابها یا بردارها با مقادیر صحیح و غیر منفی $x_1, \ldots\ldots, x_k$ به صورت $[x_1, x_2, \ldots\ldots, x_k]$ و $x_1 + \ldots\ldots + x_k = n$ برابر است با :

$$\binom{n+k-1}{n} \text{ یا } \binom{n+k-1}{k-1}$$

نتیجه ۱: تعداد جوابها یا برادرهای با مقادیر صحیح و غیر منفی و غیر صفر معادله قضیه ۳ برابر است با :

$$\binom{n-1}{k-1}$$

نکته ۱:

$$(x+y)^n = \binom{n}{n} x^n y^0 + \binom{n}{n-1} x^{n-1} y + \ldots\ldots + \binom{n}{0} x^0 y^n$$

$$\text{a)} \begin{cases} x=1 \\ y=1 \end{cases} \Rightarrow 2^n = \binom{n}{0} + \binom{n}{1} + \ldots\ldots + \binom{n}{n}$$

تعداد زیرمجموعه‌های یک مجموعه n عضوی $= 2^n$

$$\text{b)} \begin{cases} x=1 \\ y=-1 \end{cases} \Rightarrow 0 = \binom{n}{n} - \binom{n}{n-1} + \ldots\ldots$$

$$\Rightarrow \binom{n}{0} + \binom{n}{2} + \ldots\ldots = \binom{n}{1} + \binom{n}{3} + \ldots\ldots$$

نکته ۲: با n شی می‌خواهیم ترکیب‌هایی حداکثر rتایی و حداقل 1تایی که در آن‌ها تکرار مجاز است بسازیم، تعداد حالت ممکن عبارت است از:

$$n + n^2 + n^3 + \ldots\ldots n^r = \frac{n^r - 1}{n-1} \times n$$

مثال: با عددهای ٤ و ٣ و ٢ و ١ و ٦ می‌خواهیم اعدادی بسازیم که حداکثر ١٠ رقمی باشد.

$$r=10 \ , n=5 \Rightarrow 5^1 + 5^2 + \ldots\ldots + 5^{10} = \frac{5^{10}-1}{5-1} \times 5$$

ترتیب‌های ناسازگار :

اگر سه حرف abc مفروض باشند، به هر ترتیب قرار گرفتن سه حرف مذکور به طوری که هیچ یک از آنها در جایگاه فعلیشان قرار نگیرند، ترتیب ناسازگار گفته میشود:

قضیه ٤: تعداد ترتیب‌های ناسازگار n شی متمایز برابر است با : $n!\sum_{k=0}^{n} \frac{(-1)^k}{k!}$

نتیجه ۱: تعداد ترتیب‌های ناسازگار r شی از n شیء متمایز برابراست با :

$$\frac{n!}{(n-r)!} \sum_{k=0}^{r} \frac{(-1)^k}{k!}$$

r ناسازگاری معادل n-r سازگاری است یعنی :

$$\binom{n}{r} \times r! \sum_{k=0}^{r} \frac{(-1)^k}{k!} = \frac{n!}{r!(n-r)!} \times r! \sum_{k=0}^{r} \frac{(-1)^k}{k!} = \frac{n!}{(n-r)!} \sum_{k=0}^{r} \frac{(-1)^k}{k!}$$

مثال: برای 10 نفر 10 نامه می‌فرستیم. مطلوب است تعداد حالاتی که فقط هفت نفر نامه خودشان را دریافت نمایند.

$$\binom{10}{3} \times 3! \sum_{k=0}^{3} \frac{(-1)^k}{k!}$$

احتمال:

تعریف: اندازه امکان وقوع حادثه A را با P(A) نشان داده که احتمال حادثه بوده و

$$0 \le P(A) \le 1$$

نکته ۱: حوادث از نقطه نظر احتمال وقوع عبارتند از:

۱) غیر ممکن $\leftarrow P(A)=0$

۲) تصادفی $\leftarrow 0<P(A)<1$

۳) یقینی (حتمی) $\leftarrow P(A)=1$

نکته ۲: حوادث با هم به صورت‌های زیر در نظر گرفته می‌شوند:

الف) حوادث هم تراز:

به حوادثی که احتمال وقوع یکسان و برابر با هم را داشته باشند حوادث هم تراز می گوئیم.

مثال: نتایج پرتاپ یک تاس $\begin{cases} P(A_i)=\frac{1}{6} \\ i=1,2,\dots,6 \end{cases}$

مثال: نتایج پرتاب یک سکه خط و شیر $\begin{cases} P(A_i)=\frac{1}{2} \\ i= \end{cases}$

حوادث به طور پیش فرض هم تراز فرض می‌شوند و n حادثه هم تراز به طور پیش فرض هر کدام احتمال $\frac{1}{n}$ دارند

ب) حوادث مستقل:

هر گاه وقوع یا عدم وقوع یک حادثه تاثیری در وقوع یا عدم وقوع حادثه دیگر نداشته باشد. دو حادثه را مستقل گویند. و خواهیم داشت:

$$A,B \text{ مستقل} \Leftrightarrow P(A\cap B)=P(A)\times P(B)$$

مثال: اگر P(A)=0.3 و P(B)=0.2 و $P(A\cap B)=0.06$، رویدادهای (حوادث) A و B چگونه‌اند.

۱) مکمل ۲) مستقل ۳) ناسازگار ٤) وابسته

حل:

$$P(A\cap B)=P(A).P(B) \Rightarrow 0.06=0.3\times 0.2 \rightarrow \text{A و B مستقل‌اند}$$

گزینه ۲ صحیح می‌باشد.

مثال: در پرتاب یک تاس و یک سکه ظاهر شدن 2 و شیر چه خواهد بود.

2 آمدن تاس = A

شیر آمدن سکه = B

$$B,A \xRightarrow{\text{مستقل}} P(A\cap B)=P(A)\times P(B)=\frac{1}{6}\times\frac{1}{2}=\frac{1}{2}$$

* در احتمال " و " را با $\cap$ و "یا" را با $\cup$ نمایش می دهند.

مثال: شرط استقلال سه واقعه A , B و C تعریف شده در یک فضای نمونه از یکدیگر چیست؟

۱) $P(A\cap B\cap C)=P(A).P(B).P(C)$

۲) $P(A\cap B)=P(A).P(B)$

$P(A\cap C)=P(A).P(C)$

$P(B\cap C)=P(B).P(C)$

۳) $P(A\cap B\cap C)=P(A).P(B).P(C)$

$P(A\cap B)=P(A).P(B)$

$P(A\cap C)=P(A).P(C)$

$P(B\cap C)=P(B),P(C)$

٤) هیچکدام

حل:

سه پیشامد B,A و C را دو به دو مستقل گویند، هر گاه داشته باشیم:

$$(1)\begin{cases}P(A\cap B)=P(A).P(B)\\P(A\cap C)=P(A).P(C)\\P(B\cap C)=P(B).P(C)\end{cases}$$

در برخی از مسائل اتفاق می‌افتد که:

$$(2)P(A\cap B\cap C)=P(A).P(B).P(C)$$

در این حالت نمی‌توان گفت که هر سه پیشامد دو به دو مستقل‌اند، زیرا نمی‌توان شرط (1) را از شرط (2) نتیجه گرفت.

گزینه ۳ صحیح می‌باشد.

مثال: $P(A)=0.1$ و $P(B)=0.2$ و $P(A\cap B)=0.02$ ، A و B نسبت به هم چگونه‌اند؟

۱) ناسازگار ۲) مستقل ۳) وابسته ۴) مکمل

مثال: اگر در پرتاب یک تاس:

$A=\{1,2\}$ ظاهر شدن 1 یا 2 $\Rightarrow P(A)=\frac{2}{6}=\frac{1}{3}$

$B=\{1,3,5\}$ ظاهر شدن 1 یا 3 یا 5 $\Rightarrow P(B)=\frac{3}{6}=\frac{1}{2} \Rightarrow P(A\cap B)=P(A)\times P(B)$

$A\cap B=\{1\}\Rightarrow P(A\cap B)=\frac{1}{6}$

در نتیجه A و B مستقل هستند.

مثال: اگر در پرتاب یک تاس:

$A\{1,2\}$ ظاهر شدن 1 یا 2

$B=\{1,3,5,7\}$ ظاهر شدن 1 یا 3 یا 5 یا 7 $\Rightarrow P(A\cap B)\neq P(A)\times(B)$

$A\cap B=\{1\}$

در نتیجه A و B مستقل نیستند.

ج) حوادث ناسازگار:

هر گاه وقوع همزمان دو حادثه غیر ممکن باشند، حوادث را ناسازگار گوئیم و در نتیجه:

$$A, B \Leftrightarrow \begin{cases} P(A\cap B)=0 \\ A\cap B=\varnothing \end{cases}$$ ناسازگارند.

مثال۱: $P(A)=0.1$ و $P(B)=0.2$ و $P(A\cap B)=0$ ، A و B نسبت به هم چگونه‌اند؟

۱) مستقل ۲) ناسازگارند ۳) وابسته‌اند ۴) هیچکدام

نکته: در دو حادثه مستقل A و b هر گاه احتمال وقوع یکی از حوادث 0 شود آنگاه دو حادثه ناسازگارند.

$$P(A\cap B)=0 \quad \overset{P(A)=0}{\underset{P(B)=0}{\Rightarrow}} \quad B \Leftrightarrow P(A\cap B)=P(A)\times P(B)$$ و A مستقل A و B ناسازگارند.

مثال۲: در پرتاب یک سکه و یک تاس (مستقل) احتمال ظاهر شدن 7 و شیر کدام است؟

ظاهر شدن B=7 ظاهر شدن شیر = A

$$P(A\cap B)=P(A)\times P(B)=\frac{1}{2}\times 0=0$$

گروه کامل حوادث:

اگر حادثه A فقط و فقط به یکی از نتایج A_1 و A_2 و ..... و A_k ختم شود، آنگاه A_1 و A_2 و ..... و A_k را گروه کامل حوادث گفته و خواهیم داشت: $P(A_1\cup A_2.....\cup A_k)=1$

مثال: در پرتاب یک تاس نتایج 1 و 2 و 3 و 4 و 5 و 6 گروه کامل حوادث هستند

محاسبه احتمال:

۱ـ احتمال کلاسیک : $P(A)=\dfrac{\text{تعداد مساعد}}{\text{تعداد کل حالات}}$

مثال: در پرتاب یک تاس احتمال ظاهر شدن عدد 1 یا 2 ؟ $P(A)=\frac{2}{6}$

مثال: در پرتاب یک تاس احتمال ظاهر شدن عدد زوج؟ $P(A)=\frac{3}{6}$

۲ـ احتمال هندسی: $P(A)=\dfrac{\text{مساحت مساعد}}{\text{مساعد کل}}$

مثال: در صورتیکه یک نقطه داخل یک مربع به ضلع 2 انتخاب شود، احتمال آنکه داخل دایره محاط در آن به شعاع 1 باشد؟

$$P(A)=\frac{\text{مساحت دایره}}{\text{مساحت مربع}}=\frac{\pi\times 1^2}{2^2}=\frac{\pi}{4}$$

۳ـ احتمال آماری:

فضای نمونه‌ای:

مجموعه نتایج ممکن یک آزمایش تصادفی را فضای نمونه‌ای گویند و با S نشان می‌دهند.

مثال:

الف) فضای نمونه پرتاب یک سکه عبارت است از:

S = { خط و شیر } = $\{H, T\}$

ب) فضای نمونه پرتاب یک تاس عبارتست از:

$S = \{1, 2, 3, 4, 5, 7\}$

پیشامد:

هر زیر مجموعه از فضای نمونه را یک پیشامد می‌گویند.

مثال: در پرتاب یک تاس فضای نمونه‌ای 6 عضو دارد. پس $2^6 = 64$ پیشامد قابل تصور است، مانند :

$A_1 = \{1, 2\}$ یا $A_2 = \{2, 5, 6\}$ یا ....

مسائل مهم احتمال:

1) **پرتاب تاس و سکه:** (در پرتاب m تاس فضای نمونه 6^m و در پرتاب n سکه 2^n است.)

الف) در پرتاب دو تاس احتمال ظاهر شدن مجموع 5 چقدر است؟

حالات مساعد: $\{(1,4),(4,1),(2,3),(3,2)\}$

حالات کل: 6^2 $\Rightarrow P(A)=\frac{4}{36}=\frac{1}{9}$

ب) در پرتاب دو تاس احتمال ظاهر شدن مجموع 6 چقدر است؟

حالات مساعد: $\{(1,5),(5,1),(4,2),(2,4),(3,3)\}$

حالات کل: 6^2 $\Rightarrow P(A)=\frac{5}{36}$

ج) در پرتاب دو تاس احتمال ظاهر شدن مجموع 6 و تاس اول کمتر از 3 چقدر است؟

حالات مساعد: $\{(1,5),(2,4)\}$

حالات کل: 6^2 $\Rightarrow P(A)=\frac{2}{36}$

د) در پرتاب دو تاس احتمال آنکه مجموع کمتر از 5 باشد، چقدر است؟

حالات مساعد: $\{(1,1),(1,2),(2,1),(2,2),(1,3)(3,1)\}$

حالات کل: 6^2 $\Rightarrow P(A)=\frac{6}{36}=\frac{1}{6}$

هـ) در پرتاب دو تاس احتمال آنکه مجموع کمتر از 5 باشد، و یکی از تاس‌ها کمتر از 2 باشد چقدر است؟

حالات مساعد: $\{(1,3),(1,2),(1,1),(2,1),(3,1)\}$

حالات کل: 6^2

$$\Rightarrow P(A)=\frac{5}{36}$$

و) در پرتاب 3 تاس احتمال آنکه نتایج متفاوت باشد؟ $\frac{6}{6}\times\frac{5}{6}\times\frac{4}{6}$

د) در پرتاب 3 تاس احتمال آنکه تاس اول و سوم یکسان باشد؟ $\frac{6}{6}\times\frac{6}{6}\times\frac{1}{6}$

ذ) در پرتاب 3 تاس احتمال آنکه تاس اول و سوم یکسان و با دوم متفاوت باشد؟ $\frac{6}{6}\times\frac{5}{6}\times\frac{1}{6}$

ل) در پرتاب 4 تاس احتمال آنکه تاس اول عدد 4 باشد چقدر است؟ $\frac{1}{6}=\frac{1}{6}\times\frac{6}{6}\times\frac{6}{6}\times\frac{6}{6}$

م) در پرتاب 4 تاس احتمال آنکه تاس اول عدد 4 باشد و تاس دوم و چهارم برابر باشد؟ $\frac{1}{36}=\frac{1}{6}\times\frac{6}{6}\times\frac{6}{6}\times\frac{1}{6}$

✱ **نکته:** در پرتاب چند تاس اگر بخواهیم نتایج یکسان در مراحل مختلف داشته باشیم کافی است مرحله اول را محاسبه کرده و به جای مراحل دیگر یکسان 1 می‌گذاریم.

ن) در پرتاب 4 تاس احتمال آنکه تاس اول و سوم یکسان و تاس دوم و چهارم یکسان باشند و با هم متفاوت باشند؟

$\frac{6}{6}$	$\times\frac{5}{6}$	$\times\frac{1}{6}$	$\times\frac{1}{6}$
اول	دوم	سوم	چهارم

پرتاب سکه:

الف) در پرتاب 2 سکه احتمال ظاهر شدن نتایج یکسان چقدر است؟

حالات مساعد: { (خ و خ) و (ش و ش) }

حالات کل: 2^2

$$\Rightarrow P(A)=\frac{2}{4}$$

ب) در پرتاب 3 سکه احتمال ظاهر شدن حداقل یک خط کدام است؟

P (حداقل یک خط) = 1-P (همه شیر) = 1 − $\frac{\text{(ش و ش و ش)}}{2^3}$ $=1-\frac{1}{8}=\frac{7}{8}$

ج) در پرتاب 4 سکه احتمال آنکه سکه اول خط ظاهر شود؟

P (سکه اول خط) = $\frac{1}{2}$ (تاس اول مستقل از تاس‌های دیگر است.)

د) در پرتاب 2 سکه احتمال آنکه نتایج متفاوت باشد، چقدر است؟

P (نتایج متفاوت) = $\frac{\text{(ش و خ) (خ و ش)}}{2^3}$ $=\frac{2}{4}$

پرتاب تاس و سکه: (تاس و سکه مستقل از هم بررسی می‌شوند)

الف) در پرتاب یک تاس و یک سکه احتمال آنکه تاس 5 و سکه خط ظاهر شود؟

$P(\text{سکه خط و تاس } 5) = P(\text{سکه خط}) \times P(\text{تاس } 5) = \frac{1}{6}\times\frac{1}{2}=\frac{1}{12}$

ب) در پرتاب 3 تاس و یک سکه احتمال آنکه سکه خط و تاس اول و سوم یکسان ظاهر شوند؟

$\underbrace{\frac{1}{2}}_{\text{سکه خط}}\times\frac{6}{6}\times\frac{6}{6}\times\frac{1}{6}=\frac{1}{12}$

مسئله مهره‌ها:

۱ـ کیسه‌ای شامل 10 مهره از شماره 1 تا 10 است مهره‌ای انتخاب می‌کنیم.

الف) احتمال آنکه مهره 4 باشد؟ $P(A)=\frac{1}{10}$

ب) احتمال آنکه مهره یک عدد بین 1 تا 10 باشد؟ $P(A)=\frac{1}{10}$

ج) احتمال آنکه مهره زوج باشد؟ $P(A)=\frac{5}{10}$ $(2,4,6,8,10)$ حالات مساعد

د) احتمال آنکه مهره زوج و کمتر از 6 باشد؟ $P(A)=\frac{2}{10}$ $(2,4)$ حالات مساعد

۲ـ کیسه‌ای شامل 4 مهره قرمز است و 5 مهره آبی است، مهره‌ای از آن خارج می‌کنیم احتمال آنکه قرمز باشد؟ $\frac{4}{9}$

۳ـ کیسه‌ای شامل 4 مهره قرمز و 5 مهره آبی است، 2 مهره آبی از آن خارج می‌کنیم، سپس مهره‌ای خارج می‌کنیم احتمال آنکه مهره قرمز باشد؟

با خارج کردن 2 مهره آبی فضای نمونه شامل 4 مهره قرمز و 3 مهره آبی می‌شود بنابراین احتمال قرمز بودن مهره آخر $\frac{4}{7}$ است.

۴ـ کیسه‌ای شامل 4 مهره قرمز و 5 مهره آبی است، یک مهره را از آن خارج می‌کنیم سپس مهره‌ای دیگر از آن خارج می‌کنیم احتمال آنکه مهره آخر خارج شده قرمز باشد؟

$\frac{4}{9}\times\frac{3}{8}+\frac{5}{9}\times\frac{4}{8}=\frac{32}{72}$

قرمز اولی، قرمز دومی، آبی اولی، قرمز دومی

۵ـ کیسه‌ای شامل 4 مهره قرمز و 5 مهره مشکی است، یک مهره از آن خارج کرده و به همراه یک مهره همرنگ آن دو بار داخل ظرف می‌گذاریم سپس مهره‌ای خارج می‌کنیم احتمال آنکه این مهره مشکی باشد؟

$\frac{5}{9}\times\frac{6}{10}+\frac{4}{9}\times\frac{5}{10}=\frac{50}{90}$

مشکی اولی، مشکی دومی، قرمز اولی، مشکی دومی

۶ـ کیسه‌ای شامل 5 مهره قرمز و 4 مهره آبی است، تاسی را پرتاب می‌کنیم اگر 2 آمد، 2 مهره قرمز در غیر این صورت 3 مهره آبی به کیسه اضافه می‌کنیم، سپس مهره‌ای از ظرف خارج می‌کنیم احتمال آنکه این مهره قرمز باشد؟

$$\frac{1}{6} \times \frac{7}{11} + \frac{5}{6} \times \frac{5}{12}$$

تاس 2 ظاهر شود — قرمز — تاس 2 ظاهر نشود

۷ـ کیسه A شامل 5 مهره قرمز و 2 مهره آبی و کیسه B شامل 7 مهره قرمز و 5 مهره آبی است مهره‌ای از ظرف A خارج کرده‌ایم و به ظرف B ریخته‌ایم، سپس مهره‌ای از ظرف B خارج می کنیم، مطلوبست احتمال اینکه این مهره قرمز باشد.

$$\frac{2}{7} \times \frac{7}{13} + \frac{5}{7} \times \frac{8}{13} = \frac{54}{91}$$

آبی (A) — قرمز (B) — قرمز (A) — قرمز (B)

مثال: در پرتاب دو تاس چقدر احتمال دارد مجموع دو شماره 7 بشود.

حل: تعداد حالات ممکن برابر $6 \times 6 = 36$ می‌باشد و

$$A = \{(1,6)\ (6,1)\ (2,5)\ (5,2)\ (3,4)\ (4,3)\}$$

$$P(A) = \frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6}$$

دقت شود به همین ترتیب می‌توان احتمال آنکه مجموع دو شماره a بشود را یافت در صورتی که $a = 2,3,4,...,11,12$ باشد در واقع :

a	2	3	4	5	6	7	8	9	10	11	12
احتمال مجموع برابر a	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

یعنی در پرتاب 2 تاس احتمال آنکه مجموع دو شماره 7 بشود بیشتراز سایر مقادیر است.

مثال: در پرتاب 5 تاس چقدر احتمال دارد :

(الف) همه شماره‌ها فرد باشند :

حل:

$$\frac{3 \times 3 \times 3 \times 3 \times 3}{6 \times 6 \times 6 \times 6 \times 6} = \frac{1}{32}$$

(ب) شماره‌های اول و سوم و پنجم یکسان باشند :

$$\frac{6 \times 6 \times 1 \times 6 \times 1}{6^5} = \frac{1}{36}$$

(ج) فقط شماره‌های اول و سوم و پنجم یکسان باشند :

$$\frac{6 \times 5 \times 1 \times 4 \times 1}{6^5} = \frac{20}{6^4}$$

(د) فقط سه شماره یکسان باشند :

$$\frac{\binom{5}{3} 6 \times 5 \times 1 \times 4 \times 1}{6^5} = \frac{200}{6^4}$$

(هـ) همه شماره‌ها متفاوت باشند :

$$\frac{6 \times 5 \times 4 \times 3 \times 2}{6^5} = \frac{61}{6^5}$$

مثال: از میان 10 مرد و 5 زن می‌خواهیم 4 نماینده انتخاب کنیم. مطلوبست محاسبه احتمال در حالت‌های زیر:

(الف) همه از یک جنسیت باشند :

$$\frac{\binom{5}{4}+\binom{10}{4}}{\binom{15}{2}}$$

(ب) دو نماینده زن و دو نماینده مرد باشند :

$$\frac{\binom{5}{2}\binom{10}{2}}{\binom{15}{4}}$$

(ج) لااقل یک نماینده زن باشد :

$$P(\text{لااقل یک زن}) = 1-P(\text{هیچ زن}) = 1-\frac{\binom{10}{4}}{\binom{15}{4}}$$

انواع بیان احتمال:

۱ـ احتمال کلاسیک

$$P(A)=\frac{\text{تعداد حالات مساعد}}{\text{تعداد حالات ممکن}}=\frac{m}{n}$$

۲ـ احتمال هندسی

$$P(A)=\frac{\text{طول ، سطح یا حجم مساعد}}{\text{طول، سطح یا حجم کل}}$$

۳ـ احتمال آماری . f_i مقدارفراوانی نمونه گیری‌ها $\lim_{n\to\infty} f_i = P(A)$

n تعداد نمونه‌ها است.

تبصره: اگر احتمال وقوع حادثه A ، K برابر احتمال وقوع حادثه B باشد و این دو حادثه کل را پوشانند، $(P(A)+P(B))=1$

$$P(A)=\frac{K}{k+1} \quad ,P(B)=\frac{1}{k+1}$$

قضیه ۱: اگر A، B دو حادثه باشند.

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

نتیجه ۱: اگر A و B مستقل باشند.

$$P(A\cup B)=P(A)+P(B)-P(A)P(B)$$

مثال: احتمال این که یک مساله ریاضی را حسن حل کند، 0.4 و احتمال این که حسین حل کند، 0.5 است. احتمال این که مساله حل شود برابر است با:

۱) 0.2 ۲) 0.8 ۳) 0.16 ٤) 0.36

حل:

A: پیشامد آن که حسن مسئله را حل کند.

B: پیشامد آن که حسین مسئله را حل کند.

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)=P(A)+P(B)-P(A).P(B)$$
$$=0.4+0.5-(0.4\times 0.5)=0.7$$

گزینه ۲ صحیح می‌باشد.

مثال: یک شرکت حفاری نفت فقط امکانات لازم برای حفر دو چاه را دارد. اگر در حفر اولین چاه به نفت برسد کار را تمام می‌کند وگرنه چاه دوم را حفر می‌کند. اگر احتمال این که در حفر هر چاه به نتیجه برسد، 0.2 باشد، احتمال این که شرکت حفاری به نتیجه برسد کدام است؟ (حفاری چاه‌ها به طور مستقل از هم صورت می‌گیرد)

۱) 0.2 ۲) 0.8 ۳) 0.16 ٤) 0.36

حل:

P (حفاری به نتیجه برسد) =P (حفاری اول به نتیجه برسد) + P (حفاری دوم به نتیجه برسد)

$$=0.2+(0.8\times 0.2)=0.36$$

گزینه ٤ صحیح می‌باشد.

نتیجه ۲: اگر A و B ناسازگار باشند.

$$P(A\cup B)=P(A)+P(B)$$

نتیجه ۳: اگر $A_1, A_2, \ldots\ldots A_K$ ناسازگار باشند.

$$P(A_1\cup A_2\ldots\ldots\cup A_K)=P(A_1)+\ldots\ldots\ldots+P(A_K)$$

مثال: اگر $A_1, A_2, \ldots, A_n$ رویدادهای با $A_i\cap A_j=\varnothing$ باشند، در این صورت:

۱) $P\left(\bigcup_{n=1}^{\infty}A_n\right)=0$ ۲) $P\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P(A_n)$

۳) $P\left(\bigcup_{n=1}^{\infty}A_n\right)=P(A_1)$ ٤) $P\left(\bigcup_{n=1}^{\infty}A_n\right)>\sum_{n=1}^{\infty}P(A_n)$

حل:

گزینه ۲ صحیح می‌باشد.

تعریف : اگر A پیشامد وقوع یک حادثه باشد. آنگاه مکمل A که با A' یا A^C یا $\overline{A}$ نشان داده می‌شود. عدم وقوع حادثه را نشان می‌دهد. روابط زیر برای آنها برقرار است: (A و A' ناسازگارند.)

$$1)\ P(A\cup A')=P(A)+P(A')$$

$$2)P(A)+P(A')=1 \Rightarrow P(A)=1-P(A'),P(A')=1-P(A)$$

(A و A' یک گروه کامل حوادث هستند.)

$$3)(A\cup B)'=(A'\cap B')$$
$$(A'\cup B')'=(A\cap B)$$
$$4)(A\cap B)'=(A'\cup B')$$
$$(A'\cap B')'=(A\cup B)$$

نکته مهم: اگر A و B دو حادثه باشند. روابط زیر چون $(A\cap B')$ و $(A\cap B)$ ناسازگارند برقرار می‌باشند:

$$1)\ P(A)=P(A\cap B)+P(A\cap B')$$

$$2)\ P(B)=P(B\cap A)+P(B\cap A')$$

نکته: روابط ۱ و ۲ برای A' و B' نیز برقرار است یعنی:

$$4)\ P(A')=P(A'\cap B)+P(A'\cap B')$$

$$5)\ P(B')=P(B'\cap A)+P(B'\cap A')$$

نکته: اگر از بین جفت‌های (A,B) ، (A,B')،(A',B') ، (A',B)، یک جفت مستقل باشند آن‌گاه بقیه نیز مستقل‌اند.

مثال: اگر $P(A)=0.3$ و $P(B)=0.2$ ، A و B مستقل باشند، $P(\bar{A}\cap\bar{B})$ برابر است با:

۱) 0.44 ۲) 0.50 ۳) 0.56 ۴) 0.667

حل:

A و B مستقل $\rightarrow P(\bar{A}\cap\bar{B})=P(\bar{A})P(\bar{B})=[1-P(A)].[1-P(B)]=0.7\times0.8=0.56$

گزینه ۳ صحیح می‌باشد.

نکته: احتمال تفاضل دو پیشامد:

$$P(B-A)=\begin{cases}P(B)-P(A\cap B)\\P(B\cap A')\end{cases}$$ احتمال وقوع فقط B

$$P(A-B)=\begin{cases}P(A)-P(A\cap B)\\P(A\cap B')\end{cases}$$ احتمال وقوع فقط A

تفاضل متقارن = احتمال وقوع فقط یک حادثه بین A و B=
$$P(A-B)+P(B-A)=P(A\cap B')+P(A'\cap B)$$

مثال: سه نفر C, B , A به ترتیب با احتمال 0.4 , 0.7 ,0.5 یک مساله را حل می‌کنند مطلوب است:

الف) احتمال آن‌که فقط یکی مساله را حل کند.

$$AB'C'+A'BC'+A'B'C$$

$$P=0.5\times0.3\times0.6+0.5\times\times0.7\times0.6+0.5\times0.3\times0.4=0.36$$

ب) احتمال آن‌که مساله حل شود.

حل شدن مساله بدین معنی است که حداقل یکی مساله را حل کند. بنابراین از متمم آن استفاده می‌کنیم.

$$P=1-(1-0.5)(1-0.7)(1-0.4)=1-0.5\times0.3\times0.6=0.91$$

نکته : معمولاً وقتی به $\cap$ بر می‌خوریم ، مستقل بودن را چک می‌کنیم . و وقتی به $\cup$ بر می‌خوریم ناسازگار بودن را چک می‌کنیم .

مثال: احتمال به صدا درآمدن هر یک از سه آژیر خطر مستقلی که در یک فروشگاه نصب شده‌اند. به هنگام آتش‌سوزی برابر 0.95 است. احتمال آن‌که به هنگام بروز آتش‌سوزی حداقل یکی از سه آژیر خطر به صدا درآید، چقدر است؟

۱) 0.15 ۲) 0.95^3 ۳) $1-(0.05)^3$ ٤) $1-(0.95)^3$

حل:

A: پیشامد آن که حداقل یکی از سه آژیر خطر به صدا درآید.

A' : پیشامد آن که هیچ یک از سه آژیر خطر به صدا در نیاید.

$$P(A)=1-P(A')=1-(0.05\times0.05\times0.05)=1-(0.05)^3$$

گزینه ۳ صحیح می باشد.

نکته:

هر دو اتفاق بیفتند	$A\cap B$
فقط A اتفاق افتد	$A\cap B'$
فقط B اتفاق افتد	$A'\cap B$
هیچ کدام اتفاق نیفتد	$A'\cap B'$

مثال مهم:

فرض کنید احتمال آنکه در بیست سال آینده زن و شوهری زنده بمانند به ترتیب 0.7 و 0.4 باشد مطلوبست محاسبه احتمال:

زنده ماندن شوهر=B زنده مانده مرد=A

الف) هر دو زنده بمانند: $P(A\cap B)=P(A)\times P(B)=0.28$

ب) هیچکدام زنده نمانند: $P(A'\cap B')=P(A')\times P(B')=0.3\times0.6=0.18$

ج) احتمال آنکه حداقل یکی زنده بماند (در بیست سال آینده شخصی زنده بماند)

راه حل اول $P(A\cup B)=1-P(A'\cap B')=1-P(\text{هیچکدام زنده نمانند})=1-0.3\times0.6=0.82$

راه حل دوم $P(A\cup B)=P(A)+P(B)-P(A\cap B)=0.4+0.7-0.4\times0.7=0.82$

راه حل اول توصیه می‌شود

د) فقط یکی زنده بماند $P(A'\cap B)+P(A\cap B')=0.4\times0.3+0.6\times0.7=0.54$

ه) فقط زن زنده بماند. $P(A'\cap B)=0.6\times0.7=0.42$

و) زن زنده بماند. $P(B)=0.7$

مثال: کدام یک از عبارات زیر بیان قانون اعداد بزرگ به صورت برنولی می‌باشد؟

۱) $\lim_{n\to\infty}P\left(\left|\frac{m}{n}-P\right|<\varepsilon\right)=0$ ۲) $\lim_{n\to\infty}P\left(\left|\frac{m}{n}-P\right|\geq\varepsilon\right)=1$

۳) $\lim_{n\to\infty} P\left(\left|\frac{m}{n}-P\right|<\varepsilon\right)=1$

٤) $\lim_{n\to\infty} P\left(\left|\frac{m}{n}-P\right|\geq\varepsilon\right)=0$

حل:

گزینه ۲ صحیح می‌باشد.

مثال: در آزمایش حادثه A باعث وقوع حادثه B می‌گردد. کدام یک از عبارات زیر درباره احتمال‌های این حوادث صحیح است؟

۱) $P(A)>P(B)$ ۲) $P(A)\leq P(B)$ ۳) $P(A)\neq P(B)$ ٤) $P(A)\geq P(B)$

حل:

چون حادثه A باعث وقوع حادثه B شده است پس:

$$A\subset B\Rightarrow P(A)\leq P(B)$$

گزینه ۲ صحیح می‌باشد.

احتمال شرطی:

بنابر تعریف احتمال وقوع حادثه A ، به شرط آنکه بدانیم حادثه B رخ داده، بصورت زیر بیان می‌شود:

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \rightarrow$$ احتمال A به شرط B

همچنین می‌توان احتمال وقوع حادثه B را به شرط وقوع حادثه A بصورت زیر بیان کرد:

$$P(B|A)=\frac{P(A\cap B)}{P(A)} \rightarrow$$ احتمال B به شرط A

نکته ۱: در صورتی که حوادث A و B مستقل ، ناسازگار ، یا وابسته باشند به نتایج زیر می‌رسیم:

$$P(A|B)=\begin{cases} 0 & \text{ناسازگار} \\ P(A) & \text{مستقل} \\ \frac{P(A\cap B)}{P(B)} & \text{وابسته} \end{cases}$$

$$P(B|A)=\begin{cases} 0 & \text{ناسازگار} \\ P(B) & \text{مستقل} \\ \frac{P(A\cap B)}{P(A)} & \text{وابسته} \end{cases}$$

نکته ۲: $\boxed{P(A'|B)=1-P(A|B)}$

مثال‌ها:

۱-اگر $P(A)=0.3$ و $P(B)=0.5$ و $P(A|B)=0.3$ باشد. می‌توان گفت A و B هر دو:

۱) مستقل ✓ ۲) ناسازگار ۳) وابسته ۴) شرطی‌اند

۲-اگر $P(A)=0.3$ و $P(B)=0.7$ و $P(A|B)=0$ باشند. می‌توان گفت A و B هر دو:

۱) مستقل ۲) ناسازگار ✓ ۳) وابسته ۴) شرطی‌اند

۳-اگر $P(A)=0.3$ و $P(B)=0.7$ و $P(A|B)=0.1$ باشد می‌توان گفت A و B هر دو:

۱) مستقل ۲) ناسازگار ۳) وابسته ✓ ۴) مکمل

۴-اگر $P(A)=0.4$ و $P(B)=0.6$ و $P(B|A)=0.1$ باشد آنگاه $P(A|B)$ کدام است؟

۱) 0.0153 ۲) 0.04 ۳) 0.05 ۴) 0.667 ✓

$$P(B|A)=\frac{P(A\cap B)}{P(A)} \Rightarrow P(A\cap B)=0.1\times 0.4=0.04$$

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{0.04}{0.6}=0.667$$

۵-اگر $P(A) = 0.5$ و $P(B) = 0.4$ و $P(A \mid B) = 0.1$ باشد آنگاه؟ $P(A\cup B) = ?$ کدام است؟

۱) 0.75　۲) 0.86 ✓　۳) 0.8　۴) 0.9

$$P(A \mid B) = \frac{P(A\cap B)}{P(B)} \longrightarrow P(A\cap B) = 0.1\times 0.4 = 0.04$$

$$P(A\cup B) = P(A)+P(B)-P(A\cap B) = 0.5+0.4-0.04 = \underline{0.86}$$

۶-با توجه به جدول توام زیر مطلوبست احتمال A به شرط B:

	B	$\bar{B}$
A	?	0.2
$\bar{A}$	0.1	0.5

$\longrightarrow$ جمع احتمال‌ها باید 1 شود

	B	$\bar{B}$
A	0.2	0.2
$\bar{A}$	0.1	0.5

$\longrightarrow$

$P(A\cap B)$

	B	$\bar{B}$	
A	0.2	0.2	0.4 $P(A)$
$\bar{A}$	0.1	0.5	0.6 $P(\bar{A})$
	0.3 $P(B)$	0.7 $P(\bar{B})$	1

$$P(A \mid B) = \frac{P(A\cap B)}{P(B)} = \frac{0.2}{0.3} = 0.667$$

۱) 0.04　۲) 0.2　۳) 0.333　۴) 0.667 ✓

۷-اگر $P(A) = \frac{1}{5}$ و $P(B) = \frac{1}{2}$ و $P(A \mid B) = \frac{1}{4}$ باشد، $P(A'\cup B')$ کدام است؟

۱) $\frac{19}{20}$　۲) $\frac{7}{8}$ ✓　۳) $\frac{21}{20}$　۴) $\frac{3}{8}$

$$P(A \mid B) = \frac{P(A\cap B)}{P(B)} \Rightarrow P(A\cap B) = \frac{1}{8}$$

$$P(A'\cup B') = P(A\cap B)' = 1-P(A\cap B) = 1-\frac{1}{8} = \frac{7}{8}$$

۸-اگر برای دو پیشامد A و B داشته باشیم $P(A) = \frac{1}{2}$ و $P(A|B) = \frac{1}{3}$ و $P(B|A') = \frac{1}{4}$ مقدار $P(B)$ کدام است؟

۱) $\frac{3}{16}$ ✓　۲) $\frac{1}{4}$　۳) $\frac{1}{2}$　۴) $\frac{3}{4}$

$$P(A \mid B) = \frac{P(A\cap B)}{P(B)} \Rightarrow P(A\cap B) = \frac{1}{3}P(B) \qquad (I)$$

$$P(B \mid A') = \frac{P(A'\cap B)}{P(A')} \Rightarrow P(A'\cap B) = \frac{1}{8} \qquad (II)$$

می‌دانیم $P(B) = P(A\cap B)+P(A'\cap B) \overset{II,I}{\Rightarrow} P(B) = \frac{1}{3}P(B)+\frac{1}{8}$

$$\boxed{P(B) = \frac{3}{16}}$$

۹-با فرض آن که احتمال آمدن برف در امروز 0.2 و فردا 0.22 باشد. احتمال برف آمدن فردا به شرط آن که امروز برف بیاید، 0.7 است.

احتمال برف نیامدن فردا به شرط آن‌که امروز برف نیاید، چقدر است؟

$P(B|A)=0.7$, $P(A)=0.2$, $P(B)=0.22$ — $B =$ برف آمدن فردا — $A =$ برف آمدن امروز

۱) 0.3 ۲) 0.72✓ ۳) 0.78 ۴) 0.9

$$P(B'|A')=\frac{P(A'\cap B')}{P(A')}=?$$

$$(\mathrm{I})\qquad P(A')=1-P(A)=1-0.2=\boxed{0.8}$$

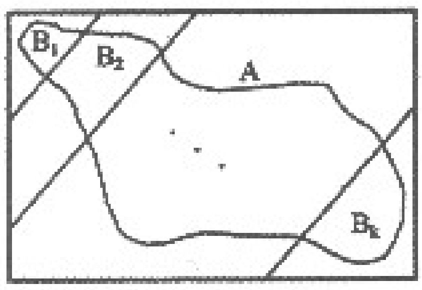
$$(\mathrm{II})\qquad P(B\,|\,A)=\frac{P(A\cap B)}{P(A)}\ \rightarrow\ P(A\cap B)=0.4$$

$$P(A'\cap B')=P(A\cup B)'=1-P(A\cup B)=1-\left[P(A)+P(B)-P(A\cap B)\right]=\boxed{0.72}$$

$$\mathrm{I}\ ,\ \mathrm{II}\ \Rightarrow\ P(B'\,|\,A')=\frac{P(A'\cap B')}{P(A')}=\frac{0.72}{0.8}=0.9$$

☑ احتمال متوسط:

هر گاه حادثه A در نتیجه وقوع هر یک از حوادث B_1 ، B_2 ،...، B_k بتواند اتفاق بیفتد آنگاه وقوع حادثه A به طور متوسط به شرح زیر بررسی می‌شود.



$$P(A)=P(B_1\cap A)+P(B_2\cap A)+\dots\dots+P(B_k\cap A)$$
$$=\sum_{i=1}^{k}P(A\cap B_i)$$
$$P(A)=\sum_{i=1}^{k}P(B_i)P(A\,|\,B_i)$$

مثال:

۱-اگر $P(A)=0.2$ و $P(B)=0.4$ و $P(E\,|\,A)=P(E\,|\,B)=0.1$ مطلوبست محاسبه $P(E)$؟

$$P(E)=\overbrace{P(A)\times P(E|A)}^{P(A\cap E)}+\overbrace{P(B)\times P(E|B)}^{P(B\cap E)}=0.2\times 0.1+0.4\times 0.1=\underline{0.06}$$

۱) 0.2 ۲) 0.1 ۳) 0.06✓ ۴) 0.7

۲-اگر $P(A)=0.3$ و $P(B)=0.4$ و $P(E\,|\,A)=0.1$ و $P\left(E^c\,|\,B\right)=0.8$ مطلوبست $P(E)$؟

$$P\left(E^c\,|B\right)=1-P(E|B)\ \Rightarrow\ P(E|B)=0.2$$

$$P(E)=\underbrace{P(A)\times P(E|A)}_{P(A\cap E)}+\underbrace{P(B)\times P(E|B)}_{P(B\cap E)}==0.3\times 0.1+0.4\times 0.2=0.11$$

۱) 0.11✓ ۲) 0.18 ۳) 0.3 ۴) 0.35

۳-اگر $P(A) = P(B) = 0.4$ و $P(G \mid B) = P(G \mid A) = 0.1$ مقدار $P(G)$ چقدر است؟

$$P(G) = P(A) \times P(G \mid A) + P(B) \times P(G \mid B) = 0.4 \times 0.1 + 0.4 \times 0.1 = 0.08$$

۱) 0.08 ✓ ۲) 0.21 ۳) 0.51 ۴) 0.9

۴-دو تلفنچی شماره ۱ و شماره ۲ به ترتیب 40 % و 60 % تلفن‌های شرکت را وصل می‌کنند، تلفنچی شماره ۱ در 0.02 موارد و تلفنچی شماره ۲ در 0.05 موارد دچار خطا می‌شوند چند درصد تلفن‌های شرکت اشتباهاً وصل شده‌اند؟

۱) % 2.7
۲) % 3.8 ✓
۳) % 7
۴) % 1.07

(۱) % 40
- اشتباه 0.02
- صحیح 0.98

$\Rightarrow 0.4 \times 0.02 + 0.6 \times 0.05 = \%3.8$

(۲) % 60
- اشتباه 0.05
- صحیح 0.95

۵-حسابدار رتبه یک 0.60 حساب‌های یک شرکت را ثبت می‌کند و حسابدار رتبه دو 0.40 حساب‌های یک شرکت را ثبت می‌کند. هر یک از آن‌ها در 0.02 موارد در ثبت حساب‌های خود دچار اشتباه می‌شوند. احتمال آن‌که حساب‌های ماه گذشته شرکت درست ثبت شده باشند، چقدر است؟

(۱) % 60
- اشتباه 0.02
- صحیح 0.98

(۲) % 40
- اشتباه 0.02
- صحیح 0.98

روش ۱: $P(\text{ثبت درست}) = 0.60 \times 0.98 + 0.40 \times 0.98 = 0.98$

روش ۲: $P(\text{ثبت غلط}) = 0.60 \times 0.02 + 0.40 \times 0.02 = 0.02$

$P(\text{ثبت درست}) = 1 - P(\text{ثبت غلط}) = 1 - (0.60 \times 0.02 + 0.40 \times 0.02) = 0.98$

۱) 0.98 ✓ ۲) 0.80 ۳) 0.66 ۴) 0.96

متغیر تصادفی

متغیر تصادفی کمیتی است که مقادیر خود را با احتمال دریافت می‌کند.

مثال: x = نتایج پرتاپ تاس

$$x=1 \rightarrow P(x=1)=\frac{1}{6}$$

$$\vdots$$

$$x=6 \rightarrow P(x=6)=\frac{1}{6}$$

✓ انواع متغیرهای تصادفی:

۱- **متغیر تصادفی گسسته:** تعداد نقاط متناهی یا نامتناهی شمارش‌پذیر

۲- **متغیر تصادفی پیوسته:** هر فاصله زمانی یا مکانی بین دو نقطه دلخواه a و b

مثال:

گسسته: تعداد مشتریان در یک ساعت – تعداد آزمایشات لازم برای اولین پیروزی – ...

پیوسته: مدت زمان لازم برای اولین اتفاق – ...

✓ تابع توزیع متغیر تصادفی:

۱- تابع توزیع احتمال گسسته:

فرمول یا جدول شامل مقادیر مختلف یک متغیر تصادفی گسسته x به همراه احتمالات متناظر آن، تابع توزیع گسسته نامیده شده و با f(x) نشان داده و باید خواص زیر را داشته باشد.

$1)\ \forall x\ \ f(x)=P(x)$ f(x) در هر نقطه احتمال در آن نقطه است.

$2)\sum f(x)=\sum P(x)=1$

مثال: با توجه به جدول موارد خواسته شده را حساب کنید:

x_i	0	1	2	3	
$P(x_i)=f(x_i)$	0.1	a	0.2	0.3	$\sum f_i=1$

الف) a=?

ب) $P(x=3)$ یا $f(x=3)$

ج) $P(x>2)$

د) $f\left(x>\sqrt{5}\right)$

هـ) $f(x>5)$

و) $f(1.5<x<2.5)$

الف) $\sum f=1 \Rightarrow 0.1+a+0.2+0.3=1$ $\Rightarrow a=0.4$	ب) $P(x=3)=0.3$
ج) $P(x>2)=P(x=3)=0.3$	د) $P(x>\sqrt{5})=P(x=3)=0.3$
هـ) $f(x>5)=0$	و) $f(1.5<x<2.5)=f(x=2)=0.2$

نکته: محاسبه تابع توزیع $y=g(x)$ از روی تابع توزیع گسسته $f(x)$:

برای این تبدیل فقط $y=g(x)$ را بدست می‌آوریم ولی $f(x)$ تغییر نمی‌کند به جز در نقاطی که $g(x)$ یکسان می‌شود که در آن نقاط $f(x)$ها را با هم جمع می‌کنیم.

x_i	-1	0	1	2	
$f(x_i)$	0.1	0.2	0.3	0.4	$\sum f(x)=1$

$\overset{y=x^2}{\Rightarrow}$

$y=x^2$	0	1	4	
$f(y)$	0.2	0.4	0.4	$\sum f(y)=1$

x_i	0	1	2	3	
$f(x_i)$	0.1	0.2	0.3	0.4	$\sum f(x)=1$

$\overset{y=x^2}{\Rightarrow}$

$y=x^2$	0	1	4	9	
$f(y)$	0.1	0.2	0.3	0.1	$\sum f(y)=1$

۲- تابع چگالی متغیر پیوسته:

تابع توزیع f را تابع چگالی متغیر تصادفی پیوسته x گوئیم هرگاه شرایط زیر را داشته باشد:

$f(x)$ تابع چگالی متغیر تصادفی پیوسته x در فاصله $[a,b]$:

۱) $f(x)\geq 0$

۲) $\int_a^b f(x)dx=1$

۳) به ازاء نقاط پیوسته $x=\alpha \Rightarrow P(x=\alpha)=0$ (احتمال در هر نقطه پیوسته صفر است)

۴) (احتمال در هر فاصله انتگرال در آن فاصله است.) $P(\alpha\leq x\leq\beta)=\int_\alpha^\beta f(x)\,dx$

مثال: تابع چگالی احتمال $f(x)$ که توزیع احتمال یک متغیر تصادفی پیوسته را توصیف می‌کند، کدام ویژگی را ندارد؟

۱) به ازای هر نقطه مانند a ، $P(x=a)\neq 0$.
۲) $f(x)\geq 0$ تابع چگالی مثبت یا صفر است.
۳) مساحت کل زیر منحنی چگالی برابر یک است.
۴) سطح زیر منحنی چگالی بین a و b برابر است با $P(a<x<b)$.

گزینه ۱ صحیح می‌باشد.

مثال: تابع احتمال متغیر تصادفی به صورت زیر داده شده است، مقدار k کدام است؟

$$f(x)=\begin{cases}\dfrac{x}{k} & 0\leq x\leq 4\\ 0 & \text{سایر مقادیر}\end{cases}$$

۱) 2　　۲) 4　　۳) 8　　۴) 16

$$\int_{-\infty}^{+\infty} f(x)dx=\int_0^4 \frac{x}{k}dx=1 \Rightarrow \frac{x^2}{2k}\Big]_0^4=1 \Rightarrow k=8$$

گزینه ۳ صحیح می‌باشد.

مثال: تابع چگالی احتمال متغیر تصادفی X به صورت زیر داده شده است:

$$f(x)=\begin{cases} x & 0\le x\le 1 \\ k-x & 1\le x\le 2 \\ 0 & \text{برای سایر مقادیر } x \end{cases}$$

مقدار **k** چقدر است؟

۱) $\frac{1}{2}$ ۲) 1 ۳) 2 ۴) 3

$$\int_{-\infty}^{+\infty} f(x)dx=1 \Rightarrow \int_0^1 x\,dx+\int_1^2 (k-x)dx=1 \Rightarrow \frac{1}{2}x^2\Big]_0^1+\left(kx-\frac{1}{2}x^2\right)\Big]_1^2=1$$

$$\Rightarrow \frac{1}{2}+2k-2-k+\frac{1}{2}=1 \Rightarrow k=2$$

گزینه ۳ صحیح می‌باشد.

مثال: چگالی احتمال‌های کمیت تصادفی x توسط تابع زیر بیان شده است:

$$f(x)=\begin{cases} \frac{2}{900}x & 0<x\le 30 \\ 0 & \text{سایر} \end{cases}$$

احتمال آنکه متغیر تصادفی x مقدار خود را در فاصله $(5,10)$ اختیار کند چقدر است؟

۱) $\frac{65}{900}$ ۲) $\frac{70}{900}$

۳) $\frac{75}{900}$ ۴) $\frac{92}{900}$

$$P(5<x<10)=\int_5^{10}\frac{2}{900}x\,dx=\left[\frac{x^2}{900}\right]_5^{10}=\frac{75}{900}$$

گزینه ۳ صحیح می‌باشد.

مثال: کمیت تصادفی X در جامعه‌ای بر طبق قانون نمایی توزیع شده است:

$$f(x)=\frac{1}{20}e^{-\frac{x}{20}} \quad 0<x<\infty$$

احتمال این که کمیت تصادفی X ، مقداری مساوی با 125 ، اختیار کند، چقدر است؟

۱) $\frac{1}{20}$ ۲) 1 ۳) 0 ۴) 0.1

$P(X=125)=0$.

گزینه ۳ صحیح می‌باشد.

مثال: X به عنوان یک متغیر تصادفی معرف طول عمر لامپی است که بین صفر تا 160 ساعت کار می‌کند. احتمال این‌که این لامپ دقیقاً 80 ساعت کار کند برابر است با:

۱) صفر

۲) 0.5

۳) این احتمال را تا زمانی که تابع چگالی X مشخص نباشد، نمی‌توان محاسبه کرد.

۴) این احتمال را تا زمانی که میانگین و واریانس X مشخص نباشد، نمی‌توان محاسبه کرد.

متغیر تصادفی X (طول عمر لامپ) پیوسته است، لذا:

$$P(X=80)=0$$

گزینه ۱ صحیح می‌باشد.

نکته: در تابع چگالی متغیر تصادفی پیوسته x رابطه زیر برقرار است:

$$P(\alpha \le x \le \beta)=P\left(\cancel{x=\alpha}_{0}\right)+P(\alpha<x<\beta)+P\left(\cancel{x=\beta}_{0}\right)=P(\alpha<x<\beta)$$

***نکته:** در نقاط پیوسته $P(x=\alpha)=0$

تابع توزیع تجمعی یا C df :

طبق روابط زیر برای هر متغیر تصادفی گسسته یا پیوسته x با تابع توزیع یا چگالی F(x) تابع توزیع تجمعی را با $F_x(x)$ نشان می‌دهند.(یا F(x)) :

گسسته $F_x(x)=P(X\le x)=\sum_{t\le x} f_x(t)$

پیوسته $F_x(x)=P(X\le x)=\int_{\text{حد پایین}}^{x} f(t)\ d\ t$

مثال حالت گسسته:

با توجه به جدول مقابل مطلوب است: $F_x(\sqrt{5})$

x	-1	0	1	2	4
f(x)	0.1	0.2	0.1	A	0.4

از آن‌جایی که $\sum f(x)=1$ می‌باشد در نتیجه a=0.2، و می‌دانیم که $f_x(\sqrt{5})=0$ است (چون نقطه $\sqrt{5}$ وجود ندارد) همچنین بنا به تعریف داریم:

$$F_x(\sqrt{5})=\sum_{\forall x\le\sqrt{5}} f(x)=0.1+0.2+0.1+0.2=\boxed{0.6}$$

۳. تابع چگالی احتمال‌ها برای کمیت X به صورت زیر بیان شده است:

$$f(x)=\frac{2x+3}{18} \qquad 0<x<3$$

تابع توزیع کمیت تصادفی X کدام است؟

۱) $F_X(x)=\frac{2x^2+3x}{18}$ ۲) $F_X(x)=\frac{x^2+3x}{18}$ ۳) $F_X(x)=\frac{x^2}{18}$ ۴) $F_X(x)=\frac{x+3x^2}{18}$

$$F(x)=\int_0^x f(x)dx=\int_0^x \frac{2x+3}{8}dx=\frac{x^2+3x}{18}$$

گزینه ۲ صحیح می‌باشد.

تابع توزیع تجمعی دارای خواص زیر است:

الف) $F_X(x)$ غیر نزولی است.

ب) $0\le F_X(x)\le 1$

ج) $F_X(x)$ از سمت راست پیوسته است.

د) F(حد پایین)=0 و F(حد بالا)=1

مثال: اگر تابع توزیع احتمال تجمعی متغیر تصادفی X به صورت زیر باشد، مقدار k چقدر است؟

$$f(x)=\begin{cases}0 & x<1\\ \frac{1}{k}(x-1)^3 & 1\le x<3\\ 1 & 3\le x\end{cases}$$

۱) 1 ۲) 2 ۳) 4 ۴) 8

باید دقت کنیم که در اینجا f، F بوده، چون F (حد پائین) = 0 و F (حد بالا) = 1. در نتیجه داریم:

$$F(3)=1 \Rightarrow \frac{1}{K}(3-1)^3=1 \Rightarrow k=8$$

گزینه ۴ صحیح می‌باشد.

نکته : در مورد متغیر تصادفی پیوسته در صورتیکه $F_X(x)$ پیوسته باشد. $F'_X(x)=f(x)$

مثال: تابع توزیع (تجمعی احتمال) متغیر تصادفی x به قرار زیر است:

$$F_X(x)=\begin{cases}0 & x\le 0\\ \frac{11}{25}x^2-2x & 0<x\le 5\\ 1 & x>5\end{cases}$$

تابع چگالی احتمال x کدام است؟

۱) $\frac{11}{25}x-2$ ۲) $\frac{11}{75}x^3-x^2$ ۳) $\frac{11}{50}x+2$ ۴) $\frac{22}{25}x-2$

$$f(x)=\left(F_X(x)\right)'=\frac{22}{25}x-2$$

گزینه ۴ صحیح می‌باشد.

نکته :

$$P(a < x \le b) = F(b) - F(a)$$

مثال: اگر تابع توزیع کمیت تصادفی X به صورت زیر باشد:

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x}{10} & 0 < x \le 10 \\ 1 & x > 10 \end{cases}$$

احتمال $P(2 < x < 8)$ کدام است؟

۱) 0.2 ۲) 0.4 ۳) 0.6 ۴) 0.8

$$P(2 < X < 8) = F(8) - F(2) = \frac{8}{10} - \frac{2}{10} = 0.6$$

گزینه ۳ صحیح می‌باشد.

نکته:

$$P(x = a) = \left| F_x(a^+) - F_x(a^-) \right|$$

که در صورت پیوسته بودن در نقطه a، $P(x = a) = 0$ است.

نکته:

$$P(a \le x \le b) = P(x = a) + P(a < x < b) + P(x = b)$$
$$= F_x(a^+) - F_x(a^-) + F_x(b^-) - F_x(a^+) + F_x(b^+) - F_x(b^-)$$
$$= F_x(b^+) - F_x(a^-)$$

مثال: اگر تابع توزیع کمیت تصادفی ناپیوسته X به صورت زیر باشد:

$$F_X(x) = \begin{cases} 0 & x \le 1 \\ \dfrac{2}{10} & 1 < x \le 4 \\ 1 & x > 4 \end{cases}$$

تابع احتمال‌های آن کدام است؟

۱)

x	1	4
$P_X(x)$	0.2	0.8

۲)

x	1	4
$P_X(x)$	0.8	0.4

۳)

x	0	1	4
$P_X(x)$	0.2	0.5	0.3

۴)

x	0	1	4
$P_X(x)$	0.3	0.5	0.2

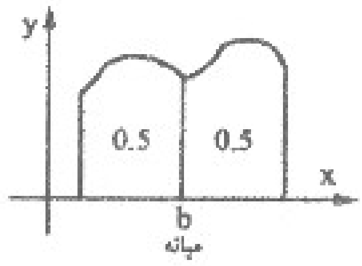
$$P(X = 1) = F(X = 1^+) - F(X = 1^-) = 0.2 - 0 = 0.2$$

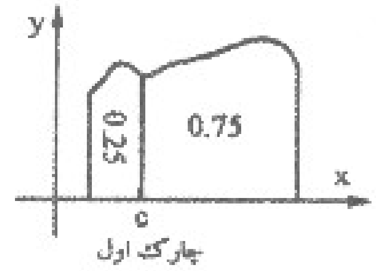
$$P(X = 4) = F(X = 4^+) - F(X = 4^-) = 1 - 0.2 = 0.8$$

گزینه ۱ صحیح می‌باشد.

نکته: محاسبه میانه -چارک- دهک و صدک

طبق تعریف برای محاسبه هر یک از موارد بالا می‌توانیم به صورت زیر عمل کنیم:





$$F_X(x) = p(X \le x) = \int_{\text{پایین}}^{x} f(x)dx = \frac{1}{2}$$ x میانه

$$= \int_{\text{پایین}}^{x} f(x)dx = \frac{a}{4} \quad a = 1,2,3$$ x چارک aام

$$= \int_{\text{پایین}}^{x} f(x)dx = \frac{a}{10} \quad a = 1,2,\dots,9$$ x دهک aام

$$= \int_{\text{پایین}}^{x} f(x)dx = \frac{a}{100} \quad a = 1,2,\dots,99$$ x صدک aام

مثال: تابع چگالی احتمال‌ها برای کمیت تصادفی X به صورت تعریف شده است:

$$f(x) = \frac{1}{2}x \qquad 0 < x < 2$$

میانه را حساب کنید.

۱) $-\sqrt{2}$ ۲) 1 ۳) $\sqrt{2}$ ۴) $\pm\sqrt{2}$

$$\int_0^{me} \varphi(x)dx = \frac{1}{2} \Rightarrow \int_0^{me} \frac{1}{2}x\,dx = \frac{x^2}{4}\Big]_0^{me} = \frac{1}{2} \Rightarrow me = \sqrt{2}$$

گزینه ۳ صحیح می‌باشد.

مثال: یک تابع توزیع احتمالی دارای چگالی f(x)=1 است. اگر حد پایین توزیع 3.4 باشد میانه توزیع چقدر است؟

۱) 3.7 ۲) 3.9 ۳) 4 ۴) 6.8

$$F(x) = \int_{-\infty}^{x} f(x)dx = \frac{1}{2} \Rightarrow \int_{3.4}^{x} dx = \frac{1}{2} \Rightarrow x\Big]_{3.4}^{x} = \frac{1}{2} \Rightarrow x = md = 3.9$$

گزینه ۲ صحیح می‌باشد.

مثال: در تابع چگالی زیر صدک 80 چقدر است؟

$$\left(f(x) = \frac{1}{8}x\,,\ 0 < x < 4\right)$$

۱) 2.38 ۲) 3.20 ۳) 3.58 ۴) 12.82

$$F(x) = \int_0^x \frac{1}{8}x\,dx = \frac{1}{16}x^2\Big]_0^x \Rightarrow F(x) = \frac{1}{16}x^2$$

$F(x)=\frac{80}{100} \Rightarrow \frac{1}{16}x^2=\frac{4}{5} \Rightarrow x=3.58 \rightarrow 80$ صدک

$\int_0^c \frac{1}{8}x\,dx=\frac{80}{100} \Rightarrow \frac{1}{16}x^2\Big]_0^c=\frac{4}{5} \Rightarrow \frac{c^2}{16}=\frac{4}{5} \Rightarrow c=3.58$

گزینه ۳ صحیح می‌باشد.

مثال: تابع توزیع کمیت تصادفی X به صورت زیر بیان شده است:

$$F_X(x)=\begin{cases}0 & x\le 2\\ \frac{x-2}{3} & 2<x\le 5\\ 1 & x>5\end{cases}$$

میانه توزیع کدام است؟

۱) 3.5 ۲) 3 ۳) 4 ۴) 4.5

$F(x)=\frac{1}{2} \Rightarrow \frac{x-2}{3}=\frac{1}{2} \Rightarrow 2x=7 \Rightarrow x=me=3.5$

گزینه ۱ صحیح می‌باشد.

محاسبه مد یا نما (mo) = محلی که بیشترین مقدار فراوانی در تابع چگالی احتمال یا تابع جرم احتمال وجود دارد و به صورت زیر محاسبه می‌شود.

$f'(x)=0 \Rightarrow x=mo$

$(F(x))''=0 \Rightarrow x=mo$

مثال: تابع توزیع $F(x)=\frac{1}{3}\left(-\frac{x^3}{3}+3x^2-5\right)$ در دامنه $0\le x\le 3$ تعریف شده است. نما (مد) آن برابر است با:

۱) 0.6 ۲) 6 ۳) 3 ۴) 4

$\varphi(x)=F'(x)=-\frac{3x^2}{9}+2x$

$\varphi'(x)=0 \Rightarrow -\frac{6x}{9}+2=0 \Rightarrow x=3 \Rightarrow mo=3$

گزینه ۳ صحیح می‌باشد.

مثال: تابع توزیع تجمعی $F_X(x)=\begin{cases}0 & x<0\\ \frac{1}{2}x^2(3-x) & 0\le x\le 1\\ 1 & x>1\end{cases}$ داده شده است. نما تابع توزیع مزبور برابر است با:

۱) 0 ۲) 1 ۳) 2 ۴) 3

$f(x)=F'(x)=\frac{6x-3X^2}{2}$ ، $f'(x)=0 \Rightarrow \frac{6-6x}{2}=0 \Rightarrow x=mo=1$

گزینه ۲ صحیح می‌باشد.

نکات مهم در مورد روابط $f_x(x), F_x(x)$

۱) در صورتیکه $F_X(x)$ داده شده باشد در حالت متغیرهای پیوسته به شرط آنکه هیچ نقطه گسستگی وجود نداشته باشد.

$$(F_x(x))' = f_x(x)$$
$$\int_{\text{پایین}}^{x} f_x(x)\,dx = F_x(x) = P(X \le x)$$

۲) در صورتیکه نقطه گسستگی داشته باشیم . برای محاسبه f_x از روی F_X :

$$f_x(x) = \begin{cases} (F(x))' & \text{نقاط پیوسته} \\ f_x(x) = p(x) = |F_x(x^+) - F_x(x^-)| & \text{نقاط گسسته} \end{cases}$$

در مثال زیر نقاط گسسته $x = 1,2,3$ و $x = 0$ پیوسته است.

مثال : با توجه به $F(x)$ مطلوبست $P(x=2)$ ؟

$$F_x(x) = \begin{cases} 0 & x<0 \\ \frac{x}{4} & 0 \le x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ \frac{x}{3} & 2 \le x < 3 \\ 1 & 3 \le x \end{cases}$$

$$f(x) = \begin{cases} \frac{1}{4} & 0<x<1 \\ \frac{1}{4} & x=1 \\ \frac{1}{6} & x=2 \\ \frac{1}{3} & 2<x<3 \end{cases}$$

با توجه به f بدست آمده در نقاط گسسته $P(x=2) = f(x=2) = \frac{1}{6}$.

امید ریاضی(میانگین – مقدار متوسط – ارزش انتظاری- $E(x) - \mu_x$ **):** امید ریاضی متغیر تصادفی x ، حد متوسطی است که انتظار می‌رود برای متغیر تصادفی x اتفاق بیافتد. در یک مجموعه از مقادیر یک جامعه امید ریاضی مرکز ثقل جامعه است؛

امید ریاضی متغیر تصادفی x به صورت زیر نشان داده میشود:

$E(x) = \sum_{\forall x} x f_x(x)$ گسسته (به شرط همگرائی $\sum$)

$E(x) = \int_{-\infty}^{+\infty} x f_x(x) dx$ پیوسته (به شرط موجود بودن $\int$)

نکته:

$$-\infty < E(x) < +\infty$$

امید ریاضی یک تابع از متغیر تصادفی x با تابع توزیع احتمال $f(x)$ به فرم زیراست :

$E[k(x)] = \sum_{\forall x} k(x) f_X(x)$ گسسته

$E[k(x)] = \int_{-\infty}^{+\infty} k(x) f_X(x)\, d_X$ پیوسته

نکته: اگر x و y دو متغیر تصادفی و a و b و c مقادیر ثابت باشند. روابط زیر برای امید ریاضی می‌توانند برقرار باشند:

$$1)\begin{cases} E(a)=a \\ E(bx)=bE(x) \end{cases} \Rightarrow E(ax+b)=aE(x)+b$$

اگر $h(x)$ و $g(x)$ توابعی از x باشند:

$$2)\ E[g(x)\pm h(x)]=E[g(x)]\pm E[h(x)]$$

$$3)\ E(\ldots(E(x)))=E(x)$$

$$4)\ E(\pm ax\pm by\pm c)=\pm aE(x)\pm bE(x)\pm c$$

نکته: دقت کنید هنگام محاسبه $E(g(x))$ تابع توزیع $f(x)$ به هیچ شکل تغییر نمی‌کند.

مثال: تابع چگالی احتمالی متغیر تصادفی پیوسته X به صورت $f(x)=1-\frac{1}{2}x$ برای $0\le x\le 2$ داده شده است. میانگین و $P(1\le X<2)$ به ترتیب از راست به چپ برابر است با:

۱) 1 و $\frac{3}{4}$ ۲) $\frac{2}{3}$ و $\frac{1}{4}$ ۳) 1 و $\frac{1}{4}$ ۴) $\frac{2}{3}$ و $\frac{3}{4}$

$$E(x)=\int_0^2 xf(x)\,dx=\int_0^2 x\left(1-\frac{1}{2}x\right)dx=\int_0^2\left(x-\frac{1}{2}x^2\right)dx$$

$$E(x)=\left(\frac{x^2}{2}-\frac{x^3}{6}\right)\Bigg]_0^2=\frac{2}{3}$$

$$P(1\le X<2)=\int_1^2\left(1-\frac{1}{2}x\right)dx=\left(x-\frac{1}{4}x^2\right)\Bigg]_1^2=\frac{1}{4}$$

گزینه ۲ صحیح می‌باشد

مثال: کدام تساوی در مورد عمل کننده امید ریاضی غلط است. (a و c مقادیر ثابت هستند)؟

۱) $E[c(a+X)]=A+cE(X)$ ۲) $E(a+X)=a+E(X)$

۳) $E[(cX)]=cE(X)$ ۴) $E(a)=a$

$$E[c(a+X)]=E(ca+cX)=E(ca)+E(cX)=ca+cE(X)$$

گزینه ۱ صحیح می‌باشد

مثال: کمیت تصادفی X در جامعه‌ای بر طبق قانون نرمال با امید ریاضی $E(X)=\mu=40$ و واریانس $D(X)=\sigma^2=16$ توزیع شده است. امید ریاضی کمیت تصادفی Y که بر طبق رابطه $Y=3X+2$ از کمیت X تبعیت می‌کند، کدام است؟

۱) 122 ۲) 40 ۳) 16 ۴) 120

$$E(Y)=E(3X+2)=E(3X)+E(2)=3E(X)+2$$

$$E(Y)=(3\times 40)+2=122$$

گزینه ۱ صحیح می‌باشد

مثال: تابع چگالی کمیت تصادفی X به صورت زیر بیان شده است:

$$f(x)=\frac{2x}{9}\quad 0<x<3$$

امید ریاضی کمیت تصادفی Y که بر طبق رابطه y=2x+1 از کمیت X تبعیت می‌کند، چقدر است؟

۱) 3.5　　۲) 4　　۳) 2　　۴) 5

$$E(X)=\int_0^3 x\,.\,\frac{2x}{9}dx=\int_0^3\frac{2}{9}x^2\,dx=\frac{2}{27}x^3\Big]_0^3=2$$

$$E(Y)=E(2x+1)=2E(x)+1=4+1=5$$

گزینه ۴ صحیح می‌باشد

مثال: تابع احتمال (قانون توزیع احتمال‌ها)، به صورت زیر تعریف شده است:

x_i	-1	0	1	2
$P(x_i)$	0.3	0.3	0.3	0.1

امید ریاضی X^2، یعنی $E(X^2)$ کدام است؟

۱) 0.04　　۲) 0.2　　۳) 1　　۴) 2

$$E(X^2)=\sum x^2.\,P(x)=(-1)^2(0.3)+0+(1)^2(0.3)+(2)^2(0.1)=1$$

گزینه ۳ صحیح می‌باشد.

مثال: متغیر تصادفی x می‌تواند یکی از سه مقدار 5 و 4 و x_3 را انتخاب کند که احتمال آن‌ها به ترتیب 0.2، 0.5 و P_3 است. اگر میانگین متغیر تصادفی x برابر 6 باشد مقدار x_3 چقدر است؟

۱) 2　　۲) 5　　۳) 7　　۴) 10

x_i	4	5	x_3	
$P_i=f_i$	0.5	0.2	P_3	$\sum P_i=1$

$$(1)\quad E(x)=\text{میانگین}=\sum x_i\times P(x_i)=6\Rightarrow 4\times0.5+5\times0.2+x_3\times P_3=6\Rightarrow x_3\times P_3=3$$

$$(2)\quad \sum P_i=1\Rightarrow 0.5+0.2+P_3=1\Rightarrow P_3=0.3$$

$$(1),(2)\Rightarrow x_3=10$$

گزینه ۴ صحیح می‌باشد.

مثال: جدول احتمال متغیر تصادفی X به صورت زیر است.

x_i	0	1	2	3
P_i	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	a

امید ریاضی X کدام است؟

۱) 1　　۲) $\frac{4}{3}$　　۳) $\frac{5}{3}$　　۴) 2

می‌دانیم جمع احتمال‌ها برابر با یک است، یعنی:

$$\sum p(x)=1\Rightarrow\frac{1}{6}+\frac{1}{4}+\frac{1}{3}+a=1\Rightarrow a=\frac{1}{4}$$

پس:

$$E(x)=\sum x_i \times P(x_i) \Rightarrow E(X)=0\times\frac{1}{6}+1\times\frac{1}{4}+2\times\frac{1}{3}+3\times\frac{1}{4}=\frac{5}{3}$$

گزینه ۳ صحیح می‌باشد.

☑ واریانس (پراش) – $V(x)$ – $Var(x)$ – $D(x)$ – $\sigma^2(x)$ – امید مجذور انحراف معیار انحرافات از میانگین

بنابر تعریف واریانس متغیر تصادفی از رابطه زیر بدست می‌آید.

$$\sigma^2=E\left[(x-E(x))^2\right]$$
$$=E(x^2)-(E(x))^2$$

خواص واریانس:

اگر a و b ثابت باشند:

$$1)\ \sigma^2(ax+b)=a^2\ \sigma^2(x)$$
$$2)\ \sigma^2(b)=0$$

در نتیجه انحراف معیار (σ) یا خطای معیار که همواره مقدار مثبتی است دارای خواص زیر خواهد بود:

$$1)\ \sigma(ax+b)=|a|\ \sigma(x)$$
$$2)\ \sigma(b)=0$$

مثال: اگر X یک متغیر تصادفی نرمال باشد، تبدیل $Y=a+bX$ که در آن a و b ثابت هستند، کمیت تصادفی Y :

۱) دارای توزیع نرمال با میانگین $a+bE(X)$ و واریانس $a^2+b^2\sigma_x^2$ است.

۲) دارای توزیع نرمال با میانگین $bE(X)$ و واریانس $b^2\sigma^2$ است

۳) دارای توزیع نرمال با میانگین $a+bE(X)$ و واریانس $b^2\sigma_x^2$ است

۴) دارای توزیع نرمال نیست ولی میانگین و واریانس آن معلوم است.

هر ترکیب خطی از توزیع نرمال همیشه دارای توزیع نرمال است.

$$E(Y)=E(a+bX)=E(a)+E(bX)=a+bE(X)$$

$$Var(Y)=Var(a+bX)=Var(a)+Var(bX)=b^2Var(X)=b^2\delta_x^2$$

گزینه ۳ صحیح می‌باشد.

مثال: اگر میانگین و انحراف معیار x برابر 2 باشد، میانگین x^2 چقدر است؟

۱) 4   ۲) 8   ۳) 16   ۴) 32

$$E(X)=\mu=2 \quad ; \quad \sigma_X=2 \Rightarrow Var(X)=4$$
$$Var(X)=E(X^2)-\mu^2 \Rightarrow 4=E(X^2)-4 \Rightarrow E(X^2)=8$$

گزینه ۲ صحیح می‌باشد.

مثال: واریانس متغیر تصادفی x با چگالی $\left\{f(x)=\frac{2}{3}\,|-1<x<\frac{1}{2}\right\}$ کدام است؟

۱) $\frac{1}{16}$ ۲) $\frac{3}{16}$ ۳) $\frac{4}{9}$ ۴) $\frac{9}{4}$

$$\delta_x^2=E(x^2)-E(x)^2=\int_{-1}^{\frac{1}{2}}x^2\times\frac{2}{3}dx-\left(\int_{-1}^{\frac{1}{2}}x\times\frac{2}{3}dx\right)^2$$

$$=\left[\frac{2}{9}x^3\right]_{-1}^{\frac{1}{2}}-\left(\left[\frac{2}{6}x^2\right]_{-1}^{\frac{1}{2}}\right)^2=\frac{3}{16}$$

گزینه ۲ صحیح می‌باشد.

مثال: متغیر تصادفی X دارای میانگین 5 و واریانس 9 می‌باشد. میانگین و واریانس $\frac{X-5}{3}$ به ترتیب (از راست به چپ) کدام است؟

۱) 0 و 1 ۲) 1 و 0 ۳) 5 و 3 ۴) 5 و 9

$$E\left(\frac{X-5}{3}\right)=\frac{1}{3}E(X)-\frac{5}{3}=\frac{1}{3}\times5-\frac{5}{3}=0$$
$$D\left(\frac{X-5}{3}\right)=\frac{1}{9}D(X)=\frac{1}{9}\times9=1$$

گزینه ۱ صحیح می‌باشد.

مثال: تابع چگالی احتمال x به صورت $f(x)=\begin{cases}\frac{2x}{c^2} & 0\le x\le c\\ 0 & \text{سایر نقاط}\end{cases}$ است. c چه مقدار باشد تا این که $\sigma_x^2=2$ گردد.

۱) c=2 ۲) c=4 ۳) c=6 ۴) c=9

$$\delta_x^2=2\Rightarrow E(x^2)-E(x)^2=\int_0^c x^2\frac{2x}{c^2}dx-\left(\int_0^c x\frac{2x}{c^2}dx\right)^2=2$$

$$\Rightarrow\left[\frac{2x^4}{4c^2}\right]_0^c-\left(\left[\frac{2}{3}\frac{x^3}{c^2}\right]_0^c\right)^2=2\ \Rightarrow\frac{c^2}{2}-\left(\frac{2}{3}c\right)^2=2$$

$$\Rightarrow\frac{c^2}{2}-\frac{4}{9}c^2=2\ \Rightarrow c=6$$

گزینه ۱ صحیح می‌باشد.

مثال: در صورتیکه $P(x)=\frac{|x|+1}{5}$ برای $x=-1,0,1$ تابع احتمال متغیر تصادفی ناپیوسته x باشد. آنگاه امید ریاضی و واریانس x به ترتیب از راست به چپ برابر چیست؟

۱) $\frac{4}{5},0$ ۲) $\frac{1}{5},1$ ۳) $\frac{2}{5},0$ ۴) $\frac{4}{25},\frac{4}{5}$

x	-1	0	1
P(x)	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

$\sum P(x)=1$

$$E(x)=\sum x_i\times P(x_i)=-1\times\frac{2}{5}+0\times\frac{1}{5}+1\times\frac{2}{5}=\boxed{0}$$

$$\delta^2(x)=E(x^2)-(E(x))^2=\sum x_i^2\times P(x_i)-\left(\sum x_i\times P(x_i)\right)^2$$
$$=\left[-1^2\times\frac{2}{5}+0^2\times\frac{1}{5}+1^2\times\frac{2}{5}\right]-\left[-1\times\frac{2}{5}+0\times\frac{1}{5}+1\times\frac{2}{5}\right]^2$$
$$=\frac{4}{5}-0^2=\boxed{\frac{4}{5}}$$

گزینه ۱ صحیح می‌باشد.

توزیع احتمال توام ، چگالی احتمال توام (مشترک)

اگر x و y دو متغیر تصادفی باشند، احتمال پیشامد همزمان این دو توزیع احتمال توأم x و y نام دارد و آن را با $f(x,y)$ نشان می‌دهند. (یا $f_{x,y}(x,y)$) و دارای خواص زیر است:

$$f\,x,y\ \geq 0 \qquad ;\forall(x,y)$$

الف) گسسته
$$\begin{cases}\sum\limits_{\forall x}\ \sum\limits_{\forall y} f(x,y)=1\\ P\{(x,y)\in A\}=\sum\limits_{\forall (x,y)\in A}\sum f(x,y)=f_{x,y}(x,y)\end{cases}$$

ب) پیوسته
$$\begin{cases}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy=1\\ P\{(x,y)\in A\}=\int\int_A f(x,y)\,dx\,dy=f_{x,y}(x,y)\end{cases}$$

مثال: دو متغیر تصادفی ناپیوسته (گسسته) X و Y با قانون توزیع (تابع احتمال‌ها) توام $P(X=x,Y=y)=f(x,y)$ در نظر گرفته می‌شود، کدام یک از روابط زیر برای توزیع فوق صادق است؟

۱) $\sum\limits_x f(x,y)=1$ ۲) $\sum\limits_x\sum\limits_y f(x,y)=1$

۳) $\sum\limits_y f(x,y)=1$ ۴) هیچ کدام

گزینه ۲ صحیح می‌باشد، گزینه ۱، $\sum\limits_x f(x,y)=f(y)$ و گزینه ۳، $\sum\limits_y f(x,y)=f(x)$ می‌باشد.

توزیع حاشیه ای یا کناره‌ای x , y : تابع توأم یا تابع چگالی احتمال توأم یکی از دو متغیر است، وقتی که تمام ناحیه تغییرات متغیر دیگر در نظر گرفته شود.

اگر x و y گسسته باشند :

توزیع حاشیه‌ای x $=f(x)=\sum\limits_{\forall y} f(x,y)$ توزیع حاشیه‌ای y $=f(y)=\sum\limits_{\forall x} f(x,y)$

اگر x ، y پیوسته باشند:

توزیع حاشیه‌ای y $=f(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx$ توزیع حاشیه‌ای x $=f(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy$

نتیجه ۱:

$$\boxed{X,Y \text{ مستقل } \Leftrightarrow f(x,y)=f(x)\times f(y)}$$

مثال: تابع احتمال زیر را در نظر بگیرید.

X \ Y	1	2
-1	$\frac{1}{4}$	$\frac{1}{4}$
0	$\frac{1}{4}$	$\frac{1}{4}$

کدام گزینه صحیح است؟

۱) دارای ارتباط مثبت هستند

۲) دارای ارتباط غیر خطی هستند

۳) دارای ارتباط منفی هستند

۴) دو متغیر مستقل هستند

x و y مستقل‌اند

X \ Y	1	2	f (x)
-1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$
f (y)	$\frac{2}{4}$	$\frac{2}{4}$	1

$$f(-1,1)=f_x(-1)f_y(1) \Rightarrow \frac{1}{4}=\frac{2}{4}\times\frac{2}{4}$$

$$f(-1,2)=f_x(-1)f_y(2) \Rightarrow \frac{1}{4}=\frac{2}{4}\times\frac{2}{4}$$

$$f(0,1)=f(0)f(1) \Rightarrow \frac{1}{4}=\frac{2}{4}\times\frac{2}{4}$$

$$f(0,2)=f(0)f(2) \Rightarrow \frac{1}{4}=\frac{2}{4}\times\frac{2}{4}$$

گزینه ۴ صحیح می‌باشد.

مثال ۱: تابع احتمال توزیع X و Y برابر است با ؟

$$P_{XY}(x,y)=\begin{cases}\frac{1}{3} & x,y=(-1,0),(0,1),(1,0)\\ 0 & \text{سایر مقادیر}\end{cases}$$

حل: با توجه به جدول زیر x,y مستقل نیستند.

y \ x	-1	0	1	f (y)
0	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
1	0	$\frac{1}{3}$	0	$\frac{1}{3}$
f(x)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

$$f(x)=\sum_{\forall y_i} f(x,y_i)$$

$$f(y)=\sum_{\forall x_i} f(x_i,y)$$

y ,x مستقل نیستند.

چون‌که برای استقلال y,x برقراری تساوی $f(x,y)=f_x(x)f_y(y)$ برای تمام xها و yها لازم می‌باشد.

مثال ۲: با توجه به جدول زیر x,y مستقل‌اند.

Y \ x	0	1	f (y)
0	0.15	0.35	0.5
0	0.15	0.35	0.5
f (x)	0.3	0.7	1

مثال پیوسته:

۱ـ با توجه به تابع چگالی توام روبرو آیا x,y مستقل‌اند؟

$$\begin{cases} f(x,y)=e^{-(x,y)} \\ 0 \le x \\ 0 \le y \end{cases}$$

$$f(x)=\int_0^{+\infty} f(x,y)dy=e^{-x}$$

$$f(y)=\int_0^{+\infty} f(x,y)dx=e^{-y}$$

با توجه به $f(x),f(y)$ بدست آمده دیده می‌شود که $f(x,y)=f(x)f(y)$ **در نتیجه** x,y مستقل‌اند.

توزیع احتمال شرطی : توزیع احتمال شرطی متغیر تصادفی x در صورتیکه Y=y باشد، چنین است:

$$f_{x|Y}(x|y)=\frac{f(x,y)}{f_Y(y)} \quad ;f_Y(y)>0$$

$$f_{Y|x}(y|x)=\frac{f(x,y)}{f_X(x)} \quad ;f_X(x)>0$$

مثال: توزیع احتمال مشترک دو متغیر تصادفی ناپیوسته y و x به شرح جدول زیر است:

x \ y	-3	2	4
1	0.1	0.2	0.2
3	0.3	0.1	0.1

به نظر شما $P(x=3|y=-3)$ برابر است با؟

۱) $\frac{3}{10}$ ۲) $\frac{2}{5}$ ۳) $\frac{3}{5}$ ۴) $\frac{3}{4}$

$$P(x=3|y=-3)=\frac{P(x=3\ ,\ y=-3)}{P(y=-3)}=\frac{0.3}{0.4}=\frac{3}{4}$$

x \ y	-3	2	4	P(x)
1	0.1	0.2	0.2	0.5
3	0.3	0.1	0.1	0.5
P(y)	0.4	0.3	0.3	1

گزینه ۴ صحیح می‌باشد.

مثال: اگر در فواصل $0 \le y \le 1$ و $0 \le x \le 2$ تابع توزیع مشترک این دو متغیر $f(x,y)=\frac{1}{3}(x+y)$ باشد، $f(y|x)$ در همین فواصل برابر است با:

۱) $\frac{2+3y}{2}$ ۲) $\frac{1}{3}(x+y^2)$ ۳) $\frac{x+y}{x+\frac{1}{2}}$ ۴) $\frac{x+2y}{3}$

$$f(y|x)=\frac{f(x,y)}{f(x)}$$

$$f(x)=\varphi(x)=\int_0^1 \frac{x+y}{3}dy=\frac{xy+\frac{y^2}{2}}{3}\Big]_0^1=\frac{2x+1}{6}$$

$$f(y|x)=\frac{\frac{x+y}{3}}{\frac{2x+1}{6}}=\frac{2x+2y}{2x+1}=\frac{2(x+y)}{2\left(x+\frac{1}{2}\right)}=\frac{x+y}{x+\frac{1}{2}}$$

گزینه ۳ صحیح می‌باشد.

مثال: با توجه به جدول احتمال توام زیر، احتمال A به شرط B چقدر است؟

	B	$\bar{B}$
A	?	0.20
$\bar{A}$	0.10	0.50

۱) 0.040 ۲) 0.20 ۳) 0.333 ۴) 0.667

مجموع احتمال داخل جدول برابر 1 است.

$$P(A\cap B)=1-(0.2+0.1+0.5)=0.2$$

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{0.2}{0.3}=0.667$$

گزینه ۴ صحیح می‌باشد.

مثال: تابع چگالی احتمال مشترک متغیرهای تصادفی y, x به صورت زیر تعریف شده است:

$$P_{\xi,\eta}(x,y)=\begin{cases}\frac{1}{3}(x+y) & 0\le x\le 1\ ,\ 0\le y\le 2\\ 0 & \text{در غیر اینصورت}\end{cases}$$

چگالی احتمال مشروط $P_{x,y}(x|y)$ کدام است؟

۱) $\frac{x+y}{2y+1}$ ۲) $\frac{x+y}{2(2y+1)}$ ۳) $\frac{2(x+y)}{2y+1}$ ۴) $\frac{x+y}{3(2y+1)}$

گزینه ۳ صحیح می‌باشد.

$$P(x|y)=f(x|y)=\frac{f(x,y)}{f(y)}=\frac{\frac{1}{3}(x+y)}{\int_0^1\frac{1}{3}(x+y)dx}=\frac{x+y}{\frac{1}{2}+y}=\boxed{\frac{2(x+y)}{2y+1}}$$

☑ امید شرطی

تعریف ـ امید شرطی y به شرط X=x طبق رابطه زیر تعریف میشود:

$$E\left(y \mid X=x\right)=\sum_{\forall y} y\frac{f(x,y)}{f_x(x)}$$ x,y گسسته

$$E\left(y \mid X=x\right)=\int_{-\infty}^{+\infty} y\frac{f(x,y)}{f_x(x)}dy$$ x,y پیوسته

و نیز همین‌طور امید شرطی x به شرط Y=y :

$$E\left(x \mid Y=y\right)=\sum_{\forall x} x\frac{f(x,y)}{f_y(y)}$$ x,y گسسته

$$E(X \mid Y=y)\int_{-\infty}^{+\infty} x\frac{f(x,y)}{f_y(y)}dy$$ x,y پیوسته

مثال: توزیع احتمال مشترک دو متغیر تصادفی X و Y توسط جدول زیر بیان شده است. میانگین شرطی Y بر حسب X=6 کدام است؟

X \ Y	2	3	4	5	f(x)
4	0.10	0.05	0.03	0.02	0.20
6	0.05	0.20	0.20	0.05	0.50
8	0.05	0.10	0.10	0.05	0.30
f(y)	0.020	0.35	0.33	0.12	1

۱) 4.2 ۲) 3.8 ۳) 3.5 ۴) 3

$$E\left(Y \mid X=6\right)=\frac{\sum y_i \times f\left(x=6,y_i\right)}{f\left(x=6\right)}=\frac{(2\times 0.05)+(3\times 0.20)+(4\times 0.20)+(5\times 0.05)}{0.05+0.20+0.20+0.05}=\frac{1.75}{0.5}=3.5$$

گزینه ۳ صحیح می‌باشد.

مثال: توزیع احتمال مشترک دو متغیره X و Y به صورت جدول زیر است:

x \ y	0	1	2
0	0.20	0.15	0.05
1	0.05	0.20	0.05
2	0.05	0.05	0.20

$E\left(Y \mid X=1\right)$ برابر است با:

۱) 2 ۲) 1 ۳) 0.5 ۴) 1.5

$$E\left(Y \mid X=1\right)=\frac{\sum y_i f\left(x=1\ ,\ y_i\right)}{f\left(x=1\right)}=\frac{0\times P\left(x=1,y=0\right)+1\times P\left(x=1,y=1\right)+2\times P\left(x=1,y=2\right)}{0.05+0.2+0.05}$$

$$=\frac{0\times 0.05+1\times 0.20+2\times 0.05}{0.3}=\frac{0.3}{0.3}=1$$

X \ Y	0	1	2	f(x)
0	0.20	0.15	0.05	0.4
1	0.05	0.20	0.05	0.3
2	0.05	0.05	0.20	0.3
f(y)	0.3	0.4	0.3	1

گزینه ۲ صحیح می‌باشد

نکته مهم: اگر x, y مستقل از هم باشند، آنگاه:

$$\begin{cases} f[x \mid y] = f[x] \\ f[y \mid x] = f[y] \end{cases} \qquad \begin{cases} E[x \mid y] = E[x] \\ E[y \mid x] = E[y] \end{cases} \Leftrightarrow$$ x و y مستقل اند.

$E(xy) = E(x).E(y) \Leftrightarrow$ x و y مستقل اند.

$$E[xy] = \begin{cases} \sum\sum x_i y_k f(x_i, y_k) & \text{در حالت گسسته} \\ \iint xy f(x,y)dxdy & \text{در حالت پیوسته} \end{cases}$$

☑**امید ریاضی تابعی از دو متغیر تصادفی x و y به صورت k(x,y) که دارای تابع توزیع توام $F_{x,y}(x,y)$ است به فرم زیر نشان داده میشود:**

$$E[k(x,y)] = \sum_x \sum_y k(x,y) f(x,y) \quad \text{گسسته}$$

$$E[k(x,y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k(x,y) f(x,y) dx\, dy \quad \text{پیوسته}$$

کوواریانسی (هم پراش) : $(cov(x,y) = \delta_{xy} = sp_{xy})$

$$CoV(x,y) = \delta_{xy} = E[(x-\mu_x)(y-\mu_y)]$$
$$= E[xy] - E[x]E[y] = E[xy] - \mu_x \mu_y$$

*** نکته مهم:** در مورد میانگین همیشه دو قاعده مهم زیر وجود دارد:

۱) **مجموع انحرافات از میانگین همیشه صفر است و همین‌طور میانگین آن یعنی** $E(x-E(x)) = 0$

۲) **امید مجذور انحرافات از میانگین که به آن واریانس گفته می‌شود همیشه مینیمم است** $E\left[(x-E(x))^2\right] \le E\left[(x-a)^2\right]$

$-\infty <$ **کوواریانس = هم پراش** $= Cov(x,y) < +\infty$

کوواریانس یا هم پراشی معیاری است که ارتباط بین x و y را نشان می‌دهد.

اگر $\delta_{xy} = Cov(x,y) > 0$: بین x و y رابطه مستقیم وجود دارد، یعنی با افزایش x , y افزایش پیدا می کند.

اگر $\delta_{xy} = Cov(x,y) < 0$: بین x و y رابطه معکوس وجود دارد.

اگر $\delta_{xy} = Cov(x,y) = 0$: x و y نا همبسته‌اند (به عبارتی یا مستقل اند یا ارتباط خطی با یکدیگر ندارند)

نکته مهم:

اگر y و x مستقل باشند. $Cov(x,y) = \delta_{xy} = 0$ عکس مطلب صحیح نیست یعنی ممکن $Cov(x,y) = 0$ باشد و x و y مستقل نباشد.

نکته مهم : در صورتی که x,y مستقل باشند حتماً $E(xy) = E(x).E(y)$ است و در نتیجه $Cov(x,y) = 0$ اما ممکن است $E(xy) = E(x).E(y)$ یا $Cov(x,y) = 0$ باشد ولی x و y مستقل نباشند.

خواص کواریانس : اگر a,b,c,a',b',c' ثابت باشد.

1) $Cov(x,y) = Cov(y,x)$

2) $Cov(x,x) = \delta_x^2 = Var(x)$

3) $Cov(x,a) = Cov(a,x) = 0$

4) $Cov(x_1 + x_2 + \ldots\ldots, y) = Cov(x_1, y) + \ldots + Cov(x_k, y) + \ldots\ldots$

5) $Cov(ax+b, cy+d) = ac\,Cov(x,y)$

6) $Cov(ax+by+c, a'x+b'y+c') = aa'\ \delta_x^2 + bb'\delta_y^2 + ab'Cov(x,y) + a'b\,Cov(x,y)$

مهم:

$$E[ax \pm by \pm cz \pm d] = aE[x] \pm bE[y] \pm cE[z] \pm [d]$$

$$var(ax \pm by \pm cz + d) = a^2 var(x) + b^2 var(y) + c^2 var(z) \pm 2ab\,cov(x,y) \pm 2ac\,cov(x,z) \pm 2bc\,cov(y,z)$$

$$Var(ax \pm by + c) = a^2\ Var(x) + b^2\ Var(y) \pm 2ab\ Cov(x,y)$$

اگر متغیرهای $x_1, x_2, \ldots, x_n$ مستقل باشند آنگاه

$$var(\pm a_1x_1 \pm a_2x_2 \pm \ldots \pm a_nx_n) = a_1^2\ var(x_1) + a_2^2\ var(x_2) + \ldots + a_n^2\ var(x_n)$$

نکته: $-\infty <$ **هم پراش** $=$ **کوواریانس** $= Cov(x,y) < +\infty$

$Cov(x,y) = 0$ ناهمبسته
x, y از هم مستقل بوده یا اصلاً رابطه خطی با هم نداشته بلکه رابطه غیرخطی دارند.

$Cov(x,y) < 0$
y, x رابطه معکوس و منفی دارند

$Cov(x,y) > 0$
x,y رابطه مستقیم و مثبت دارند.

نکته: برای محاسبه $cov(x,y)$ از رابطه زیر استفاده می‌کنیم:

$$cov(x,y) = E\big[(x - E[x])(y - E[y])\big]$$

$$= E\big[xy - E[x]E[y]\big]$$

نکته مهم: قبل از محاسبه کواریانس بهتر است ابتدا مستقل بودن x,y را چک کنیم چون اگر مستقل باشند $Cov(x,y)=0$ است.

y, x مستقل باشند $\Rightarrow \underbrace{E[x]E[y] = E[xy] \Leftrightarrow cov(x,y) = 0}_{}$

x, y ناهمبسته‌اند

مثال: در جدول ۱ x,y مستقل‌اند در نتیجه $cov(x,y)=0$ است

جدول ۱

Y \ x	0	1	f (y)
0	0.15	0.35	0.5
0	0.15	0.35	0.5
f (x)	0.3	0.7	1

در جدول ۲ x,y مستقل نیستند ولی باز هم $cov(x,y)=0$ شده است.

جدول ۲

Y \ x	-1	0	1	f(y)
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
f(x)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

$$E(x)=-1\times\frac{1}{3}+0\times\frac{1}{3}+1\times\frac{1}{3}=0$$
$$E(y)=0\times\frac{1}{3}+1\times\frac{2}{3}=\frac{2}{3}$$
$$E(xy)=-1\times0\times0+0\times0\times\frac{1}{3}+1\times0\times0+-1\times1\times\frac{1}{3}+0\times1\times0+1\times1\times\frac{1}{3}=0$$

$$\operatorname{cov}(x,y)=E[xy]-E[x]E[y]=0$$

در عین حال دیده می‌شود که x,y مستقل نیستند اما

$$\operatorname{cov}(x,y)=0$$

در نتیجه x,y ناهمبسته‌اند.

مثال: مقدار کوواریانس تابع احتمال توام زیر چقدر است؟

X \ Y	0	1
-1	0.30	0.30
0	0.30	0.10

۱) صفر ۲) 1 ۳) -0.30 ۴) -0.06

ابتدا مستقل بودن را کنترل می‌کنیم اما در اینجا X و Y مستقل نیستند چون در صورت مستقل بودن کوواریانس 0 خواهد بود.

X \ Y	0	1	f(x)
-1	0.30	0.30	0.60
0	0.30	0.10	0.40
f(y)	0.6	0.4	1

$$E(X)=\sum x.f(x)=-0.6+0=-0.6$$
$$E(Y)=\sum y.f(y)=0+0.4=0.4$$
$$E(XY)=\sum\sum xyf(x,y)=(0\times0.3)+(-1\times0.3)+(0\times0.3)+(0\times0.3)=-0.3$$
$$\operatorname{Cov}(X,Y)=E(XY)-E(X)E(Y)=-0.3-(0.6\times0.4)=-0.06$$

گزینه ۴ صحیح می‌باشد.

مثال: توزیع احتمال مشترک دو متغیر تصادفی X ، Y به صورت زیر است:

X \ Y	1	2
5	0.1	0.3
10	0.4	0.2

کوواریانس این دو متغیر برابر است با؟

۱) 1.5 ۲) 1 ۳) 0.5 ۴) -0.5

X \ Y	1	2	f(x)
5	0.1	0.3	0.4
10	0.4	0.2	0.6
f(y)	0.5	0.5	1

$$E(X)=\sum x_i f(x_i)=5\times0.4+10\times0.6=8$$
$$E(Y)=\sum y_i.f(y_i)=1\times0.5+2\times0.5=1.5$$
$$E(XY)=\sum\sum x_i y_i f(x_i,y_i)$$

$=(5\times1\times0.1)+(5\times2\times0.3)+(10\times1\times0.4)+(10\times2\times0.2)=0.5+3+4+4=11.5$

$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=11.5-(8\times1.5)=-0.5$

گزینه ۴ صحیح می‌باشد.

مثال: اگر متغیرهای تصادفی X و Y و Z دارای کمیت‌های $\mu_X=2,\mu_Y=-3,\mu_Z=4$ و واریانس $\sigma_X^2=1$ ، $\sigma_Y^2=2,\sigma_Z^2=2$ و کوواریانس‌های $\mathrm{Cov}(X,Y)=1,\mathrm{Cov}(X,Z)=1,\mathrm{Cov}(Y,Z)=1$ باشند، میانگین و واریانس $W=3X-Y+2Z$ کدام است؟

۱) 14 و 15 ۲) 21 و 17 ۳) 18 و 17 ۴) 18 و 18

$E(W)=E(3X-Y+2Z)=3E(X)-E(Y)+2E(Z)$

$E(W)=3(2)-(-3)+2(4)=6+3+8=17$

$D(W)=D(3X-Y+2Z)=9D(X)+D(Y)+4D(Z)-6\mathrm{Cov}(X,Y)$

$+12\,\mathrm{Cov}(X,Z)-4\mathrm{Cov}(Y,Z)$

$D(W)=9+2+8-6+12-4=21$

گزینه ۲ صحیح می‌باشد

مثال: در صورت صادق بودن کدام یک از شرایط زیر دو متغیر تصادفی X و Y مستقل هستند؟

۱) $E(XY)=E(X)E(Y)$ ۲) $\mathrm{Cov}(X,Y)=0$

۳) $P(X,Y)$ ۴) $f(X,Y)=f(X)f(Y)$

تنها رابطه‌های دو طرفه با استقلال به شرح زیر است:

$$x,y \text{ مستقل} \Leftrightarrow f\begin{cases} f(x,y)=f(x)\times f(y) \\ f(x|y)=f(x)\,,f(y|x)=f(y) \\ E(x|y)=E(x)\,,E(y|x)=E(y) \end{cases}$$

گزینه ۴ صحیح می‌باشد.

مثال: اگر کوواریانس بین X و Y مساوی با صفر باشد، کدام یک از عبارات زیر صحیح است؟

۱) آن دو متغیر مستقل‌اند. ۲) دارای رابطه غیر خطی می‌باشند.

۳) موارد ۱ و ۲ هر دو صحیح است. ۴) یا دارای رابطه غیر خطی هستند یا هیچ رابطه‌ای با همدیگر ندارند ✓

گزینه ۴ صحیح می‌باشد.

مثال: اگر $V(X)=9$ ، $V(Y)=2$ و $\mathrm{Cov}(X,Y)=-3$ باشد، واریانس $Z=-\frac{1}{3}X-2Y+18$ برابر است با؟

۱) 5 ۲) 77 ۳) 85 ۴) 87

گزینه ۱ صحیح می‌باشد.

مثال: کدام یک از موارد زیر برای V(X-Y) صحیح است؟ (اقتصاد ـ ۷۸)

۱) $V(X)+V(Y)-2Cov(X,Y)$
۲) $V(X)-V(Y)-Cov(X,Y)$
۳) $V(X)+V(Y)+2Cov(X,Y)$
۴) $V(X)+V(Y)-Cov(X,Y)$

گزینه ۱ صحیح می‌باشد.

مثال: چنانچه $W = a + bx$ و $Z = c + dy$ باشد، $Cov(Z , W)$ بر حسب X و Y عبارت است از:

۱) $(a+b)(c+d)Cov(x,y)$
۲) $bd\,Cov(x,y)$
۳) $(ac+bd)Cov(x,y)$
۴) $Cov(x,y)$

بنابر قاعده پخشی:

$$Cov(Z;W)=Cov\left(\cancel{\overset{0}{c,a}}\right)+Cov\left(\cancel{\overset{0}{c,bx}}\right)+Cov\left(\cancel{\overset{0}{dy,a}}\right)+Cov(dy,bx)=bd\,Cov(x,y)$$

گزینه ۲ صحیح می‌باشد.

مثال: اگر X , Y دو متغیر تصادفی مستقل باشند و $E(Y)=2, E(X)=3$ باشند، کدام یک از عبارت‌های زیر صحیح است؟

۱) $Cov(X,Y)=0$
۲) $E(X+Y)=5$
۳) $E(XY)=6$
۴) هر سه صحیح است

(اگر x و y مستقل نباشند فقط گزینه ۲ صحیح است)

$$E(X+Y)=E(X)+E(Y)=3+2=5$$
$$E(XY)=E(X)E(Y)=3\times2=6$$
$$Cov(X,Y)=0$$

گزینه ۴ صحیح می‌باشد.

مثال: اگر $\sigma_x^2=\frac{1}{2}$ و $\sigma_x^2=\frac{2}{3}$ و $\sigma_{x+y}^2=\frac{5}{6}$ باشد، مقدار کوواریانس کدام است؟

۱) $-\frac{1}{2}$
۲) $-\frac{1}{6}$
۳) $\frac{1}{6}$
۴) $\frac{1}{3}$

$$\sigma^2_{(X+Y)}=\sigma_x^2+\sigma_y^2+2\,Cov(X,Y) \Rightarrow \frac{5}{6}=\frac{2}{3}+\frac{1}{2}+2Cov(X,Y) \Rightarrow 2Cov(X,Y)=-\frac{2}{6} \Rightarrow Cov(X,Y)=-\frac{1}{6}$$

گزینه ۲ صحیح می‌باشد.

مثال: اگر X , Y دو متغیر تصادفی مستقل باشند و $E(Y)=2, E(X)=1$ باشند، کدام یک از عبارت‌های زیر صحیح است؟

۱) $Cov(X,Y)=0$
۲) $E(X+Y)=3$
۳) $E(XY)=2$
۴) هر سه صحیح است

(اگر x و y مستقل نباشند **فقط** گزینه ۲ صحیح است)

$E(X+Y)=E(X)+E(Y)=1+2=3$

$E(XY)=E(X)E(Y)=1\times 2=2$

$Cov(X,Y)=0$

گزینه ۴ صحیح می‌باشد.

ضریب همبستگی : معیاری است که شدت یا ضعف ارتباط بین x و y را بیان می‌کند. و به صورت زیر نشان داده می‌شود.

$$\rho_{x,y}=\rho=\frac{Cov(x,y)}{\sqrt{\delta_x^2.\delta_y^2}}=\frac{\delta_{xy}}{\delta_x.\delta_y}$$

که همواره $-1\leq\rho\leq+1$ است .

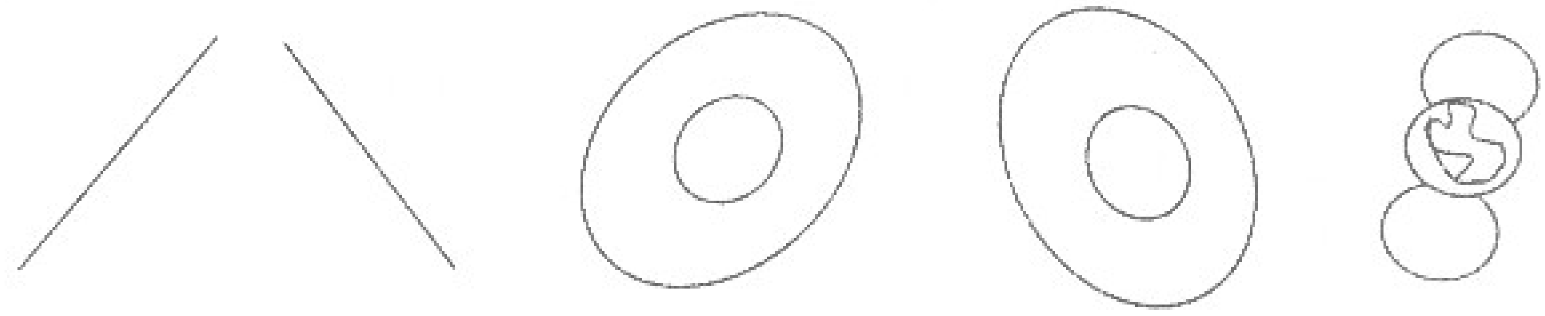
اگر $\rho=1$ باشد . بین x و y رابطه مستقیم و کامل وجود دارد.

اگر $\rho=-1$ باشد . بین x و y رابطه معکوس و کامل وجود دارد.

اگر $0<\rho<+1$ باشد. بین x و y رابطه مستقیم و ناقص وجود دارد.

اگر $-1<\rho<0$ باشد . بین x و y رابطه معکوس و ناقص وجود دارد.

اگر x و y مستقل از یکدیگر باشند. آنگاه $\rho=0$ است ولی عکس این مطلب همیشه درست نیست . یعنی اگر $\rho=0$ باشد . ممکن است x و y مستقل باشند یا نباشند. و تنها می‌توان گفت بین x و y رابطه خطی وجود ندارد.



نکته مهم: هر تغییری روی متغیر x و y تغییری در ρ (ضریب همبستگی) بوجود نمی‌آورد.

$\rho_{ax\pm b,cy\pm d}=\rho_{x,y}$

$\rho_{x,x}=1$

$\rho_{x,-x}=-1$

توضیح: $\rho_{ax\pm b,cy\pm d}=\dfrac{Cov(ax\pm b , cy\pm d)}{\sqrt{Var(ax\pm b)Var(Cy\pm d)}}=\dfrac{Cov(x,y)}{\delta_x\ \delta_y}=\rho_{x,y}$

$-1\leq\rho_{x,y}\leq 1$

خواص ضریب همبستگی:

$(1)\ P_{x,x}=1$

$(2)\ P_{x,a}=0$

$(3)\ P_{x,y}=P_{y,x}$

نکته:

$$x,y \text{ مستقل} \underset{\not\leftarrow}{\longrightarrow} \begin{cases} - \quad\quad\quad x,y \\ -Cov(x,y)=0 \\ -P_{x,y} \end{cases}$$

مثال: به شرط صفر بودن کوواریانس دو متغیر تصادفی X و Y:

۱) ضریب همبستگی این دو متغیر صفر است

۲) ضریب همبستگی این دو متغیر مثبت است

۳) ضریب همبستگی این دو متغیر منفی است

۴) این دو متغیر مستقل‌اند

$$r_{x,y}=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=\frac{0}{\sigma_X\sigma_Y}=0$$

گزینه ۱ صحیح می‌باشد

مثال: فرض کنیم تابع چگالی احتمال‌های مشترک دو متغیر X و Y نرمال باشد، در این صورت اگر کوواریانس X و Y صفر باشد؟

۱) ضریب همبستگی X و Y مثبت است

۲) X و Y می‌توانند مستقل نباشند.

۳) X و Y مستقل از هم هستند.

۴) ضریب همبستگی X و Y منفی است.

$$x,y \text{ مستقل} \underset{\not\leftarrow}{\longrightarrow} \begin{cases} - \quad\quad\quad x,y & \text{ناهمبسته‌اند} \\ -Cov(x,y)=0 & \\ -P_{x,y} & \text{ناهمبسته‌اند} \end{cases}$$

گزینه ۲ صحیح می‌باشد.

مثال: اگر $Cov(x,y)=10$ و $\sigma_x=5$ و $\sigma_y=3$ باشد، ضریب همبستگی کدام است؟

۱) $-\frac{1}{2}$ ۲) $\frac{1}{2}$ ۳) $\frac{2}{3}$ ۴) هیچکدام

$$\hat{\rho}=\frac{Cov(x,Y)}{\sigma_X.\sigma_Y}=\frac{10}{5\times3}=\frac{2}{3}$$

گزینه ۳ صحیح می‌باشد.

- ضریب تشخیص یا ضریب تعیین

ضریب تشخیص بیان کننده نسبت درصد تغییرات تابع یعنی y بوسیله تغییرات متغیر یعنی x می باشد. به عبارت دیگر ضریب تشخیص معلوم می‌کند که چند درصد از تغییرات y ناشی از تغییرات x است. ضرایب تشخیص R^2 با معلوم بودن ضریب **همبستگی** یعنی r از فرمول زیر بدست می آید:

$0\leq R^2\leq1 \qquad\qquad R^2=r^2$

مثال: اگر $Cov(X,Y)=32.4$ و $\sigma_x^2=\sigma_y^2=36$ باشد، ضریب تعیین (تشخیص) چقدر است؟

۱) 0.64 ۲) 0.72 ۳) 0.81 ۴) 0.90

$$r=\frac{Cov(X,Y)}{\sigma_X.\sigma_Y}=\frac{32.4}{6\times6}=0.9$$

$$r^2=(0.9)^2=0.81$$

گزینه ۳ صحیح می‌باشد.

مثال: اگر کوواریانس دو صفت متغیر X و Y برابر با $\sigma_{xy}=Cov(X,Y)=-27$ باشد و واریانس هر یک به ترتیب $\sigma_x^2=36$ و $\sigma_Y^2=25$ به دست آمده باشد، چند درصد تغییرات Y به وسیله X بیان شده است؟

۱) %76 ۲) %81 ۳) %95 ۴) %78

گزینه ۲ صحیح می‌باشد. بنابراین $(100\times0.81=\%81)$ تغییرات Y به وسیله X بیان شده است.

$$\hat{\rho}=\frac{Cov(x,Y)}{\sigma_X.\sigma_Y}=\frac{-27}{6\times5}=-0.9\Rightarrow r^2=(-0.9)^2=0.81$$

مثال: ضریب همبستگی بین X و Y مساوی با %70 است. چند درصد تغییرات Y تحت تأثیر X نیست؟

۱) %30 ۲) %49 ۳) %51 ۴) %70

$$r=\%70=0.7\Rightarrow r^2=(0.7)^2=0.49$$

بنابراین %51 تغییرات Y تحت تأثیر X نیست.

گزینه ۳ صحیح می‌باشد.

مثال: اگر ضریب همبستگی بین دو متغیر 0.6 و دو متغیر دیگر 0.3 باشد، می‌توان گفت همبستگی دو متغیر اول «چند برابر قوی‌تر» از دو متغیر دوم است؟

۱) دو برابر ۲) سه برابر ۳) چهار برابر ۴) نه برابر

قوت ارتباط توسط ضریب تشخیص مشخص می‌شود بنابراین $\frac{R^2_{x,y}}{R^2_{z,t}}=\frac{0.6^2}{0.3^2}=4$

گزینه ۳ صحیح می‌باشد.

همبستگی و رگرسیون

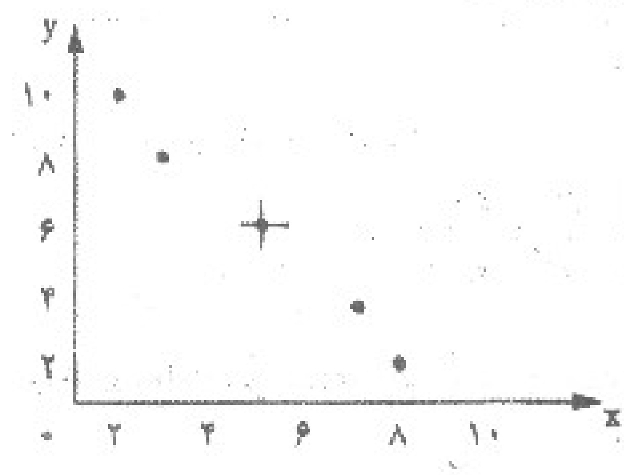
از یک جامعه یا نمونه ای از یک جامعه دو صفت متغیر x و y را در نظر می گیریم و مقادیر مختلف را اندازه گیری می کنیم. مثلاً در یک کلاس نمرات ریاضی و معدل دانشجویان را مورد بررسی قرار می دهیم. منظور از این مبحث این است که آیا بین دو صفت متغیر x و y همبستگی و رابطه خطی وجود دارد یا خیر و روش آماری آن را مورد توجه قرار دهیم.

۱- دیاگرام یا نمودار پراکندگی

اگر هر زوج مرتب (x_i,y_i) را به عنوان مختصات یک نقطه در نظر بگیریم و آنها را در صفحه مختصات رسم کنیم نمودار پراکندگی بدست خواهد آمد که با مطالعه روی آن و از نحوه قرار گرفتن نقاط می توان حدس زد که آیا رابطه و بستگی بین x و y وجود دارد یا خیر؟

مثال: در مثال فوق نمودار پراکندگی بشرح زیر است.

x	Y
2	10
3	8
5	6
7	4
8	2
Σ:25	30



$$\bar{x} = \frac{\sum x_i}{n} = \frac{25}{5} = 5$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{30}{5} = 6$$

ضمناً نقطه میانگین یعنی(۶و۵) را روی نمودار نشان داده ایم.

این نمودار نشان می دهد که احتمالاً x و y بستگی خطی دارند.

۲- ضریب همبستگی

بهترین معیار تشخیص وجود همبستگی یا عدم آن و حتی نوع و جهت و میزان همبستگی می باشد.اگر ضریب همبستگی را با r نشان دهیم فرمولهای ضریب همبستگی ساده بشرح زیر می باشند.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \qquad (a)$$

اگر صورت و مخرج این کسر را بر n تقسیم کنیم خواهیم داشت:

$$r = \frac{\frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{\sum (x_i - \bar{x})^2}{n} \cdot \frac{\sum (y_i - \bar{y})^2}{n}}}$$

با توجه به تعریف واریانس های x و y و کوواریانس x و y داریم:

$$r = \frac{E(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x . \sigma_y} = \frac{Cov(x,y)}{\sigma_x . \sigma_y} \qquad (b)$$

از بسط فرمول (a) داریم:

$$r = \frac{\sum x_i y_i - \frac{1}{n}\left(\sum x_i\right)\left(\sum y_i\right)}{\sqrt{\left[\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2\right]\left[\sum y_i^2 - \frac{1}{n}\left(\sum y_i\right)^2\right]}} =$$

$$= \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{\sqrt{\left[n\sum x_i^2 - \left(\sum x_i\right)^2\right]\left[n\sum y_i^2 - \left(\sum y_i\right)^2\right]}} \qquad (c)$$

و اگر در فرمول (c) از $\sum x_i = n\bar{x}$ و $\sum y_i = n\bar{y}$ استفاده کنیم خواهیم داشت:

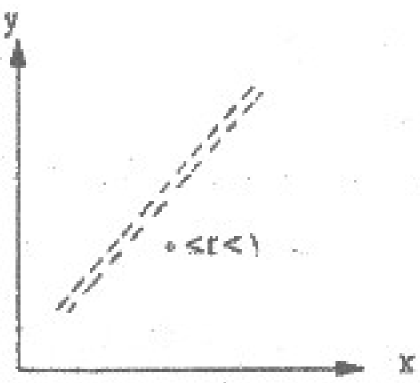
$$r = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(\sum x_i^2 - n\bar{x}^2\right)\left(\sum y_i^2 - n\bar{y}^2\right)}}$$

تبصره ۱: ضریب همبستگی همواره بین ۱- و ۱+ قرار دارد یعنی $-1 \le r \le 1$

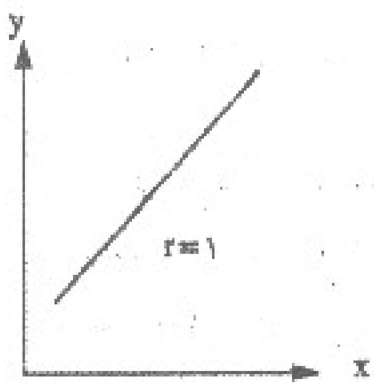
تبصره ۲: اگر $\bar{x}$ و $\bar{y}$ اعداد صحیح باشند استفاده از فرمول (a) مناسب است و اگر $\bar{x}$ و $\bar{y}$ اعداد صحیح نباشند استفاده از فرمول (c) مناسب است.

۳- انواع همبستگی‌ها

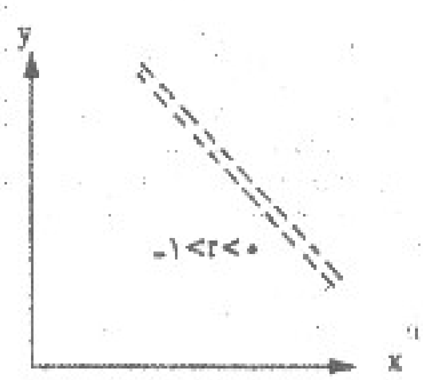
بین نمودار پراکندگی و ضریب همبستگی r که از فرمول های فوق بدست می آید رابطه بشرح زیر وجود دارد:



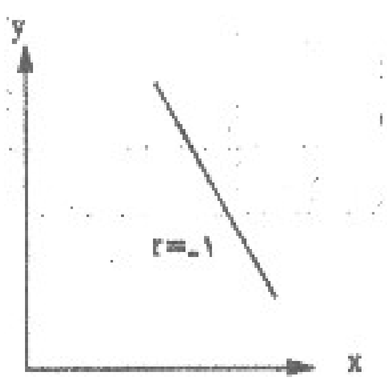
همبستگی مستقیم و ناقص است



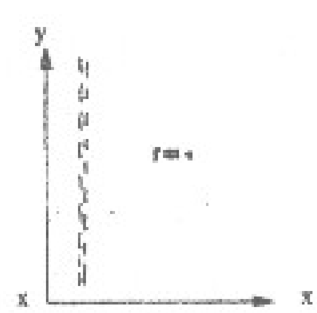
همبستگی مستقیم و کامل است



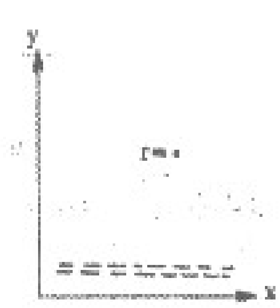
همبستگی معکوس و ناقص است



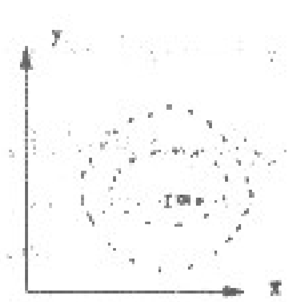
همبستگی معکوس و کامل است



عدم همبستگی



عدم همبستگی

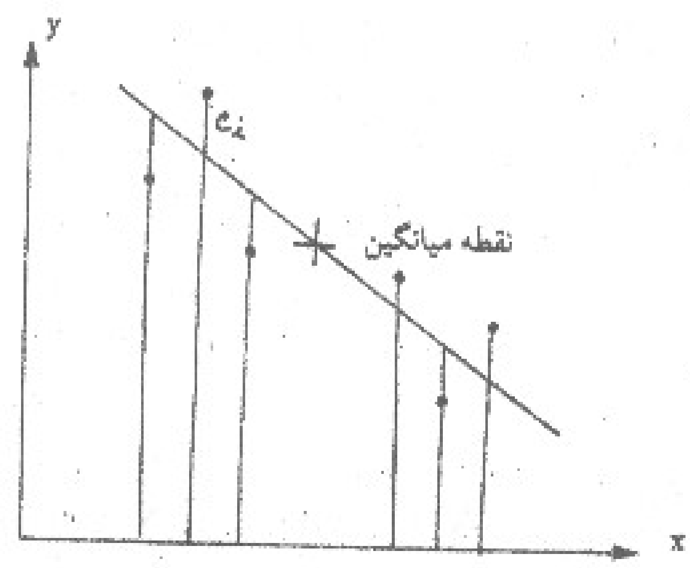


عدم همبستگی

معادله خط رگرسیون

پس از تحقیق معنی دار بودن r می توان معادله همبستگی خطی را نوشت و خط رگرسیون را رسم نمود:

معادله خط رگرسیون بطور کلی $y = a + bx$ می باشد.



خط رگرسیون به خطی می گویند که از لابلای نقاط نمودار پراکندگی می گذردو مجموع مربعات فواصل نقاط از این خط در امتداد محور yها حداقل می باشد.

تبصره – معادله خط رگرسیون همواره از نقطه میانگین یعنی $(\overline{x}, \overline{y})$ می گذرد و شیب آن b می باشد.

با توجه به توضیحات فوق ضرایب معادله خط رگرسیون بشرح زیر بدست می آیند.م

$$b = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2}$$

و چون خط رگرسیون از نقطه میانگین می گذرد یعنی $\overline{y} = a + b\overline{x}$ در نتیجه خواهیم داشت:

$$a = \overline{y} - b\overline{x}$$

با معلوم بودن $\overline{x}$ و $\overline{y}$ و b مقدار ثابت a محاسبه می گردد.

تبصره ۱: ضرایب a و b را می توان به کمک واریانس و کوواریانس بشرح زیر بدست آورد.

اگر صورت و مخرج کسر b فوق را بر n تقسیم کنیم خواهیم داشت:

$$a = \overline{y} - b\overline{x} \quad \text{یا} \quad b = \frac{\mathrm{Cov}(x,y)}{\sigma_x^2} = \frac{Sp_{xy}}{SS_x} = \frac{\sigma_{xy}}{S_x^2}$$

تبصره ۲: ضریب b را می توان با بسط صورت و مخرج کسر b بشرح زیر هم محاسبه نمود.

$$b = \frac{\sum x_i y_i - n\overline{x}\overline{y}}{\sum x_i^2 - n\overline{x}^2} \quad \text{یا} \quad b = \frac{\sum x_i y_i - \frac{1}{n}\left(\sum x_i\right)\left(\sum y_i\right)}{\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2}$$

مثال: ضریب همبستگی جامعه دو متغیر وابسته X و Y برابر است با:

۱) 2   ۲) 1   ۳) $\frac{1}{2}$ ۴) -1

X_i	Y_i	X_i^2	Y_i^2	X_iY_i
7	14	49	196	98
10	20	100	400	200
4	8	16	64	32
11	22	121	484	242
32	64	286	1144	572

$$r=\frac{n\sum XY-\sum X\sum Y}{\sqrt{\left[n\sum X^2-\left(\sum X\right)^2\right]\left[n\sum Y^2-\left(\sum Y\right)^2\right]}}$$

$$r=\frac{4(572)-(32)(64)}{\sqrt{\left[4(286)-(32)^2\right]\left[2(1144)-(64)^2\right]}}$$

$$r=\frac{240}{\sqrt{120\times480}}=\frac{240}{240}=1$$

گزینه ۲ صحیح می‌باشد.

مثال: اگر شیب معادله رگرسیون 10 – باشد، $\sum X_i=100$ ، $\bar{X}=20$ ، $\sum Y_i=20$ باشد، ثابت معادله کدام است؟

۱) 106   ۲) 110

۳) 220   ۴) 204

گزینه ۴ صحیح می‌باشد.

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=\frac{20}{5}-(-10)(20)\Rightarrow\hat{\alpha}=204$$

$$\bar{X}=\frac{\sum X_i}{n}\Rightarrow 20=\frac{100}{n}\Rightarrow n=5$$

مثال: فرض کنید $Cov(X,Y)=12$ ، $n=10$ ، $\sum X_i=\sum Y_i=50$ ، $\sigma_X=4$ و $\sigma_Y=3$ است. معادله رگرسیون Y برحسب X کدام است؟

۱) $\bar{y}_x=1.5-0.3x$

۲) $\bar{y}_x=1.25+0.75x$

۳) $\bar{y}_x=1.5+0.4x$

۴) $\bar{y}_x=3+2.2x$

$$\rho=\frac{Cov(x,Y)}{\sigma_X.\sigma_Y}=\frac{12}{4\times3}=1\Rightarrow r=1$$

$$r=\hat{\beta}\frac{\sigma_X}{\sigma_Y}\Rightarrow 1=\hat{\beta}\frac{4}{3}\Rightarrow\hat{\beta}=\frac{3}{4}=0.75$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=\frac{\sum Y_i}{n}-\hat{\beta}.\frac{\sum X_i}{n}=\frac{50}{10}-\frac{3}{4}\left(\frac{50}{10}\right)=1.25$$

$$y_x=\hat{\alpha}+\hat{\beta}X\Rightarrow y_x=1.25+0.75x$$

گزینه ۲ صحیح می‌باشد.

مثال: فرض کنید:

$n=10\ ,\ \sum X_i=\sum Y_i=50\ ,\ \sigma_X=4\ ,\ \sigma_Y=3\ ,\ \mathrm{Cov}(X,Y)=12$

مقدار رگرسیون Y برحسب X کدام است؟

۱) $Y=1.25+0.75X$ ۲) $Y=1.5+0.3X$

۳) $Y=1.5+0.4X$ ۴) $Y=3+2.2X$

$$\hat{\beta}=\frac{\mathrm{Cov}(x,Y)}{\sigma_x^2}=\frac{12}{(4)^2}=\frac{3}{4}=0.75$$

$$\alpha=\bar{Y}-\hat{\beta}\bar{X}=\frac{\sum Y_i}{n}-\left(\frac{3}{4}\right)\left(\frac{\sum X_i}{n}\right)=5-\frac{3}{4}(5)=1.25$$

$$Y=1.25+0.75X$$

گزینه ۱ صحیح می‌باشد.

مثال: اگر $SP_{xy}=20$ و $SS_x=20$ و $SS_y=20$ و $\bar{x}=5$ و $\bar{y}=4$ باشد. معادله خط رگرسیون برابر است با:

۱) $y=-x+1$ ۲) $y=x-1$

۳) $y=x+1$ ۴) $y=\frac{1}{2}x+1$

$$\hat{\beta}=\frac{\sum(x_i-\bar{X})(y_i-\bar{Y})}{\sum(x_i-\bar{X})^2}=\frac{SP_{xy}}{SS_x}=\frac{20}{20}=1$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=4-(1\times 5)=-1$$

$$y=-1+x$$

گزینه ۲ صحیح می‌باشد.

مثال: با استفاده از اطلاعات زیر معادلهٔ رگرسیون کدام است؟

$\sum Y_i=50\ ,\ \sum X_i=75\ ,\ n=25\ ,\ \sum Y_i^2=228\ ,\ \sum X_iY_i=30\ ,\ \sum X_i^2=625$

۱) $\hat{Y}=2.9-0.3X$ ۲) $\hat{Y}=2.9-0.15X$

۳) $\hat{Y}=5.8-0.3X$ ۴) $\hat{Y}=8.7-0.6X$

$$\hat{\beta}=\frac{n\sum X_iY_i-\sum X_i\sum Y_i}{n\sum X_i^2-\left(\sum X\right)^2}=\frac{(25\times 30)-(75\times 50)}{(25\times 625)-(75)^2}=\frac{-3000}{10000}=-0.3$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=\frac{\sum Y_i}{n}-\hat{\beta}.\frac{\sum X_i}{n}=\frac{50}{25}-\left(-0.3\times\frac{75}{25}\right)=2.9$$

$$\hat{Y}=\hat{a}+\hat{\beta}X\Rightarrow \hat{Y}=2.9-0.3X$$

گزینه ۱ صحیح می‌باشد.

مثال: شیب خط رگرسیون $y = a + bx$ چقدر است؟

x	0	1	3	4	5
y	0	2	5	9	11

۱) 0.31 ۲) 1.21 ۳) 1.8 ۴) 2.2

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} = \frac{(5\times108)-(13\times27)}{(5\times51)-(13)^2} = \frac{189}{86} = 2.2$$

گزینه ۴ صحیح می‌باشد.

تبصره ۳- یکی از کاربردهای خط رگرسیون پیش بینی مقدار y بازاء هر x مفروض می باشد برای این منظور x مفروض را در معادله خط رگرسیون قرار می دهیم تا y بدست آید.

مثال: رابطه بین میانگین X و Y خطی است و داده‌های نمونه به شرح زیر در دست است:

$$\sum X_i Y_i = 1150,\ \bar{Y} = 10,\ n = 20,\ \sum X_i^2 = 550$$

مقدار پیش‌بینی شده به ازای $X = 6$ برابر است با:

۱) 10 ۲) 11 ۳) 12 ۴) 13

ابتدا معادله خط رگرسیون را بدست آورده سپس $x = 6$ را در آن قرار داده y را بدست می‌آوریم.

$$\hat{\beta} = \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2} = \frac{1150 - 20(5)(10)}{550 - 20(5)^2} = \frac{150}{50} = 3$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} = 10 - 3(5) = 10 - 15 = -5$$

$$Y = -5 + 3X \xrightarrow{X=6} Y = -5 + 18 = 13$$

گزینه ۴ صحیح می‌باشد.

مثال: معادله رگرسیون $\bar{y}_x = 400 - 20x$ را در نظر بگیرید. مقدار واقعی Y به‌ازای X=15 برابر با 150 می‌باشد. ضریب همبستگی کدام است؟

۱) $0 < r < 1$ ۲) $-1 < r < 0$ ۳) $r = +1$ ۴) $r = -1$

چون $\hat{\beta} = -20$ بنابراین همبستگی منفی است. از طرفی در معادله به ازای $X = 15$، عدد 150 به دست نمی‌آید $\left(\bar{y}_x = 400 - 300 = 100\right)$، پس همبستگی کامل نیست.

گزینه ۲ صحیح می‌باشد.

مثال: اگر $Cov(X,Y)=-10$ و $\sigma_X=\sigma_Y=2$ و $\bar{X}=\bar{Y}=10$ باشد، مقدار پیش‌بینی Y به‌ازای $X=4$، چقدر است؟

۱) 25 ۲) 128.75 ۳) 32.25 ۴) 40

$$\hat{\beta}=\frac{Cov(x,Y)}{\sigma_x^2}=\frac{-10}{(2)^2}=-2.5$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=10-(-2.5)(10)=35$$

$$\bar{y}_x=35-2.5\,x$$

اگر $X=4$ باشد، داریم:

$$Y=35-2.5(4)=25$$

گزینه ۱ صحیح می‌باشد.

توزیع‌های مهم گسسته :

۱ـ توزیع یکنواخت :

در این توزیع x_iها هم تراز فرض می‌شوند. بنابراین اگر $x_1,\dots\dots,x_n$ ، n حالت متمایز باشند،

$$f_x(x_i)=P(x=x_i)=\frac{1}{n}\quad ;i=1,2,\dots,n$$

$$\mu_x=\frac{1}{n}\sum x_i\quad ;\quad \delta_x^2=\frac{1}{n}\sum(x_i-\mu)^2$$

نکته : در حالت خاص برای متغیر تصادفی با توزیع دنباله‌ای از n عدد طبیعی که از ۱ شروع می‌شوند داریم:

$$\mu=\frac{n+1}{2}\quad ;\delta^2=\frac{n^2-1}{12}$$

و اگر توزیع x روی دنباله‌ای به فرم $a,a+d,a+2d,\dots\dots,(n-1)d+a$ انجام شود آنگاه :

$$\mu=a+\frac{(n-1)d}{2}$$

۲ـ توزیع برنولی (دو نقطه‌ای):

x : تعداد پیروزی یا شکست در ۱ بار آزمایش برد و باخت دارای توزیع برنولی است.

$$f(x)=p^xq^{1-x}\quad ;x=0,1\quad ;p+q=1$$

$$\mu=p\quad ;\quad \delta^2=pq$$

x	0	1
f(x)	q	p

مثال: اگر کمیت تصادفی X بر طبق قانون دو نقطه‌ای با تابع احتمال زیر توزیع شده باشد:

$$P_x(x)=p^x q^{1-x} \qquad x=0,1 \quad q=1-p$$

امید ریاضی آنان کدام است؟

۱) $\frac{pq}{n}$ ۲) $p(1-p)$ ۳) p ۴) np

گزینه ۳ صحیح می‌باشد. توزیع برنولی (دو نقطه ای است) بنابراین امید ریاضی p است.

$$E(X)=\Sigma x\,.\,f(x)=\sum_{x=0}^{1} x\,.\,p^x q^{1-x}=0+pq^{\circ}=p$$

۳ـ توزیع دو جمله‌ای (بینم): $(X\sim Bin(n,p))$

در n بار آزمایش مستقل برنولی تعداد پیروزی یا شکست از توزیع دو جمله‌ای تبعیت می‌کند. (p پیروزی + q شکست = 1)

$$f_x(x)=\binom{n}{x}P^x q^{n-x}\ ;x=0,1,2,\ldots\ldots,n$$

$$\mu=np\ ;\ \ \delta^2=npq$$

نکته مهم:

۱) در حالت خاص n=1 به توزیع برنولی می‌رسیم .

۲) در توزیع دو جمله‌ای باید همواره به 2 نکته توجه نمود:

۱) جای هیچ برد و باختی مشخص نیست

۲) جامعه نامحدود است

مثال: از کیسه‌ای که شامل 9 گلوله سفید و یک گلوله سیاه است n = 100 بار مکرراً یک گلوله را به طور تصادفی انتخاب کرده و پس از مشاهده رنگ گلوله انتخاب شده، مجدداً آن را به کیسه باز می‌گردانند امید ریاضی تعداد گلوله‌های سیاه در این 100 آزمایش کدام است؟

۱) 90 ۲) 1 ۳) 10 ۴) 9

گزینه ۳ صحیح می‌باشد. تعداد گلوله‌های سیاه بر طبق قانون توزیع دو جمله‌ای توزیع می‌گردند.

$$p\,(\text{سیاه})=\frac{1}{10}\ ،\ p\,(\text{سفید})=\frac{9}{10}\ ،\ n=100$$

$$E(X)=n\,.\,p=100\times\frac{1}{10}=10$$

مثال: متغیر تصادفی X بر طبق قانون دو جمله‌ای با پارامتر n = 50 و p = 0.2 توزیع می‌شود. امید ریاضی آن کدام است؟

۱) 0.2 ۲) 10 ۳) 8 ۴) 50

گزینه ۲ صحیح می‌باشد.

$$E(X)=n\,.\,p=50\times 0.2=10$$

مثال: متغیر تصادفی X بر طبق قانون دو جمله‌ای با پارامتر $n = 20$ و $p = 0.5$ توزیع می‌شود. واریانس متغیر تصادفی X کدام است؟

۱) 5 ۲) 10 ۳) 0.25 ۴) 15

گزینه ۱ صحیح می‌باشد.

مثال: احتمال این که در یک خط تولید، محصول تولید شده، غیراستاندارد باشد، 0.2 است، از این خط تولید $n = 100$ واحد محصول را به طور تصادفی انتخاب می‌کنیم. متغیر (کمیت) تصادفی X عبارت است از تعداد محصولات غیراستاندارد بین محصولات انتخاب شده، قانون توزیع احتمال‌های متغیر (کمیت) تصادفی X کدام است؟

۱) $P(x)=0.2\,(0.8)^{x-1}$ ۲) $P(x)=C_{100}^{x}\left(\frac{2}{10}\right)^{x}\left(\frac{8}{10}\right)^{100-x}$

۳) $P(x)=\frac{20^{x}\times e^{-20}}{x!}$ ۴) $P(x)=0.2^{x}(1-0.2)^{1-x}$

گزینه ۲ صحیح می‌باشد. قانون توزیع دو جمله‌ای است.

$p = p\,(\text{غیر استاندارد}) = 0.2 \Rightarrow q = 1-p = 0.8 \quad , \quad n = 100$

$$f_X(x)=\binom{n}{x}p^{x}q^{n-x}=\binom{100}{x}(0.2)^{x}(0.8)^{100-x}=C_{100}^{x}(0.2)^{x}(0.8)^{100-x}$$

مثال: اگر X متغیری تصادفی با توزیع دو جمله‌ای و دو پارامتر n و p باشد، $E\left(\frac{X}{n}\right)$ برابر خواهد شد با:

۱) $n\,p$ ۲) $n\,p\,q$

۳) p ۴) $\sqrt{npq}$

گزینه ۳ صحیح می‌باشد. در توزیع دوجمله‌ای $E(x) = np$ است. درنتیجه:

$$E\left(\frac{X}{n}\right)=\frac{E(x)}{n}=\frac{np}{n}=p$$

$$V_{ar}(X)=n\,.\,p\,.\,q=20\times0.5\times0.5=5$$

مثال: میانگین و واریانس شماره فرد در 20 بار پرتاب یک تاس سالم چقدر است؟

۱) 10 و 5 ۲) 10 و 4 ۳) 10 و 10 ۴) 5 و 5

گزینه ۱ صحیح می‌باشد. در توزیع دوجمله‌ای خواهیم داشت

$p\,(\text{فرد})=\frac{1}{2}$ ، $p\,(\text{زوج})=\frac{1}{2}$ ، $n=20$

$$E(X)=np=20\times\frac{1}{2}=10$$

$$Var(X)=n\times p\times q=20\times\frac{1}{2}\times\frac{1}{2}=5$$

مثال: تابع احتمال تعداد شماره‌های فرد در 10 بار پرتاب یک تاس سالم برابر است با:

۱) $f(x)=\frac{1}{2}$

۲) $f(x)=0.24$

۳) $f(x)=C_{10}^{x}\left(\frac{1}{2}\right)^{10}$ ، $x=0,1,2,\ldots,10$

۴) $f(x)=C_{10}^{x}\left(\frac{1}{6}\right)^{x}\left(\frac{5}{6}\right)^{10-x}$ ، $x=0,1,2,\ldots,10$

گزینه ۳ صحیح می‌باشد.

$$p\,(\text{فرد})=\frac{1}{2} \Rightarrow p\,(\text{زوج})=\frac{1}{2} \Rightarrow n=10$$

$$f_X(x)=\binom{n}{x}p^x q^{n-x}=\binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}=\binom{10}{x}\left(\frac{1}{2}\right)^{10}$$

مثال: در یک توزیع دو جمله‌ای اگر $\mu_x=3$ و $\sigma_x=\sqrt{\frac{6}{5}}$ باشد، تعداد آزمایش‌ها (n) کدام است؟

۱) 3 ۲) 4 ۳) 5 ۴) 6

گزینه ۳ صحیح می‌باشد.

$$\mu_X=E(X)=np=3 \quad ، \quad \sigma_X=\sqrt{npq}=\sqrt{\frac{6}{5}}$$

$$npq=\frac{6}{5} \Rightarrow 3q=\frac{6}{5} \Rightarrow q=\frac{2}{5} \Rightarrow p=1-\frac{2}{5}=\frac{3}{5}$$

$$np=3 \Rightarrow \frac{3}{5}n=3 \Rightarrow n=5$$

مثال: اگر کمیت X برطبق قانون دوجمله‌ای (0.3 و 8) $X\sim Bin$ توزیع شده باشد و $P(X=2)=0.294$ باشد، $P(X=3)$ برابر است با:

۱) 0.126 ۲) 0.504 ۳) 0.063 ۴) 0.252

گزینه ۴ صحیح می‌باشد. برطبق قانون دو جمله‌ای داریم:

$$X\sim Bin(n\,,p) \Rightarrow Bin(8\,,0.3) \Rightarrow n=8\,,\ p=0.3\,,\ q=0.7$$

$$P(X=3)=\binom{8}{3}(0.3)^3(0.7)^5=0.252$$

مثال: چنانچه در یک توزیع دو جمله‌ای $n=5$ و $p=\frac{1}{4}$ (احتمال موفقیت) باشد، احتمال 3 موفقیت برابر است با:

۱) 0.0879 ۲) 0.0884 ۳) 0.884 ۴) 0.879

برطبق توزیع دوجمله‌ای داریم: $p=\frac{1}{4}, q=\frac{3}{4}, n=5$.

$$f_X(x)=\binom{n}{x}p^x q^{n-x} \Rightarrow f(3)=\binom{5}{3}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2=0.0879$$

گزینه ۱ صحیح می‌باشد

مثال: در یک توزیع دوجمله‌ای میانگین برابر 5 و واریانس برابر $\frac{15}{4}$ است. مقدار $P(X=0)$ در این توزیع برابر است با:

۱) $\left(\frac{1}{4}\right)^{20}$ ۲) $\left(\frac{1}{2}\right)^{10}$ ۳) $\left(\frac{1}{4}\right)^{10}$ ۴) $\left(\frac{3}{4}\right)^{20}$

گزینه ۴ صحیح می‌باشد.

$$\mu = np \Rightarrow \frac{1}{4}n = 5 \Rightarrow n = 20$$

$$\sigma^2 = npq \Rightarrow 5q = \frac{15}{4} \Rightarrow q = \frac{3}{4} \Rightarrow p = \frac{1}{4}$$

$$P(X=0) = \binom{20}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^{20} = \left(\frac{3}{4}\right)^{20}$$

نکته تستی مهم:

$$\underset{\text{یک شکست کمتر}}{\underset{\downarrow}{np-q}} \le \text{محتمل‌ترین تعداد وقوع پیروزی} \le \underset{\text{یک پیروزی بیشتر}}{\underset{\downarrow}{np+p}}$$

مثال: در یک توزیع دو جمله‌ای امید وقوع S سه برابر امید وقوع F است. اگر در این توزیع دو جمله‌ای n=10 باشد محتمل‌ترین تعداد باری که S رخ می‌دهد برابر است با؟

۱) 4 ۲) 8 ۳) 5 ۴) 2

گزینه ۲ صحیح می‌باشد.

$$\left.\begin{array}{l} p = \frac{3}{4} \\ q = \frac{1}{4} \\ n = 10 \end{array}\right\} \Rightarrow 11 \times \frac{3}{4} - 1 \le \text{S محتمل‌ترین تعداد وقوع} \le 11 \times \frac{3}{4} \Rightarrow n = 8$$

نکته مهم: اگر در جامعه‌ای وضعیت‌های $P_1, P_2, ..., P_n$ وجود داشته باشد جامعه را **چند جمله‌ای** گوییم. و رابطه زیر برای آنها برقرار است.

$$P = \binom{n}{x_1, x_2, ..., x_n} p_1^{x_1} p_2^{x_2} .....p_k^{x_k} = \frac{n!}{x_1!x_2!.....x_k!} p_1^{x_1} p_2^{x_2} ....p_k^{x_k}$$

مثال: 0.2 افراد جامعه‌ای شاغل ، 0.3 بیکار هستند. احتمال آن‌که در یک نمونه 6 نفری شاغل و یک نفر بیکار باشد را محاسبه کنید.

می‌دانیم که $\sum_{i=1}^{k} p_i = 1$ در نتیجه

$$\Rightarrow p = \frac{6!}{2!\,1!\,3!}(0.2)^2(0.3)^1(0.5)^3$$

مثال: اظهار نظر حسابرسان راجع به حساب‌های شرکتی ممکن است، قبول، مردود و مشروط باشد. در سال 0.20، 0.30 و 0.50 نظرها به ترتیب قبول، مردود و مشروط بوده است. شش شرکت به صورت تصادفی انتخاب شده‌اند. میانگین و واریانس شرکت‌های مشروط به ترتیب از راست به چپ کدام است؟

۱) (1.2 و 1.2)　　　۲) (1.8 و 1.2)

۳) (3 و 1.5)　　　۴) (3 و 3)

گزینه ۳ صحیح می‌باشد.

بر طبق قانون چند جمله‌ای $p = 0.5$ برای شرکت‌های مشروط $q = 0.2 + 0.3 = 0.5$ برای شرکت‌های غیر مشروط

$$E(X) = n\,.\,p \ \Rightarrow\ E(X) = 6 \times 0.5 = 3$$

$$Var(X) = n\,.\,p\,.\,q \ \Rightarrow\ Var(X) = 6 \times 0.5 \times 0.5 = 1.5$$

۴ـ توزیع دو جمله‌ای منفی :

x : تعداد آزمایشهای لازم برای رسیدن به r امین موفقیت یا شکست .

$$f_x(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad ; x = r, r+1, \ldots\ldots$$

$$\mu = \frac{r}{p} \ ;\ \delta^2 = \frac{r\ q}{p^2} \quad ; q = 1-p$$

مثال: احتمال این‌که هر پرتاب بازیکنی به هدف بخورد 80 % است. احتمال این‌که سومین پرتابی که به هدف می‌خورد، پنجمین پرتاب وی باشد، چقدر است؟

۱) 0.123　　　۲) 0.512

۳) 0.64　　　۴) 0.991

گزینه ۱ صحیح می‌باشد. بنابر توزیع دو جمله منفی

$$f_X(x) = \binom{x-1}{k-1} p^k q^{x-k} \quad ، \quad p = 0.8 \quad ، \quad x = 5 \quad ، \quad k = 3$$

$$f_X(5) = \binom{5-1}{3-1}\left(\frac{8}{10}\right)^3\left(\frac{2}{10}\right)^2 = 6 \times \frac{512}{1000} \times \frac{4}{100} = 0.123$$

مثال: اگر تا انهدام کامل یک هدف، به سوی آن شلیک شود و فرض کنیم که احتمال اصابت هر راکت به هدف 0.3 است، برای انهدام کامل هدف، اصابت دو راکت لازم است. احتمال این‌که با شلیک پنجمین راکت هدف کاملاً نابود شود، چند است؟

۱) 0.6225　　۲) 0.1235　　۳) 0.2425　　۴) 0.4245

گزینه ۲ صحیح می‌باشد.

$$p_{(x=5)} = \binom{4}{1} q^3 p^2 = 4(0.7)^3(0.3)^2 = \%12.35$$

مثال: احتمال قبول شدن یک فرد در امتحانی 0.2 می‌باشد. مطلوب است محاسبه احتمال آن‌که این فرد در دفعه 7 برای بار سوم قبول شود؟

$$P=\binom{6}{2}(0.8)^4(0.2)^3=\%4.91$$

چون فرد در دفعه هفتم برای دفعه سوم قبول شده ، بنابراین دو بار در دفعات قبلی نیز قبول شده است که تعداد حالات ممکن آن $\binom{6}{2}$ است.

۵ـ توزیع هندسی :

x : تعداد آزمایشهای لازم برای رسیدن به اولین موفقیت یا شکست .

از قرار دادن r=1 در توزیع دو جمله‌ای منفی به این توزیع می‌رسیم .

$$f(x)=P^1(1-P)^{x-1}=Pq^{x-1}\quad ;\quad x=1,2,.......$$
$$\mu=\frac{1}{P}\ ;\ \delta^2=\frac{q}{P^2}\ ;q=1-P\ ;\ M_x(t)=\frac{P}{1-q^{et}}$$

$E[x]=\frac{1}{P}$ = متوسط تعداد آزمایش‌های لازم برای رسیدن به اولین پیروزی

مثال: احتمال این‌که فردی از چراغ قرمز عبور کند و پلیس متوجه نشود 0.40 است. احتمال این‌که در حین عبور از چهارمین چراغ قرمز بالاخره جریمه شود، چقدر است؟

۱) 0.0384 ۲) 0.0864 ۳) 0.216 ۴) 0.60

بنابر توزیع هندسی خواهیم داشت :

$$p=0.6\ ,\ q=0.4\ ,\ x=4$$
$$f_X(x)=pq^{x-1}=(0.6)(0.4)^3=0.0384$$

گزینه ۱ صحیح می‌باشد.

مثال: 10 درصد تولیدات کارخانه‌ای معیوبند. احتمال این‌که سومین کالای کنترل شده، اولین کالای معیوب باشد، چقدر است؟

۱) 0.001 ۲) 0.081 ۳) 0.10 ۴) 0.817

برطبق قانون هندسی داریم: $P=0.10\ ,\ q=0.98$

$$f(x)=pq^{x-1}\ \Rightarrow\ f(3)=(0.1)(0.9)^2=0.081$$

گزینه ۲ صحیح می‌باشد

مثال: احتمال اصابت موشکی به یک جنگنده 0.3 است. با اصابت یک موشک، جنگنده سقوط خواهد کرد. احتمال این‌که در پرتاب پنجمین موشک، جنگنده سقوط کند، چقدر است؟

۱) 0.05 ۲) 0.005 ۳) 0.072 ۳) 0.081

گزینه ۳ صحیح می‌باشد. برطبق توزیع هندسی داریم: $p=0.3\ ,\ q=0.7\ ,\ x=5$

$$f(x)=q^{x-1}p \Rightarrow f(5)=(0.7)^{5-1}(0.3)^{1}=0.072$$

مثال: در یک ظرف 10 توپ سفید و 5 توپ سیاه داریم، یک توپ به طور تصادفی انتخاب می‌کنیم. پس از مشاهده رنگ آن، مجدداً آن را به داخل ظرف باز می‌گردانیم و این عمل را آن‌قدر تکرار می‌کنیم تا توپ سیاه انتخاب شود، متوسط تعداد انتخاب (تکرار) کدام است؟

۱) 3 ۲) 5 ۳) 6 ۴) 2

انتخاب توپ سیاه بر طبق توزیع هندسی است، زیرا احتمال ثابت است و تا انتخاب توپ سیاه، آزمایش ادامه دارد.

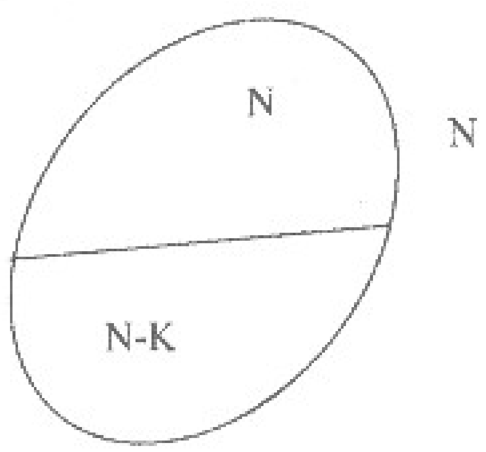
$$p=q\ (\text{سیاه})=\frac{5}{15} \quad ، \quad q=p\ (\text{سفید})=\frac{10}{15}$$

$$E(X)=\frac{1}{p}=\frac{1}{\frac{5}{15}}=3$$

گزینه ۱ صحیح می‌باشد.

۶ـ توزیع فوق هندسی :

جامعه‌ای با N عضو، را به دو قسمت K و $N-K$ عضوی تقسیم می‌کنیم . و یک نمونهٔ n تائی از آن انتخاب می‌کنیم :



x : تعداد نمونه‌هائی که متعلق به قسمت k تائی هستند.

$$f(x)=\frac{\binom{k}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \ ; x=0,1,2,\ldots\ldots,n$$

$$\mu=n\frac{k}{N},\ \delta^2=\frac{N-n}{N-1}\times\frac{k}{N}\times\left(1-\frac{k}{N}\right)\times n$$

در حالت خاص اگر $\frac{n}{N}<0.05$ باشد . آنگاه می‌توانیم به جای این توزیع از توزیع دو جمله‌ای استفاده کنیم. بصورت:

$$P=\frac{k}{N}\ ;q=1-\frac{k}{N}$$

$$\text{در نتیجه} \rightarrow \mu=nP,\delta^2=\frac{N-n}{N-1}\times n\times P\times q$$

به $\frac{N-n}{N-1}$ ضریب تصحیح توزیع فوق هندسی می‌گویند . که با 1 شدن آن دقیقاً به توزیع دو جمله‌ای می‌رسیم .

مثال: در یک کارخانه با200 قطعه تولیدی 8 قطعه معیوب است یک نمونه ده‌تایی از آن انتخاب می‌کنیم مطلوبست احتمال آن که 1 قطعه از نمونه معیوب باشد:

$$\begin{matrix}N=200\\ n=10\end{matrix} \quad P(x=1)=\frac{\binom{8}{1}\binom{192}{9}}{\binom{200}{10}}=\boxed{0.2878}$$

مثال: از جوراب‌های بسته‌بندی شده در جعبه‌ای 9 عدد سالم و 3 عدد معیوب است. یک مشتری به طور تصادفی 4 عدد را خریداری می‌کند. میانگین و واریانس تعداد جوراب‌های معیوب در این خرید به ترتیب از چپ به راست چقدر است؟

۱) $\frac{6}{11}$ و 1　　۲) $\frac{3}{4}$ و 1　　۳) $\frac{27}{12}$ و 1　　۴) $\frac{9}{12}$ و 2

گزینه ۱ صحیح می‌باشد. چون انتخاب بدون جای‌گذاری است، پس تعداد انتخاب بر طبق توزیع فوق هندسی می‌باشد.

	معیوب	سالم	کل
جامعه :	3	9	12
نمونه :	-	-	4

$N=12 \quad , \quad n=4 \quad , \quad k=3$

$$E(X)=n\,.\,p=n\times\frac{k}{N}=4\times\frac{3}{12}=1$$

$$Var=n\,.\,p\,.\,q\,.\left(\frac{N-n}{N-1}\right)=4\times\frac{3}{12}\times\frac{9}{12}\times\frac{12-4}{12-1}=\frac{6}{11}$$

مثال: کیسه‌ای با حاوی 5 مهره قرمز و 4 مهره سفید است. سه مهره بدون جایگذاری از کیسه خارج می‌شود، توزیع احتمال تعداد مهره‌های قرمز کدام یک است؟

۱) پواسن　　۲) دو جمله‌ای　　۳) فوق هندسی　　۴) یکنواخت

$$f_X(x)\frac{\binom{5}{x}\binom{4}{3-x}}{\binom{9}{3}} \quad ; x=0\,,1\,,2\,,3$$

گزینه ۳ صحیح می‌باشد.

مثال: در یک پارتیشن از محصولات تولید شده به حجم $N = 10$ که تعداد $M = 2$ محصول آن غیراستاندارد می‌باشد، به طور تصادفی 4 واحد محصول را انتخاب می‌کنیم، احتمال این که هیچ یک از محصولات غیراستاندارد نباشد، چقدر است؟

۱) $\frac{3}{4}$　　۲) $\frac{2}{3}$　　۳) $\frac{5}{6}$　　۴) $\frac{1}{3}$

بر اساس قانون توزیع فوق هندسی داریم:

$$f_X(x)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \Rightarrow f_X(0)=\frac{\binom{2}{0}\binom{8}{4}}{\binom{10}{2}}=\frac{1}{3}$$

گزینه ۴ صحیح می‌باشد.

مثال: از یک جامعه 4000 نفره یک نمونهٔ تصادفی 20 تایی انتخاب شده است. در این حالت تابع احتمال متغیر تصادفی X ، کدم است؟

۱) هندسی　　۲) دوجمله‌ای

۳) فوق هندسی　　۴) هم فوق هندسی و هم دو جمله‌ای

گزینه ۴ صحیح می‌باشد.

در صورتیکه در توزیع فوق هندسی $\frac{n}{N}\leq 0.05$ باشد توزیع فوق هندسی تقریبی از توزیع دو جمله‌ای است.

$$\frac{n}{N}=\frac{40}{4000}=0.01<0.05$$

N: تعداد اعضای جامعه

n: تعداد اعضای نمونه

۷ـ توزیع پواسن :

x: تعداد اتفاقات در یک فاصله زمانی یا مکانی

λ: متوسط اتفاقات در فاصله زمانی یا مکانی

$$f(x,\lambda)=\frac{e^{-\lambda}\ \lambda^{x}}{x\ !}\ ;\ \ x=0,1,2,.........$$

$$\mu=E\ \ (x)=\lambda\ \ ;\ \ \delta=\ \lambda;$$

نکته مهم :

۱ـ مد توزیع پواسن یک عدد صحیح است . که در رابطه زیر صدق می کند.

$$\lambda-1\le mo\le\lambda$$

۲ ـ λ با توجه به تغییر فاصله زمانی یا مکانی باید به یک نسبت تغییر کند.

نکته: در حل مسائل e برابر 2.718 در نظر گرفته می‌شود.

$\lambda=E[x]=$ تعداد متوسط اتفاقات در یک فاصله زمانی یا مکانی

نکته مهم: در حل مسائل مربوط به توزیع پواسن هر گاه λ برای یک واحد زمانی یا مکانی خاص داده شده باشد، برای محاسبه λ برای یک واحد زمانی یا مکانی دیگر باید λ اصلی را در ضریبی ضرب کرد. به مثال‌های زیر توجه کنید:

$\lambda=1$ برای نیم ساعت $\Leftarrow$ $\lambda=2$ برای یک ساعت $\lambda=0.5$ برای یک ربع $\Leftarrow$

مثال: تعداد قطعات معیوب در هر روز بر روی یک ماشین از توزیع پواسن برخوردار است و دارای **متوسط 5 قطعه معیوب** در روز است. احتمال این‌که در یک روز هیچ قطعهٔ معیوبی تولید نشود، چقدر است؟

۱) $\frac{1}{5}e^{-1.5}$ ۲) $5e^{5}$ ۳) e^{-5} ۴) $5e^{-5}$

گزینه ۳ صحیح می‌باشد. برطبق توزیع پواسن داریم: $\lambda=5$ قطعه معیوب در روز . و چون مسئله در یک روز خواسته شده λ تغییری نمی کند.

$$f(x)=\frac{\lambda^{x}.\ e^{-\lambda}}{x!}\ \ \Rightarrow\ \ f(0)=\frac{5^{0}\times e^{-5}}{0!}=e^{-5}$$

مثال: در یک نوار خاص، به طور متوسط یک عیب در هر 220 متر وجود دارد، احتمال این‌که 2 عیب در یک بسته 880 متری وجود داشته باشد، کدام است؟

۱) e^{-4} ۲) e^{-1} ۳) $4e^{-2}$ ۴) $8e^{-4}$

گزینه ۴ صحیح می‌باشد. **بر طبق توزیع پواسن $\lambda=1$ عیب در 220 متر و $\lambda=1\times4=4$ عیب در 880 متر.**

$$f_{X}(x)=\frac{\lambda^{x}e^{-\lambda}}{x!}\ \ \Rightarrow\ \ f(2)=\frac{4^{2}\times e^{-4}}{2!}=8e^{-4}$$

مثال: به طور متوسط هر 10 دقیقه یک مشتری وارد یک فروشگاه می‌شود، احتمال این که در 20 دقیقه 2 مشتری وارد شوند، چقدر است؟

۱) $4e^{-4}$ ۲) e^{-1} ۳) $3e^{-3}$ ۴) $2e^{-2}$

بر طبق توزیع پواسن $\lambda = 1$ مشتری در 10 دقیقه در نتیجه $\lambda = 1\times 2 = 2$ مشتری در 20 دقیقه.

$$f_x(x) = \frac{\lambda^x e^{-\lambda}}{x!} \Rightarrow f(2) = \frac{2^2 \times e^{-2}}{2!} = 2e^{-2}$$

گزینه ۴ صحیح می‌باشد.

مثال: در یک توزیع پواسن اگر $P(X=1)=P(X=2)$ باشد، آن‌گاه مقدار $P(X=0)$ برابر است با:

۱) $\frac{1}{2}$ ۲) $2e^{-2}$ ۳) e^{-1} ۴) e^{-2}

$$P(X=1)=P(X=2) \Rightarrow \frac{e^{-\lambda}.\lambda}{1!} = \frac{e^{-\lambda}.\lambda^2}{2!} \Rightarrow 2\lambda = \lambda^2 \Rightarrow \lambda^2 - 2\lambda = 0 \Rightarrow \lambda = 0 \text{ یا } \lambda = 2$$

$$P(X=0) = \frac{e^{-2}.2^0}{0!} = e^{-2}$$

گزینه ۴ صحیح می‌باشد.

مثال: اگر واریانس تعداد مراجعین به یک بانک از ساعت خاصی از روز 4 نفر باشند، احتمال این که در ربع اول ساعت کسی به بانک مراجعه نکند، چقدر است؟

۱) $\frac{1}{e^2}$ ۲) $\frac{1}{e}$ ۳) e^{-4} ۴) $e^{\frac{1}{2}}$

بر طبق توزیع پواسن $E(x) = \sigma^2(x) = \lambda$ در نتیجه در این سوال $\lambda = \sigma^2 = 4$ نفر در ساعت است.بنابراین $\lambda = \frac{4}{4} = 1$ نفر در یک ربع است. و خواهیم داشت:

$$p(X=x) = \frac{\lambda^x.e^{-\lambda}}{x!} \Rightarrow P(X=0) = \frac{1^0 \times e^{-1}}{0!} = e^{-1} = \frac{1}{e}$$

گزینه ۲ صحیح می‌باشد.

مثال: در یک توزیع پواسن $\frac{P(X=0)}{P(X=1)} = \frac{1}{3}$ است، میانگین این توزیع برابر است با:

۱) 2 ۲) 3 ۳) 4 ۴) 5

گزینه ۲ صحیح می‌باشد. بر طبق توزیع پواسن $E(x) = \sigma^2(x) = \lambda$ در نتیجه:

$$\frac{P(X=0)}{P(X=1)} = \frac{\frac{\lambda^0.\cancel{e^{-\lambda}}}{0!}}{\frac{\lambda^1.\cancel{e^{-\lambda}}}{1!}} = \frac{1}{3} \Rightarrow \lambda = 3 \Rightarrow E(x) = \lambda = 3$$

مثال: به طور متوسط 6 ماشین در دقیقه از یک جاده می‌گذرد، عرض جاده به اندازه عبور یک اتومبیل است. شخصی بدون توجه به عبور ماشین‌ها عرض جاده را در 10 ثانیه می‌پیماید، احتمال سالم ماندن او چقدر است؟

۱) e^{-1} ۲) $1-e^{-1}$ ۳) e^{-6} ۴) $1-e^{-6}$

گزینه ۱ صحیح می‌باشد.

برطبق قانون پواسن داریم: $6=\lambda$ ماشین در دقیقه در نتیجه $1=\frac{6}{6}=\lambda$ ماشین در 10 ثانیه برای آنکه شخص سالم بماند نباید در مدت 10 ثانیه هیچ ماشینی عبور کند. بنابراین باید $p(x=0)$ را محاسبه کنیم.

$$f_X(x)=\frac{\lambda^x e^{-\lambda}}{x!} \Rightarrow f(0)=\frac{1^0\times e^{-1}}{0!}=e^{-1}$$

مثال: مسافران هواپیمایی به صورت تصادفی و به تعداد 5 نفر در هر دقیقه وارد فرودگاه می‌شوند. احتمال این‌که در دقیقه خاصی هیچ مسافری به فرودگاه وارد نشود، چقدر است؟

۱) $(2.718)^5$ ۲) $\frac{5}{(2.718)^5}$ ۳) $\frac{5}{2.718}$ ۴) $(2.718)^{-5}$

گزینه ۴ صحیح می‌باشد.

$$f_X(x)=\frac{e^{-\lambda}.\lambda^x}{x!} \quad ، \quad \lambda=5 \quad ، \quad x=0$$

$$f_X(0)=\frac{e^{-5}\times 5^0}{0!}=\frac{e^{-5}}{1}=(2.718)^{-5}$$

مثال: یک دستگاه مکانیکی به طور متوسط 2 بار در سال نیاز به تعمیر پیدا می‌کند، احتمال این‌که در هر 6 ماه، حداقل یک بار تعمیر شود، چقدر است؟

۱) $1-e^{-1}$ ۲) $1-e^{-2}$ ۳) e^{-1} ۴) e^{-12}

گزینه ۱ صحیح می‌باشد. برطبق قانون پواسن داریم. $2=\lambda$ دستگاه در سال و $1=\frac{2}{2}=\lambda$ دستگاه در 6 ماه:

$$P(X\ge 1)=1-P(X<1)=1-P(X=0)=1-\frac{1^0\times e^{-1}}{1!}=1-e^{-1}$$

مثال: یک دستگاه مکانیکی به طور متوسط 24 بار در سال نیاز به تعمیر دارد، احتمال این‌که در ماه حداقل یک‌بار تعمیر شود، چقدر است؟

۱) $1-e^{-2}$ ۲) $1-e^{-24}$ ۳) e^{-2} ۴) e^{-24}

. برطبق توزیع پواسن داریم. $24=\lambda$ دستگاه در سال و $2=\frac{24}{12}=\lambda$ دستگاه در ماه :

$$P(X\ge 1)=1-P(X<1)=1-P(X=0)=1-\frac{2^0\times e^{-2}}{0!}=1-e^{-2}$$

گزینه ۱ صحیح می‌باشد

مثال: در یک توزیع پواسن $\frac{P(X=0)}{P(X=1)}=\frac{1}{3}$ است، میانگین این توزیع برابر است با:

۱) 2　　۲) 3　　۳) 4　　۴) 5

گزینه ۲ صحیح می‌باشد. بر طبق توزیع پواسن $E(x)=\sigma^2(x)=\lambda$ در نتیجه:

$$\frac{P(X=0)}{P(X=1)}=\frac{\frac{\lambda^0 \cdot \cancel{e^{-\lambda}}}{0!}}{\frac{\lambda^1 \cdot \cancel{e^{-\lambda}}}{1!}}=\frac{1}{3} \Rightarrow \lambda=3 \Rightarrow E(x)=\lambda=3$$

۳- بعضی اوقات می‌توان در حل مسائل به جای استفاده از توزیع دو جمله‌ای از تقریب پواسن استفاده نمود. برای این کار باید یکی از جفت شرط‌های زیر برقرار باشد:

1) $np \le 10 \quad , \quad n \ge 100$
2) $p \le 0.05 \quad , \quad n \ge 20$

در این صورت داریم:

$$p_{(x)}=\frac{e^{-np}(np)^x}{x!} \qquad \lambda=np$$

مثال: در کدام‌یک از موارد زیر توزیع پواسن تقریب خوبی برای توزیع دوجمله‌ای است؟

۱) $n=25$ و $p=0.04$　　۲) $n=50$ و $p=0.28$

۳) $n=60$ و $p=0.58$　　۴) $n=150$ و $p=0.93$

حل: گزینه ۱ صحیح می‌باشد. در توزیع دوجمله‌ای در دو وضعیت زیر می‌توانیم از تقریب توزیع پواسن استفاده کنیم.

$$\begin{cases} n \ge 20 \,, p \le 0.05 \\ n \ge 100 \,,\ np \le 10 \end{cases}$$

در این مسئله در گزینه ۱ با توجه به آنکه n=25 و p=0.04 است. در شرایط بالا صدق می‌کند. و بقیه گزینه‌ها شرط لازم را ندارند.

مثال: در صورتی که یک توزیع دو جمله‌ای دارای 100 تکرار باشد و احتمال موفقیت در هر تکرار 0.01 باشد، بهترین توزیع برای تقریب احتمال‌های آن کدام است؟

۱) نمایی　　۲) نرمال　　۳) پواسن　　۴) یکنواخت

گزینه ۳ صحیح می‌باشد.

در توزیع دوجمله‌ای در دو وضعیت زیر می‌توانیم از تقریب توزیع پواسن استفاده کنیم.

$$\begin{cases} n \ge 20 \,, p \le 0.05 \\ n \ge 100 \,,\ np \le 10 \end{cases}$$

در این مسئله $np=100\times 0.01=1\le 10$ است. در نتیجه تقریب پواسن بهتر است.

مثال: نسبت خرابی کالا در یک کارخانه برابر 0.01 است. احتمال آن‌که در 100 کالا حداکثر یک کالای خراب وجود داشته باشد، چقدر است؟

۱) e^{2} ۲) e^{-2} ۳) $2e$ ۴) $2e^{-1}$

گزینه ۴ صحیح می‌باشد.

در یک توزیع دو جمله‌ای هرگاه $\begin{cases} n \ge 20 \,,\, p \le 0.05 \\ n \ge 100 \,,\, np \le 10 \end{cases}$ درنتیجه می‌توانیم از تقریب پواسن استفاده کنیم(و $\lambda = np$ در نظر گرفته می‌شود).

در این مثال نیز: $n = 100 \;,\; p = 0.01 \Rightarrow np = 1 \le 10$ در نتیجه $\lambda = np = 1$ در نظر گرفته و خواهیم داشت:

$$P(X \le 1) = P(X=0) + P(X=1) = \frac{1^0 \times e^{-1}}{0!} + \frac{1 \times e^{-1}}{1!} = 2e^{-1}$$

توزیع‌های مهم پیوسته :

۱ـ توزیع یکنواخت پیوسته :

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{سایر نقاط} \end{cases}$$

$$\mu = \frac{a+b}{2} \quad ; \delta^2 = \frac{(b-a)^2}{12}$$

مثال: توزیع یکنواخت زیر را در نظر بگیرید:

$$f(x) = \begin{cases} \dfrac{1}{2} \,; 1 < x < 3 \\ 0 \;; \text{سایر نقاط} \end{cases}$$

مقدار واریانس این توزیع عبارت است از؟

۱) 2 ۲) $\frac{1}{2}$ ۳) $\frac{1}{3}$ ۴) 1

گزینه ۳ صحیح می‌باشد.

$$var(x) = \frac{(b-a)^2}{12} = \frac{(3-1)^2}{12} = \boxed{\frac{1}{3}}$$

مثال: واریانس متغیر تصادفی x با چگالی $\left\{ f(x) = \frac{2}{3} \,|\, -1 < x < \frac{1}{2} \right\}$ کدام است؟

۱) $\frac{1}{16}$ ۲) $\frac{3}{16}$ ۳) $\frac{4}{9}$ ۴) $\frac{9}{4}$

$f(x) = \frac{2}{3}$ تابع چگالی یکنواخت پیوسته در فاصله $-1 < x < \frac{1}{2}$ است. بنابراین واریانس آن $\frac{(b-a)^2}{12} = \frac{\frac{1}{2} - (-1)}{12} = \frac{3}{16}$ خواهد بود.

گزینه ۲ صحیح می‌باشد.

مثال: میانگین متغیر تصادفی X با این تابع چگالی احتمال‌ها چقدر است؟

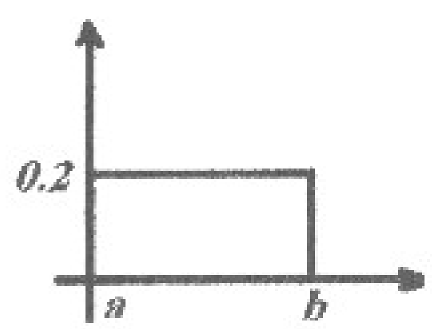
$$\varphi(x)=\frac{2}{3} \qquad -1\le x\le \frac{1}{2}$$

۱) $-\frac{3}{4}$ ۲) $-\frac{1}{4}$ ۳) $\frac{1}{4}$ ۴) $\frac{3}{4}$

$f(x)=\frac{2}{3}$ تابع چگالی یکنواخت پیوسته در فاصله $-1<x<\frac{1}{2}$ است. بنابراین میانگین آن $\frac{a+b}{2}=\frac{\frac{1}{2}+(-1)}{2}=-\frac{1}{4}$ خواهد بود.

گزینه ۲ صحیح می‌باشد.

مثال: فرض کنید در شکل زیر متغیر تصادفی X توزیع یکنواخت دارد، c را به نحوی پیدا کنید که $P(x\le c)=0.6$ باشد.



۱) 3

۲) 4

۳) 5

۴) 6

گزینه ۱ صحیح می‌باشد.

$$\varphi(x)=\frac{1}{b-a} \Rightarrow 0.2=\frac{1}{b-0} \Rightarrow b=5$$

$$P(X\le c)=0.6 \Rightarrow \int_0^c \frac{1}{5}dx=0.6 \Rightarrow \left[\frac{1}{5}x\right]_0^c=0.6 \Rightarrow c=3$$

توزیع نمائی :

x : مدت زمان لازم برای وقوع اولین اتفاق یا مدت زمان لازم برای وقوع دو اتفاق متوالی .

نکته: به طور کلی اگر λ متوسط تعداد اتفاقات در توزیع پواسن باشد $\frac{1}{\lambda}$ مدت زمان لازم برای وقوع اولین اتفاق یا دو اتفاق متوالی است که با β نشان داده می‌شود. یعنی: $\beta=\frac{1}{\lambda}$

$$f(x)=\lambda e^{-\lambda x}=\frac{1}{\beta}e^{-\frac{x}{\beta}} \;;\infty>x\ge 0$$

$$E(x)=\beta=\frac{1}{\lambda},\delta^2=\beta^2=\frac{1}{\lambda^2}$$

نکته مهم : در توزیع نمائی دارای خاصیت بی‌حافظگی یا عدم حافظه هستیم :

$$P(x>s+t \mid x>s)=P(x>t)$$

کاربرد توزیع : تعیین طول عمر قطعات و ....

نکته: β همچنین برابر متوسط اتفاقات در واحد زمانی نیز می‌باشد $(E[x])$

مثال: به طور متوسط در هر ساعت 6 نفر وارد فروشگاهی می‌شوند، پس:

$\lambda = 6$ برای یک ساعت $\Rightarrow \beta = \dfrac{\text{یک ساعت}}{6} = 10$ دقیقه

مدت زمان لازم برای ورود اولین مشتری یا ورود مشتری بعدی $= \beta = 10\ \text{min}$

مثال: توزیع ورود مشتریان به فروشگاهی دارای توزیع نمایی با میانگین 3 min است. مطلوب است محاسبهٔ احتمال آن‌که فروشنده حداقل 5 min منتظر بماند تا اولین مشتری وارد شود.

روش اول:

$\beta = 3\,\text{min}\, E[x] \Rightarrow \lambda = \frac{1}{3}$ (نفر در دقیقه)

روش اول:نمایی

$p(x \geq 5) = \int_5^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = e^{-x/3} \Big|_5^\infty = e^{-\frac{5}{3}}$ (برای 5 دقیقه محاسبه شده است)

روش دوم:پواسن

احتمال آن‌که در 5 دقیقه هیچ مشتری وارد فروشگاه نشود

$P(x=0) = \dfrac{e^{-\frac{5}{3}}\left(-\frac{5}{3}\right)^0}{0!} = e^{-\frac{5}{3}}$ (برای 5 min محاسبه شده است)

$\lambda = \frac{1}{3}$ در یک دقیقه $\Rightarrow \lambda = \frac{5}{3}$ در 5 دقیقه

مثال: در مثال فوق فروشنده 5 ساعت منتظر مانده است محاسبه احتمال آن‌که 5 دقیقه دیگر نیز منتظر بماند؟

از آن‌جایی که توزیع نمایی حافظه ندارد جواب همان مقدار بدست آمده برای مثال فوق می‌باشد.

مثال: فرض کنید زمان بین ورود هر دو مشتری به یک فروشگاه به صف صندوق دارای توزیع (پخش) نمایی با میانگین $\frac{1}{3}$ دقیقه است. در این صورت، احتمال آن‌که در 2 دقیقه 3 نفر وارد صف صندوق شوند چقدر است؟

۱) $\frac{6^2}{e^6}$

۲) $\frac{6^2}{e^6}$

۳) $\frac{2^6}{e^6}$

۴) $\frac{3^6}{e^6}$

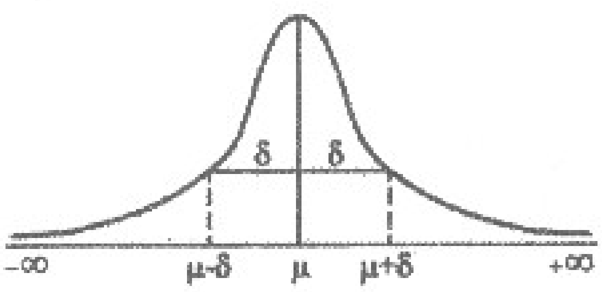
$\beta = \frac{1}{3}\text{min} \Rightarrow \lambda = 3$ نفر در دقیقه $\Rightarrow \lambda = 6$ نفر در دو دقیقه

$p(x=3) = \dfrac{e^{-6} 6^3}{3!} = \boxed{36e^{-6}}$

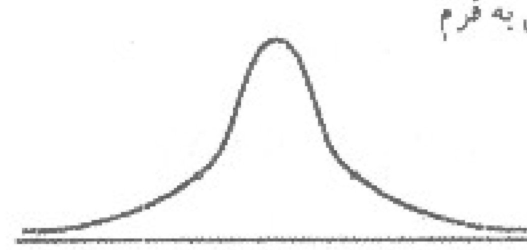
۴ـ توزیع نرمال : یکی از مهمترین توزیع‌های پیوسته می‌باشد. این توزیع یک توزیع متقارن با مشخصات زیر می‌باشد:

۱ـ خط $x = \mu$ هم میانه و هم‌مد است.

۲ـ سطح زیر منحنی برابر I است.



۳ـ تابع چگالی به فرم



$$f(x) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{(x-\mu)^2}{2\delta^2}} \quad ; -\infty < x < \infty \; ; -\infty < \mu < +\infty ; \delta > 0$$

$$E(x) = \mu \quad ; \delta^2 = Var(x) \quad ; \quad M_x(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

۶ـ مهم : هر توزیع نرمال با $N(\mu, \delta^2)$ با تغییر متغیر $Z = \frac{x-\mu}{\delta}$ به توزیع نرمال استاندارد، به فرم $f_{(Z)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ تبدیل می‌شود

نکته مهم: برای محاسبه هر احتمال از توزیع نرمال ، به شکل زیر عمل می‌کنیم:

$$P(x < \alpha) = P\left(\frac{X-\mu}{\delta} < \frac{\alpha-\mu}{\delta}\right) = P\left(z < \frac{\alpha-\mu}{\delta}\right) = \varphi\left(\frac{\alpha-\mu}{\delta}\right)$$
$$= N\left(\frac{\alpha-\mu}{\delta}\right)$$

این دو فرم مختلف نمایش احتمال در حالت استاندارد است.

عبارت $\varphi(z)$ یا $N(z)$ را به صورت زیر می‌توان نشان داد:

$$F_Z(z) = P(Z < z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\,\delta} e^{\frac{-z^2}{2}} dz = \varphi(z)$$

مثال: اگر اندازه دو نفر از جامعهٔ نرمالی 13 و 19 و اندازهٔ این دو بر حسب متغیر استاندارد، صفر و 3 باشد، میانگین و انحراف معیار به ترتیب (از چپ به راست) کدامند؟

۱) 19 و 2 ۲) 3 و 6 ۳) 2 و 13 ۴) 3 و 19

گزینه ۳ صحیح می‌باشد.

$$Z = \frac{X-\mu}{\sigma} \Rightarrow \begin{cases} 0 = \frac{13-\mu}{\sigma} \Rightarrow \mu = 13 \\ 3 = \frac{19-\mu}{\sigma} \Rightarrow 3\sigma = 6 \Rightarrow \sigma = 2 \end{cases}$$

۷ـ همیشه در توزیع نرمال استاندارد $\mu = 0$ و $\delta^2 = 1$ است یعنی :

$N(\mu, \delta^2) \to N(0,1)$ استاندارد شدن

مثال: برای متغیر استاندارد Z ، امید ریاضی (μ) و واریانس (σ^2) کدامند؟

۱) (0 و 1) ۲) (1 و 0) ۳)(0 و 0) ۴) (1 و 1)

گزینه ۲ صحیح می‌باشد.

مثال: فرض کنید X، دارای توزیع نرمال با میانگین 4 و انحراف معیار 3 باشد، اگر $Y = X - 3$ باشد، احتمال Y بزرگ‌تر از 1، یعنی $P(Y \geq 1)$ کدام است؟

۱) $\frac{1}{2}$ ۲) صفر ۳) $\frac{3}{4}$ ۴) 1

گزینه ۱ صحیح می‌باشد.

$$E(Y) = E(X) - 3 = 4 - 3 = 1$$

$$\sigma^2(Y) = \sigma^2\left(X - \overset{0}{\cancel{3}}\right) = 9 \Rightarrow \sigma(Y) = 3$$

$$P(Y \geq 1) = P\left(Z \geq \frac{1 - E(y)}{\sigma_Y}\right) = P\left(Z \geq \frac{1-1}{3}\right) = P(Z \geq 0) = \frac{1}{2}$$

مثال: اگر کمیت $X \sim N(50 \text{ و } 9)$ توزیع شده باشد، مقدار استاندارد شده $Z = -1.2$ متناظر باشد، برابر است با:

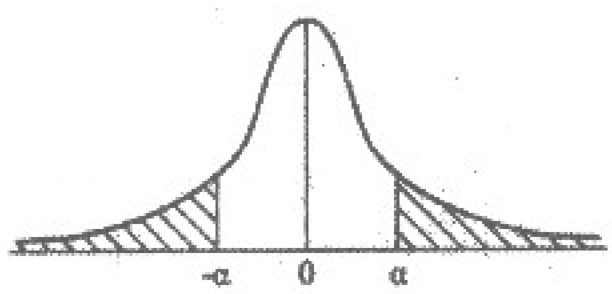
۱) 46.4 ۲) 60.8 ۳) 39.2 ۴) 53.6

گزینه ۱ صحیح می‌باشد.

$$X \sim N(\mu, \sigma^2) = X \sim N(50, 9) \Rightarrow \mu = 50 \;,\; \sigma^2 = 9 \;,\sigma = 3$$

$$Z = \frac{X - \mu}{\sigma} \Rightarrow -1.2 = \frac{X - 50}{3} \Rightarrow X = 46.4$$

۸. **مهم :** در دو نقطه متناظر α، $-\alpha$ همیشه به علت تقارن در توزیع نرمال استاندارد خواهیم داشت:



$$P(Z < \alpha) = P(Z > -\alpha)$$
$$P(z > \alpha) = P(z < -\alpha)$$

اگر $P(z > \alpha) = P(z > \beta)$ یا $P(z < \alpha) = P(z < \beta)$ شود نتیجه می‌گیریم ، $\alpha = \beta$ است. (باز هم به علت تقارن)

اگر $P(z > \alpha) = P(z < \beta)$ یا $P(z < \alpha) = P(z > \beta)$ شود نتیجه می‌گیریم ، $\alpha = -\beta$ است. (باز هم به علت تقارن)

مثال: اگر توزیع X نرمال با میانگین 100 و انحراف معیار 10 باشد، و $P(X \leq \alpha) = 0.0668$ باشد، مقدار α کدام است؟

(راهنمایی: $\int_{-\infty}^{1.5} f(z)dz = 0.9332$)

۱) 50 ۲) 85 ۳) 115 ۴) 150

گزینه ۲ صحیح می‌باشد.

$$(1)\int_{-\infty}^{1.5} f(z)dz = P(z < 1.5) = 0.9332 \Rightarrow P(z > 1.5) = 0.0668$$

$$(2)\quad P(X<a)=P\left(\frac{X-\mu}{\sigma}<\frac{a-100}{10}\right)=P\left(Z<\frac{a-100}{10}\right)=0.0668$$

$$(1),(2)\Rightarrow \frac{a-100}{10}=-1.5 \Rightarrow a=85$$

مثال: توزیع x نرمال با انحراف معیار 10 می‌باشد.اگر $P(X\geq 100)=0.975$ باشد، مقدار میانگین چقدر است؟

(راهنمایی: $\int_{1.96}^{+\infty} f(z)dz=0.025$)

۱) 60
۲) 80.4
۳) 140
۴) 119.6

گزینه ۱ صحیح می‌باشد.

$$(1)\quad \int_{1.96}^{+\infty} f(z)dz=P(Z\geq 1.96)=0.025\Rightarrow P(Z\leq 1.96)=0.975$$

$$(2)\quad P(X\geq 100)=P\left(\frac{X-\mu}{\sigma}\geq\frac{100-\mu}{10}\right)=P\left(Z\geq\frac{100-\mu}{10}\right)=0.975$$

$$(1),(2)\Rightarrow\frac{100-\mu}{10}=-1.96\Rightarrow\mu=119.6$$

مثال: توزیع کمیت تصادفی X نرمال بوده و میانگین و انحراف معیار آن به ترتیب 20 و 4 باشد، $P(X\geq 28)$، کدام است؟

(راهنمایی: $P(-2<Z<2)=0.9544$)

۱) 0.0228
۲) 0.0456
۳) 0.4772
۴) 0.9544

گزینه ۱ صحیح می‌باشد.

$$(1)\quad P(-2\leq z\leq 2)=0.9544\Rightarrow P(0\leq z\leq 2)=\frac{0.9544}{2}=0.2772\Rightarrow P(z\geq 2)=0.5-0.2772=0.0228$$

$$(2)\quad P(X\geq 28)=P\left(\frac{X-\mu}{\sigma}\geq\frac{28-20}{4}\right)=P(Z\geq 2)=0.0228$$

مثال: میانگین توزیع نمرات دانشجویان یک دانشکده 52 با انحراف معیار 10 می‌باشد. احتمال این که نمرهٔ یکی از دانشجویان کمتر از 72 باشد، چقدر است؟ (راهنمایی: $P(Z\leq -2)=0.0228$).

۱) 0.0228
۲) 0.5
۳) 0.5793
۴) 0.9772

گزینه ۴ صحیح می‌باشد.

$$(1)P(z\leq -2)=0.0228\Rightarrow P(z\geq -2)=0.9772$$

$$(2)P(x<72)=P\left(\frac{X-\mu}{\sigma}<\frac{72-52}{10}\right)=P(z<2)$$

با توجه به آنکه می‌دانیم همیشه $P(z<\alpha)=P(z>-\alpha)$ یا برعکس:

$$(1),(2)\Rightarrow P(z<2)=P(z>-2)=0.9772$$

مثال: توزیع x نرمال با میانگین 100 و انحراف معیار 10 است. اگر $P(X \ge x) = 0.0228$ باشد. مقدار x چقدر است؟

(راهنمایی: $\int_{-4}^{-2} f(z)dz = 0.0228$)

۱) 60 ۲) 80

۳) 120 ۴) 140

گزینه ۲ صحیح می‌باشد.

$$(1) \int_{-\infty}^{-2} f(z)dz = P(z \le -2) = 0.0228$$

$$(2) P(X \ge x) = P\left(\frac{X-\mu}{\sigma} \ge \frac{x-100}{10}\right) = P\left(z \ge \frac{x-100}{10}\right) = 0.0228$$

$$(1) و (2) -2 = \frac{X-100}{10} \Rightarrow X = 80$$

مثال: توزیع متغیر تصادفی X نرمال با میانگین 100 و انحراف معیار 10 است. اگر $P(X \ge x) = 0.0495$ باشد، مقدار x چقدر است؟

(راهنمایی: $\int_{-4}^{-1.65} f(z)dz = 0.0495$)

۱) 60 ۲) 83.5

۳) 116.5 ۴) 140

گزینه ۳ صحیح می‌باشد.

$$(1) \int_{-4}^{-1.65} f(z)dz = \int_{-\infty}^{-1.65} f(z)dz = P(z < -1.65) = 0.0495$$

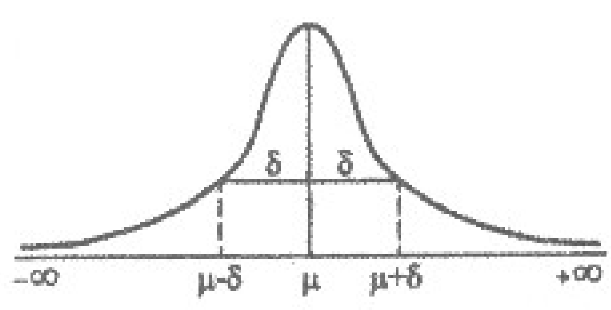
$$(2) P(X \ge x) = P\left(\frac{X-\mu}{\sigma} \ge \frac{x-100}{10}\right) = P\left(z \ge \frac{x-100}{10}\right) = 0.0495$$

$$\Rightarrow \frac{x-100}{10} = -(-1.65) \Rightarrow x = 116.5$$

نکاتی دیگر از تابع توزیع نرمال: (زنگی ـ بهنجار)

۱ـ $x = \mu$ خط تقارن منحنی و برابر میانه $\rightarrow$ Med و مد $\rightarrow$ Mod است.

$$p(x < \mu) = p(x > \mu) = \frac{1}{2}$$



۲ـ

$$f(x) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} ; \begin{cases} F(+\infty) = 1 & F(-\infty) = 0 \\ f(+\infty) = 0 & f(-\infty) = 0 \end{cases}$$

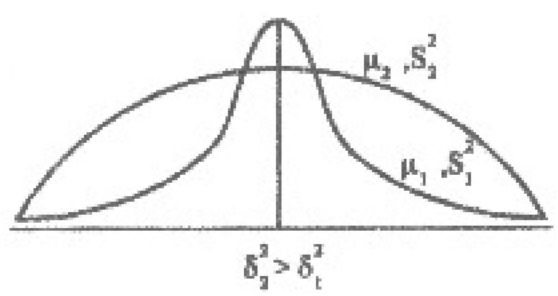
تجمع به سمت چپ

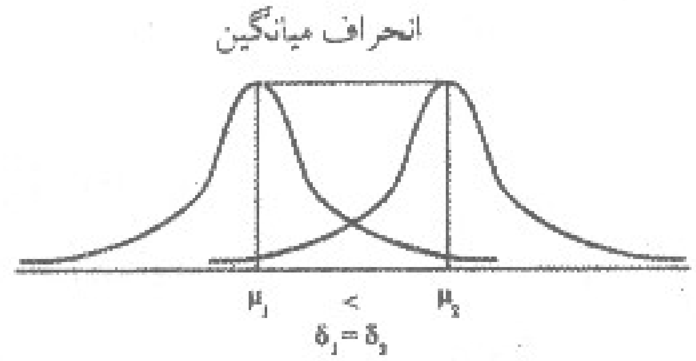
$-\infty < x < +\infty \qquad F(\mu)=\frac{1}{2}$

$-\infty < \mu < \infty \qquad f(\mu)=\frac{1}{\sqrt{2\pi}\,\delta}$

$\delta > 0 \Rightarrow \delta^2 > 0$

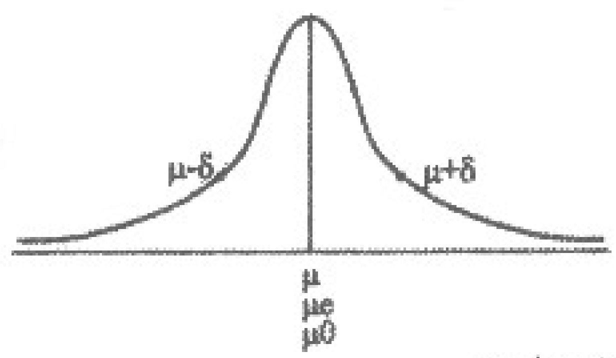
۳ـ عرض منحنی نسبت به میانگین δ . در صورت افزایش انحراف معیار ارتفاع و تمرکز کاهش یافته و پراکندگی افزایش می‌یابد.





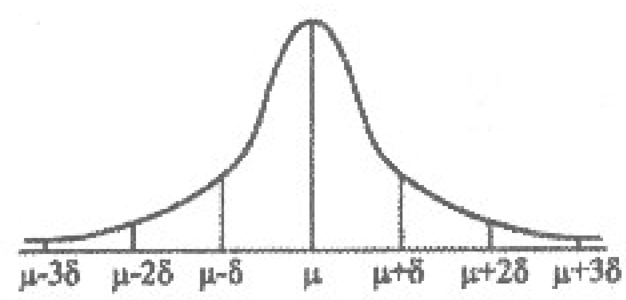
۴ـ دو نقطهٔ عطف $x=\mu\pm\delta$ می‌باشند.

$f''(x)=0$



۵ـ $f'(\mu)=0$ چون μ در توزیع نرمال همان مد است.

۶ـ فقط در توزیع نرمال داریم:



$p(|x-\mu|<\delta)=0.683$

$p(|x-\mu|<2\delta)=0.9544$

$p(|x-\mu|<3\delta)=0.997$

۷ـ در دو طرف منحنی احتمال به سمت خارج از 3 انحراف بیشتر، تقریباً صفر و به سمت داخل تقریباً 1 است

مثال:

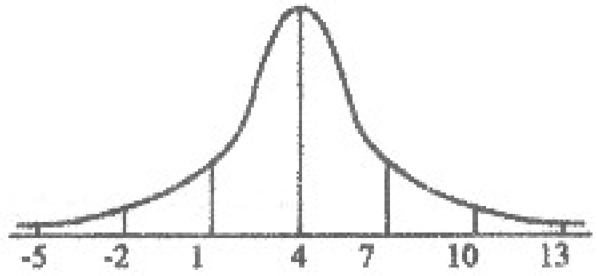
الف) $p(x\ge 4)=\frac{1}{2}$

ب) $p(x=4)=0$

ج) $p(0<x<4)=\frac{1}{2}$

د) $p(x>15)=0$

هـ) $p(x<15)=1$



۸ـ

$x \sim N(\mu, \delta^2)$

تغییر متغیر

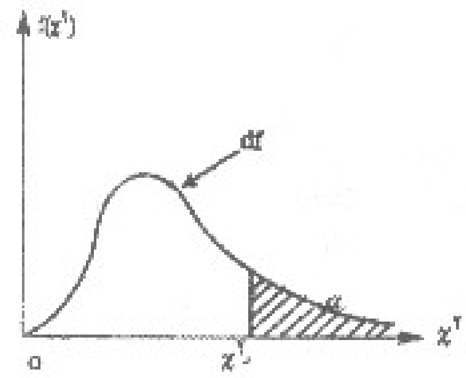
نرمال معمولی

$$f(x) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} \quad \boxed{z = \frac{x-\mu}{\delta}}$$

نرمال استاندارد

$$z \sim N(0,1);\ f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

۹ـ



$$P(y < \alpha) = \int_{-\infty}^{\alpha} f(Z)\,dZ = \varphi(\alpha) = N(\alpha)$$

$$P(Z > \alpha) = 1 - p(Z < \alpha) = 1 - \varphi(\alpha)$$

$$P(x < \alpha) = p\left(\frac{x-\mu}{\delta} < \frac{\alpha-\mu}{\delta}\right) = \varphi\left(\frac{\alpha-\mu}{\delta}\right)$$

توزیع‌های نتیجه‌گیری شده از نرمال:

ـ توزیع χ^2 یا کای اسکور: (خی دو – کای دو)

اگر در توزیع نرمال $N(\mu, \delta^2)$ باشد آنگاه:

$$\left(\frac{x-\mu}{\delta}\right)^2 = z^2$$

دارای توزیع کای اسکور با ۱ درجه آزادی است.

نکته : اگر $x_i \sim N(\mu_i, \delta_i^2)$ آنگاه:

$$\sum_{i=1}^{n} \left(\frac{x_i - M_i}{\delta_i}\right)^2 = \sum_{i=1}^{n} z_i^2$$

دارای توزیع کای اسکور با n درجه آزادی است .

نکته: در توزیع کای اسکور n درجه ,2n و $E(x) = n$ است اگر n درجه آزادی باشد. یعنی :

$$E(x) = n;\ Var(x) = 2n$$

نکته : اگر $i = 1,2,\ldots\ldots,n, x_i$ دارای توزیع نرمال و مستقل باشند. آنگاه :

$$N(\mu, \delta^2)$$

$$\sum_{i=1}^{n} x_i \sim N(n\mu, n\delta^2);$$

$$\sum_{i=1}^{n} z_i \sim N(0, n)$$

نکته : اگر z_i ها دارای توزیع نرمال استاندارد باشند و همچنین مستقل باشند.

نکته: اگر $x_i \sim N(\mu_i, \delta_i^2)$ و مستقل باشند آنگاه:

$$\sum_{i=1}^{n} x_i \approx N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \delta_i^2\right)$$

ـ توزیع t (استیودنت) :

اگر z نرمال استاندارد و V دارای توزیع x^2 با n درجه آزادی باشد آنگاه T دارای توزیع t استیودنت با n درجه آزادی است.

$$T = \frac{Z}{\sqrt{\frac{V}{n}}}$$

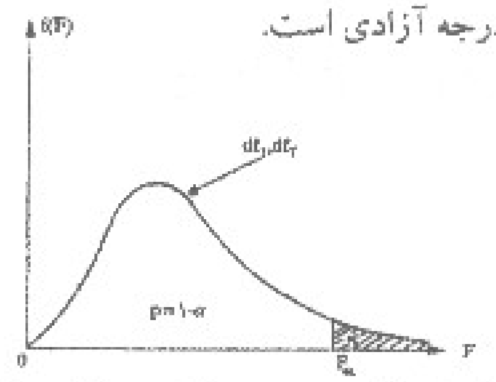
$$\mu = E(x) = 0 \;; \delta^2 = \frac{n}{n-2}$$

- توزیع F (فیشر)

اگر u و V هر دو دارای توزیع کای اسکور با m و n درجه آزادی باشند آنگاه :

$$F_{m,n} = \frac{u/m}{v/n}$$

دارای توزیع فیشر F با m و n درجه آزادی است.



1) $F_{m,n} \neq F_{n,m}$

2) $\mu = E(x) = \frac{n}{n-2}$

مثال: اگر X دارای توزیع نرمال با میانگین صفر و واریانس 1 باشد آنگاه $y = x^2$ دارای توزیع

۱) نرمال است
۲) χ^2 با یک درجه آزادی است
۳) T با یک درجه آزادی است
۴) χ^2 با $\frac{1}{2}$ درجه آزادی است

گزینه ۲ صحیح می‌باشد

$$t_{(n)} = \frac{z}{\sqrt{\frac{\chi^2_{(n)}}{n}}}$$

$$t_{(2)} = \frac{z}{\sqrt{\frac{z_1^2 + z_2^2}{2}}} \qquad t_{(1)} = \frac{z}{\sqrt{\frac{z^2}{1}}} = 1$$

$$F_{m,n} = \frac{\frac{\chi^2_{(m)}}{m}}{\frac{\chi^2_{(n)}}{n}} \qquad F_{m,n} \neq F_{n,m}$$

$$F_{1,n} = \frac{\frac{\chi^2_{(1)}}{1}}{\frac{\chi^2_{(n)}}{n}} = \frac{\frac{z^2}{1}}{\frac{\chi^2_{(n)}}{n}} = \left(\frac{z}{\sqrt{\frac{\chi^2_{(n)}}{n}}}\right)^2 = t_n^2$$

برآورد آماری

بدلیل اینکه آزمایش و اندازه گیری روی تمام عنصرهای یک جامعه مقدور نیست، بنابراین معمولا داده های آماری بطور تصادفی به وسیله نمونه گیری بدست می آید. استنباط یا استنتاج آماری عبارتست از تعمیم نتایج حاصله از نمونه روی کل جامعه.

بنابراین آمار بر دو بخش آمار توصیفی و آمار استنتاجی تقسیم بندی می شود. اختلاف اساسی بین آمار توصیفی و آمار استنتاجی در آن است که در آمار توصیفی اطلاعات جمع آوری شده خود نتیجه هستند در حالی که همین اطلاعات در آمار استنتاجی به عنوان وسیله ای در مرحله تحقیق به کار می روند.

استنباط آماری به دو روش برآورد کردن و آزمون فرضیه صورت می گیرد.

نظریه برآورد آماری

ویژگیها و معیارهای هر جامعهٔ آماری مانند:میانگین ، چارکها، انحراف معیار،واریانس و... را پارامتر می گویند. این پارامترها با استفاده از اندازه گیریهای حاصل از نمونه گیری محاسبه می شوند بنابراین اندازه پارامترها را که با استفاده از اندازه گیری بدست می آیند، برآورد آماری گویند.

برآورد آماری به دو نوع نقطه ای و برآورد فاصله ای صورت می گیرد.

برآورد نقطه ای

برآورد نقطه ای عبارتست از محاسبه مقدار مشخصی از پارامتر در نمونه که معرف بهترین حدس در مورد پارامتر جامعه باشد. برای مثال اگر برآورد میانگین سن جامعه دانشجویان عمران با استفاده از نمونه تصادفی دانشجویان کلاس استاتیک میانگین نمونه را ۲۰ محاسبه نماییم، عدد ۲۰(میانگین نمونه) برآورد نقطه ای میانگین جامعه دانشجویان این دانشکده خواهد بود.

برآورد فاصله ای:

در برآورد فاصله‌ای دو مقدار مشخص از پارامتر چنان محاسبه می گردد که با ضریب اطمینان خاصی مقدار واقعی پارامتر جامعه را در بر داشته باشد. در مثال یاد شده می توان برآورد فاصله ای زیر را عنوان کرد:

میانگین سن دانشجویان عمران با احتمال ۹۵٪ بین ۱۵ و ۲۵ است. شیوه محاسبه برآورد فاصله ای که همان نحوه ساختن فاصله اطمینان است . متعاقبا مورد بحث قرار می گیرد.

مهمترین خواص برآورد کننده ها عبارتند از:

الف - بدون تورش یا نا اریب بودن(UNBIASEDNESS):

هر تابعی از متغیرهای نمونه یعنی $f(x_1,x_2,\ldots,x_n)$ را نمی توان به عنوان برآورده کننده پارامتر جامعه انتخاب کرد. بلکه باید تابعی از متغیرهای نمونه را انتخاب نمود که مقدار عددی آن به مقدار حقیقی پارامتر جامعه θ نزدیک باشد. تخمین زن یعنی $\hat{\theta}$ باید خصوصیاتی داشته باشد که یکی از آنها نااریب بودن است.

برآورده کننده نااریب، برآورد کننده ای است که به طور متوسط مقدار واقعی پارامتر جامعه را نشان دهد. به عبارت دیگر اگر نمونه متعددی انتخاب نموده، با استفاده از هر یک از این نمونه ها برآوردی از پارامتر واقعی محاسبه کنیم. متوسط این برآوردها برابر مقدار واقعی پارامتر جامعه گردد. یعنی:

$$E(\hat{\theta}) = \theta$$

در غیر این صورت اگر $E(\hat{\theta}) \neq \theta$ باشد، $\hat{\theta}$ را برآورد کننده با اریب می گویند.

مثلاً آماره $\bar{X}$ یعنی میانگین نمونه یک برآورده کننده نااریب برای پارامتر μ یعنی میانگین جامعه می باشد چون:

$$E(\bar{X}) = \mu$$

همچنین آماره $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ یعنی واریانس نمونه یک برآورد کننده نااریب برای پارامتر σ^2 یعنی واریانس جامعه می باشد چون:

$$E(S^2) = \sigma^2$$

ب - کارائی:

هر چقدر واریانس برآورد کننده کوچک باشد آن را کاراتر گویند. میانگین نمونه و میانه نمونه هر دو برآورد کننده نااریب برای میانگین جامعه هستند ولی واریانس میانگین نمونه کوچکتر از واریانس میانه نمونه است پس کاراتر است .در واقع می توان از میان ۲ برآورد کننده نااریب یکی را بعنوان بهترین انتخاب نمود.

$\hat{\theta}_1$ کاراتر از $\hat{\theta}_2 \Rightarrow \mathrm{var}(\hat{\theta}_1) < \mathrm{var}(\hat{\theta}_2)$

ج - سازگاری یا پایداری

اگر با افزایش حجم نمونه، واریانس برآورد کننده به صفر گرایش یابد به آن برآورد کننده سازگار می گویند ، یعنی:

$$\lim_{n \to \infty} \mathrm{Var}(\hat{\sigma}) = 0$$

د - کفایت:

برآورد کننده ای کافی تلقی می شود که تمام اطلاعات نمونه را مورد استفاده قرار دهد. مثلاًمیانه که فقط از طریق مرتب کردن مشاهدات محاسبه می گردد و نه با استفاده از مقادیر که ناقل اطلاعات می باشند، دارای خاصیت کفایت نمی باشد.

برآورد فاصله‌ای

برآوردهای نقطه‌ای در عین حال که بسیار مهم‌اند و ویژگی‌های خوبی دارند دارای این مشکل هستند که نه برابر پارامتر هستند و نه میزان نزدیکی آن‌ها به پارامتر قابل اندازه‌گیری است و حتی کران خطای آن‌ها در دست نیست. برای رفع این مشکل برآوردهای فاصله‌ای می‌توانند با میزان اطمینان معین فاصله‌ای با طول معین ارائه کنند که پارامتر داخل آن قرار دارد.

تعریف: فاصله تصادفی $(\hat{\theta}_1, \hat{\theta}_2)$ را برآورد فاصله‌ای سطح اطمینان $100(1-\alpha)$% برای θ گویند اگر $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1-\alpha$

$\hat{\theta}_1$ و $\hat{\theta}_2$ برآوردکننده‌های نقطه‌ای θ هستند.

اگر $P(\hat{\theta} < \theta) = 1-\alpha$ گویند فاصله $(-\infty , \hat{\theta})$ برآورد فاصله‌ای یک طرفه برای θ در سطح اطمینان $100\ (1-\alpha)\ \%$ می‌باشد.

نکته: فاصله اطمینان‌های مهم معمولاً در سطح 99% , 95% , 90% است و مقدار Z برای آن‌ها به شرح زیر می‌باشد:

$\%90 \Rightarrow Z_{0.05} = 1.645$

$\%95 \Rightarrow Z_{0.025} = 1.96$

$\%99 \Rightarrow Z_{0.005} = 2.58$

انواع برآورد فاصله‌ای برای میانگین

برآورد فاصله‌ای برای μ در جامعه نرمال با واریانس معلوم

می‌دانیم:

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

پس برآورد فاصله‌ای در سطح $100\ (1-\alpha)\ \%$ برای μ می‌شود.

$$\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \quad و \quad \left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \quad و \quad \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$$

در شرایطی که σ یعنی انحراف معیار جامعه نامعلوم باشد شرایطی پیش می‌آید:

الف) اگر $n > 30$ باشد بازهم از رابطه بالا استفاده می‌کنیم. یعنی:

$$\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \quad و \quad \left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \quad و \quad \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$$

ب) اگر $n \le 30$ باشد آن‌گاه از توزیع t با $n-1$ درجه آزادی استفاده می‌کنیم. و برآورد فاصله‌ای به صورت زیر خواهد بود:

$$\bar{X} \pm \frac{\sigma}{\sqrt{n}} t_{n-1, \frac{\alpha}{2}} \quad و \quad \left(\bar{X} - \frac{\sigma}{\sqrt{n}} t_{n-1, \frac{\alpha}{2}} \quad و \quad \bar{X} + \frac{\sigma}{\sqrt{n}} t_{n-1, \frac{\alpha}{2}}\right)$$

مثال: اگر طول عمر لامپ‌های کارخانه‌ای توزیع نرمال با انحراف معیار 80 ساعت داشته باشد و میانگین طول عمر 81 لامپ این کارخانه 1150 ساعت باشد برآورد فاصله‌ای سطح 95 % اطمینان برای μ (متوسط طول عمر همه لامپ‌های کارخانه) بسازید.

$$\bar{X} \pm \frac{6}{\sqrt{n}} z_{\frac{\alpha}{2}} \qquad 1150 \pm \frac{80}{9} 1.96$$

و یا (1167 و 1133).

برآورد فاصله‌ای برای p (نسبت یا درصد خاصی در جامعه)

در صورتیکه $\hat{p}=\frac{m}{n}$ نسبت خاصی از یک نمونه nتائی باشد که m تعداد حالاتی از نمونه است که خاصیت مورد نظر را دارد. در این حالت برای برآورد فاصلهٔ p که نسبت جامعه است از فاصله اطمینان زیر استفاده می‌کنیم.

برآورد فاصله‌ای مذکور می‌شود

$$\hat{p} \pm \sqrt{\frac{\hat{p}\hat{q}}{n}}\; z_{\frac{\alpha}{2}}$$

مثال: فرض کنید در یک نمونه 400 نفری از ساکنین مرد بالغ یک شهر 110 نفر سیگاری هستند برآورد فاصله‌ای سطح 98 % اطمینان برای نسبت واقعی مردان سیگاری شهر بسازید.

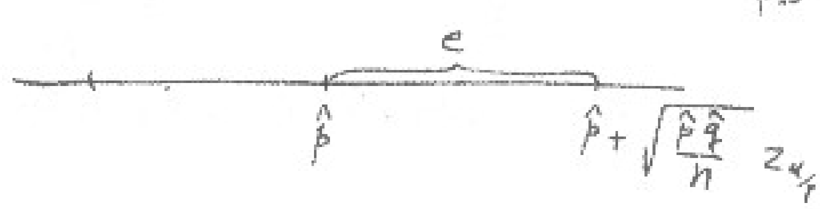
$$\hat{p}=\frac{110}{400}=0.275 \quad , \quad z_{0.01}=2.33$$

$$0.275 \pm \sqrt{\frac{(0.275)(0.725)}{400}}\; 2.33$$

و یا (0.327 و 0.223)

برآورد حجم نمونه

می‌دانیم در برآورد فاصله‌ای برای p داریم



در واقع e که شعاع فاصله است حداکثر خطای ناشی از قرار دادن $\hat{p}$ به جای p در براورد فاصله‌ای سطح 100 $(1-\alpha)$ % است پس با داشتن e می‌توان تعداد نمونه لازم برای به وجود آمدن این شرایط را بدست آورد.

با توجه به آن که

$$e=\sqrt{\frac{\hat{p}\hat{q}}{n}}\; z_{\frac{\alpha}{2}}$$

پس

$$n=\hat{p}\,\hat{q}\left(\frac{z_{\frac{\alpha}{2}}}{e}\right)^2$$

تعداد نمونه لازم برای آن است که در سطح اطمینان 100$(1-\alpha)$% حداکثر خطای ناشی از قرار دادن $\hat{p}$ به جای p برابر e بشود. این فرمول در صورتی استفاده می‌شود که $\hat{p}$ و $\hat{q}$ را از یک نمونه مقدماتی بدست آورده باشیم. در صورتی که نمونه مقدماتی نخواهیم بر اساس

آن که $\hat{p}+\hat{q}=1$ می‌دانیم $\hat{p}\hat{q}\le\frac{1}{4}$ پس $n=\frac{1}{4}\left(\frac{z_{\frac{\alpha}{2}}}{e}\right)^2$ تعداد نمونه لازم برای آن است که در سطح اطمینان $100(1-\alpha)\%$ حداکثر خطای ناشی از قرار دادن $\hat{p}$ به جای p برابر e بشود بدون آن که نمونه مقدماتی بگیریم.

مثال: فرض کنید در یک نمونه 200 نفری از دانشجویان دانشگاهی 18 نفر مشروط بوده‌اند اگر بخواهیم با حداکثر خطا 0.02 در سطح اطمینان 98 % نسبت مشروطی‌ها را برآورد کنیم به دو روش تعداد نمونه لازم را بدست آورید.

الف) با نمونه مقدماتی

$$\hat{p}=\frac{18}{200}=0.09 \quad\rightarrow\quad \hat{q}=0.91$$

$$n_1=(0.09)(0.91)\left(\frac{2.33}{0.02}\right)^2=1111.5 \quad\rightarrow\quad n_1=1112$$

ب) بدون نمونه مقدماتی

$$n_2=\frac{1}{4}\left(\frac{2.33}{0.02}\right)^2=3393.06 \qquad\qquad n_2=3394$$

همانطور که ملاحظه می‌گردد انتخاب نمونه مقدماتی مقرون به صرفه است.

در برآورد فاصله‌ای μ داریم

$$\bar{X}\pm\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}$$

$$\overbrace{\bar{X} \qquad\qquad \bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}}^{e}$$

پس تعداد نمونه لازم برای آن که حداکثر خطای ناشی از قرار دادن $\bar{X}$ به جای μ در سطح اطمینان $100\,(1-\alpha)\,\%$ برابر e بشود برابر است با:

$$n=\left(\frac{\sigma}{e}z_{\frac{\alpha}{2}}\right)^2$$

مثال: فرض کنید انحراف معیار طول عمر یک قطعه 10 ساعت است اگر بخواهیم در سطح اطمینان 95 % حداکثر خطای ناشی از قرار دادن $\bar{X}$ بجای μ برابر 1 ساعت بشود چه تعداد نمونه لازم داریم.

حل: $n=\left(\frac{10}{1}1.96\right)^2=384.16$ پس تعداد 385 نمونه لازم داریم.

زیرا $z_{0.025}=1.96$

آزمون فرضیه

کلیات

در بسیاری از موارد باید **تصمیم** گیری براساس اطلاعات حاصل از یک نمونه انجام شود. **مثلاً مدیر** کنترل کیفیت کارخانه باید تعیین کند که تولید مطابق استاندارد می باشد یا خیر؟

برای این منظور شخص تصمیم گیرنده برای پارامتر مورد نظر عددی را فرض می کند و آنرا H_0 (فرضیه صفر)می نامد. فرضیه مقابل آن را H_1 می نامد. اکنون با انتخاب یک نمونه تصادفی از جامعه و اعمال روش های آماری فرضیه را، با اطمینان معین p(در سطح $\alpha=1-P$) آزمون می کند. اگر دلیلی بر رد H_0 پیدا نشود آنرا می پذیرد و در غیر اینصورت اگر H_0 رد شود فرضیه مقابل یعنی H_1 پذیرفته می شود.

خطای نوع اول و خطای نوع دوم

در هر آزمون و تصمیم گیری دو نوع خطا وجود دارد.

الف – خطای نوع اول – احتمال خطای نوع اول(α) عبارتست از احتمال رد کردن فرضیه صفر در حالیکه واقعاً فرضیه H_0 درست باشد.

ب – خطای نوع دوم – احتمال خطای نوع دوم (β) عبارتست از احتمال پذیرفتن فرضیه H_0 در حالی که واقعاً فرضیه H_1 درست باشد.

با کم کردن مقدار یکی از دو خطا مقدارخطای دیگر زیاد می شود و تنها راه برای کاهش هر دو نوع خطا افزایش حجم نمونه می‌باشد.

آزمونهای یک طرفه و دو طرفه.

در انتخاب فرضیه های آماری، برای فرضیه H_0 باید عبارتی را در نظر گرفت که شامل قسمت مساوی باشد. معمولاً فرضیه‌های آماری بصورت های زیر انتخاب می شوند.

الف) $H_0:\theta=\theta_0, H_1:\theta\neq\theta_0$

ب) $H_0:\theta\leq\theta_0, H_1:\theta>\theta_0$

ج) $H_0:\theta\geq\theta_0, H_1:\theta<\theta_0$

به آزمونهائی که فرضیه‌های آن بصورت (الف) باشد آزمون های دو طرفه و به آزمونهائی که فرضیه‌های آن بصورت (ب)و (ج) باشد آزمونهای یک طرفه می گویند.

مراحل مختلف یک آزمون

برای انجام یک آزمون فرضیه، مراحل زیر باید صورت گیرد.

۱– تشکیل فرضیه صفر H_0 و فرضیه مقابل آن H_1که معمولاً به یکی از صورتهای مذکور در بند قبل میباشد.

۲– تعیین مقدار α یعنی احتمال خطای نوع اول که به آن سطح تشخیص یا میزان معنی دار بودن می گویند.

۳– انتخاب آماره آزمون برای پارامتر θ

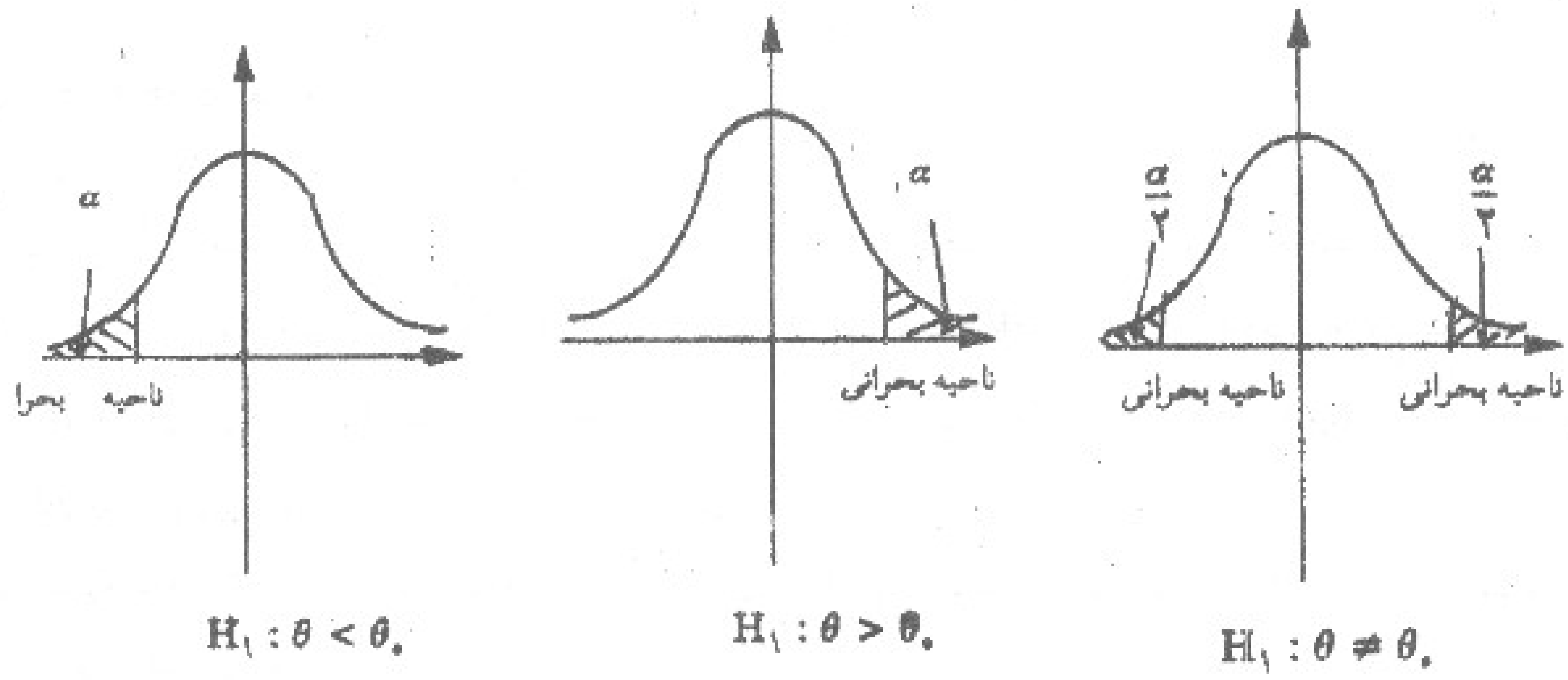
در اغلب موارد بطور کلی‌آماره آزمون بصورت $z=\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}$ یا $t=\frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}$ می باشد که در آن θ و $\hat{\sigma}$ انحراف معیار آماره $\hat{\theta}$ در تمام نمونه‌های ممکن به حجم nمی باشد.

با رعایت حالات خاص بطور کلی z برای نمونه های بزرگ (n>30) و t برای نمونه های کوچک($n\leq 30$) بکار می رود.

۴– به کمک اطلاعات بدست آمده از نمونه تصادفی مقدار آماره آزمون طبق فرمولهای بند ۳ محاسبه می گردد.

۵– نقطه بحرانی با در نظر گرفتن مقدار α و فرضیه H_1 از جدول مربوطه استخراج می گردد و منطقه بحرانی در زیر منحنی توزیع مربوطه مشخص می گردد.

اگر آزمون دو طرفه باشد یعنی $H_1:\theta \neq \theta_0$ منطقه بحرانی در دو طرف منحنی توزیع(در هر طرف $\frac{\alpha}{2}$) و اگر آزمون یک طرفه باشد برای $H_1:\theta>\theta_0$منطقه بحرانی(α) سمت راست و برای $H_1:\theta<\theta_0$ منطقه بحرانی(α) در سمت چپ منحنی توزیع انتخاب می شود.



۶- اگر آماره آزمون در منطقه بحرانی قرار بگیرد فرضیه H_0 با احتمال $P=1-\alpha$ رد می شود و فرضیه مقابل یعنی H_1 پذیرفته می شود. در غیر اینصورت فرضیه H_0رد نمی شود.

تبصره: معمولاً مقدار α را ۰/۱۰ یا ۰/۰۵ یا ۰/۰۱ و یا ۰/۰۰۱ در نظر می گیرند. اگر مقدار α را مشخص نکنند. معمولاً آن را ۰/۰۵ اختیار می کنند.

آزمون میانگین جامعه

فرض می کنیم در جامعه ای میانگین یعنی μ مجهول باشد. می خواهیم آزمون نمائیم که آیا می توان مقدار مفروض μ_0 را با احتمال p میانگین این جامعه دانست یا خیر؟

فرضیه های آماری را بصورت زیر در نظر می گیریم:

$$H_0:\mu=\mu_0$$
$$H_1:\mu\neq\mu_0$$

ابتدا نمونه ای تصادفی به حجم n انتخاب می کنیم و میانگین و واریانس آن را بشرح زیر حساب می کنیم.

$$\bar{x}=\frac{\sum x_i}{n} \quad , \quad s^2=\frac{\sum (x_i-\bar{x})}{n-1}$$

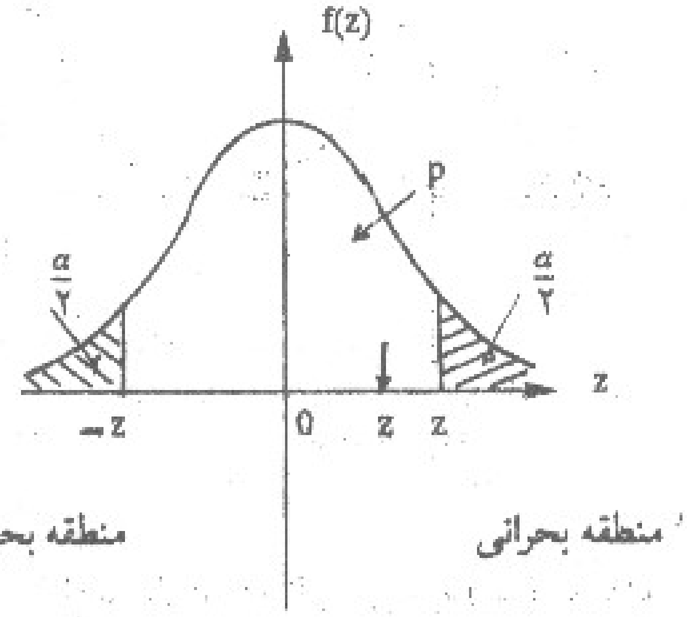
اکنون دو حالت در نظر می گیریم.

حجم نمونه بزرگ است n>30

ابتدا آماره آزمون را بصورت $z=\frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}}$ تشکیل می دهیم. سپس با توجه به احتمال p مقدار z بحرانی را از جدول توزیع نرمال استاندارد بشرح زیر حساب می کنیم و منطقه بحرانی را مشخص می کنیم.

$$S_{-z}^{z} = p \Rightarrow S_0^z = \frac{p}{2} \Rightarrow Z_{\frac{\alpha}{2}} = a$$

a عدد است.



تصمیم گیری – اگر آماره آزمون در منطقه بحرانی قرار بگیرد فرضیه H_0 رد می شود و فرضیه H_1 را می پذیریم یعنی با احتمال p مقدار μ_0 را نمی توان به عنوان میانگین جامعه پذیرفت. اگر آماره آزمون در منطقه بحرانی قرار نگیرد فرضیه H_0 رد نمی شود یعنی در نمونه دلیلی بر رد عدد μ_0 به عنوان میانگین جامعه وجود ندارد و آن را می توان پذیرفت.

تبصره: هنگامی که حجم نمونه بزرگ است نرمال بودن جامعه اصلی ضرورتی ندارد. اگر انحراف معیار جامعه معلوم باشد ترجیحاً از فرمول $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ استفاده می گردد ولی اگر انحراف معیار جامعه (σ) معلوم نباشد از برآورد کننده آن S یعنی انحراف معیار نمونه استفاده می گردد.

مثال: سازنده نوعی محصولات داروئی روی بسته های دارو ذکر کرده است که این بسته حاوی ۱۶ میلی گرم ویتامین C می باشد برای تحقیق در صحت این ادعا یک نمونه تصادفی شامل ۴۹ بسته انتخاب و آزمایش مقادیر $\bar{x} = 15.82$ و $S^2=0.49$ را نشان داده است صحت این ادعا را با احتمال ۹۰٪ آزمون نمائید.

فرضیه های آماری عبارتند از:

$H_0: \mu=16$
$H_1: \mu \neq 16$

آماره آزمون را تشکیل می دهیم.$(n=49>30)$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{15.82 - 16}{\frac{0.7}{\sqrt{49}}} = -1.8$$

مقدار بحرانی را از جدول نرمال استاندارد با توجه به احتمال $P=0.90$ و فرضیه H_1(دو طرفه)حساب می کنیم.

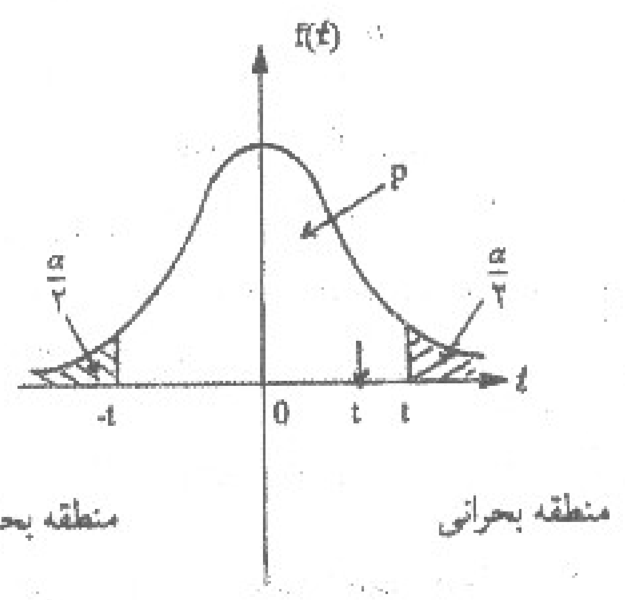
$$S_{-z}^{z} = 0.9 \Rightarrow S_0^z = 0.45 \Rightarrow Z_{0.05} = 1.645$$

منطقه بحرانی $z>1.645$ و $z<-1.645$ می باشد.

تصمیم گیری – چون آزمون در منطقه بحرانی قرار می گیرد$(-1.8<-1.645)$ بنابراین فرضیه H_0 با احتمال 90% رد می شود و ادعا را نمی توان پذیرفت.

حجم نمونه کوچک است $30 \le n$

در این حالت شرط نرمال بودن جامعه الزامی است. ابتدا آمار آزمون را بصورت $t = \dfrac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}}$ تشکیل می دهیم. سپس با توجه به احتمال p مقدار بحرانی t را از جدول توزیع t استیودنت بشرح زیر حساب می کنیم و منطقه بحرانی را مشخص می کنیم.



$$\begin{cases} df = n-1 \\ \alpha = 1-p \end{cases} \Rightarrow t_{\frac{\alpha}{2}} = a$$

a عدد است.

تصمیم گیری - اگر آماره آزمون در منطقه بحرانی قرار بگیرد فرضیه H_0 رد می شود در غیر این صورت فرضیه H_0 رد نمی شود.

مثال: ادعا شده است که میانگین جامعه ای نرمال برابر ۳/۵ می باشد. برای تحقیق در صحت این ادعا نمونه ای به حجم n=5 انتخاب گردید و مقادیر مشاهده شده عبارتند از: x_i:4,3,2,3,3 فرضیه را در سطح تشخیص α=0.05 آزمون نمائید.

فرضیه های آماری را تشکیل می دهیم.

$H_0: \mu = 3.5$

$H_1: \mu \neq 3.5$

بر اساس داده های نمونه $\overline{x}$ و s را محاسبه می کنیم:

X_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
4	1	1
3	0	0
2	-1	1
3	0	0
3	0	0
$\sum 15$	0	2

$$\overline{x} = \frac{\sum x_i}{n} = \frac{15}{5} = 3$$

$$S^2 = \frac{\sum (x_i - \overline{x})^2}{n-1} = \frac{2}{5-1} = 0.5$$

$$S = \sqrt{0.5} = 0.7$$

اکنون آماره آزمون را تشکیل می دهیم.

$$z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{3-3.5}{\frac{0.7}{\sqrt{5}}}=-1.6$$

مقدار بحرانی را از جدول t استیودنت بشرح زیر حساب می کنیم.

$$\begin{cases} df=n-1=5-1=4 \\ \alpha=0.05 \end{cases} \Rightarrow t=2.776$$

منطقه بحرانی $t>2.776$ و $t<-2.776$ می باشد.

تصمیم گیری – تحلیل این آزمون این است که چون آماره آزمون در منطقه بحرانی قرار نمی گیرد($-2.776<-1.6<2.776$) بنابراین فرضیه H_0 با احتمال ۹۵٪ رد نمی شود و می توان عدد ۳/۵ را به عنوان میانگین جامعه پذیرفت.

تبصره ۱: هنگامی که حجم نمونه کوچک است نرمال بودن جامعه الزامی است به آزمون هائی که نرمال بودن جامعه ضرورت دارد آزمونهای پارامتریک می گویند و اگر نرمال بودن جامعه ضرورتی نداشته باشد به آن آزمون ناپارامتریک می گویند.

تبصره ۲: اگر حجم نمونه کوچک و توزیع جامعه اصلی نرمال و انحراف معیار جامعه یعنی σ معلوم باشد علیرغم کوچک بودن حجم نمونه باید از متغیر z استفاده نمود و آماره آزمون بصورت $z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ خواهد بود.

تبصره ۳: چون جداول توزیع t به صورتهای یک طرفه و دو طرفه وجود دارد(یعنی α تماماً در سمت راست منحنی یا در دو طرف منحنی در هر طرف $\frac{\alpha}{2}$ قرار دارد) در هنگام استفاده از جدول باین مطلب باید توجه نمود.

تبصره ٤: اگر فرضیه مقابل بصورت $\mu\neq\mu_0$ یا $\mu>\mu_0$ یا $\mu<\mu_0$ باشد منطقه بحرانی را باید در دو طرف و یا تماماً در سمت راست و یا تماماً در سمت چپ منحنی توزیع t در نظر گرفت.

مثال: در صورتی که $Z=1.96$ جدول و $-3=(Z$ محاسبه شده) باشد چه قضاوتی می‌توان درباره H_0 و H_1 داشت؟

۱) H_0 در ناحیه بحرانی قرار نمی‌گیرد.　　۲) H_0 قبول می‌شود.

۳) H_0 رد می‌شود.　　۴) H_1 رد می‌شود.

چون در حالت پیش‌فرض آزمون را دوطرفه فرض می‌کنیم بنابراین نواحی بحرانی برای z جدول $z>1.96$ و $z<-1.96$ خواهد بود z محاسبه شده یعنی 3 - در ناحیه بحرانی قرار می‌گیرد و فرض H_0 در می‌شود و H_1 تأیید می‌شود.

مجموعه سوالات کنکوری

۱ـ در دو نمونه 10تائی و 30تائی، میانگین‌ها به ترتیب برابر 1 و 5 می‌باشد. میانگین کل برای نمونه 40تایی چند است؟

۱) 2.5 ۲) 3 ۳) 3.5 ۴) 4

۲ـ انحراف استاندارد داده‌های 2 , 5 , 4 , 1 , 4 , 2 کدام است؟

۱) $\sqrt{2}$ ۲) $\sqrt{3}$ ۳) 2 ۴) $\sqrt{\frac{3}{2}}$

۳ـ کدامیک از تعاریف زیر، در مورد علم آمار جامع‌تر است؟

۱) شمارش و اندازه‌گیری

۲) جمع‌آوری اطلاعات عددی

۳) جمع‌آوری، تلخیص، رده‌بندی و ارزیابی نتایج

۴) جمع‌آوری اعداد و ارقام

۴ـ در یک جدول توزیع فراوانی مجموع فراوانی ها برابر 25 می‌باشد. در صورتی که فراوانی مطلق طبقه آخر برابر 4 باشد، درصد فراوانی تجمعی طبقه ماقبل آخر چیست؟

۱) 21 ۲) 84 ۳) %21 ۴) $\frac{21}{25}$

۵ـ اگر انحراف معیار داده‌های آماری $X_1 , X_2 , \dots , X_n$، برابر 4 باشد، واریانس داده‌های $(2X_1+1), (2X_2+1), \dots, (2X_n+1)$ کدام است؟

۱) 8 ۲) 16 ۳) 32 ۴) 64

۶ـ از سه شیمیدان و چهار فیزیکدان، چند کمیته از یک شیمیدان و دو فیزیکدان می‌توان تشکیل داد؟

۱) 6 ۲) 9 ۳) 18 ۴) 36

۷ـ اگر $P(A) = 0.59$ و $P(B) = 0.3$ و $P(A \cap B) = 0.21$، آن‌گاه مقدار $P(A \cap \bar{B})$ کدام است؟ ($\bar{B}$ متمم B می‌باشد.)

۱) 0.56 ۲) 0.38 ۳) 0.28 ۴) 0.18

۸ـ کیسه‌ای شامل 5 مهره سفید، 2 مهره سیاه و 3 مهره قرمز است. یک مهره به تصادف از کیسه خارج می‌کنیم و می‌دانیم که سیاه نیست، احتمال این که آن مهره سفید باشد، کدام است؟

۱) $\frac{1}{2}$ ۲) $\frac{4}{9}$ ۳) $\frac{5}{8}$ ۴) $\frac{5}{9}$

۹ـ اگر X یک متغیر تصادفی پیوسته با توزیع $f(x) = \begin{cases} kx^2 + x & 0<x<1 \\ 0 & x \notin (0,1) \end{cases}$ باشد، مقدار k کدام است؟

۱) $\frac{3}{2}$ ۲) $\frac{2}{3}$ ۳) $\frac{20}{3}$ ۴) $\frac{30}{2}$

۱۰ ـ میانگین یک توزیع دو جمله ای $\mu = 8$ و $n = 25$ می‌باشد، مقدار p کدام است؟

۱) 0.08 ۲) 0.16 ۳) 0.32 ✓ ۴) 0.64

۱۱ ـ در یک برآورد کننده، n را به سمت بی‌نهایت ببریم، اریبی و واریانس به سمت صفر میل کرد، برآورد کننده مورد نظر کدام است؟

۱) سازگار ✓ ۲) کارآ

۳) ناسازگار ۴) نا اریب

۱۲ ـ در یک دانشکده، 50% دانشجویان فوتبال، 40% بسکتبال، 30% هم فوتبال و هم بسکتبال بازی می‌کنند. احتمال آن‌که دانشجویی در این دانشکده ورزش نکند، کدام است؟

۱) 0 ۲) 0.1 ۳) 0.4 ۴) 0.6

۱۳ ـ تاسی را 144 بار می‌ریزیم. واریانس دفعاتی که عدد 6 بالاتر قرار گیرد، کدام است؟

۱) $\sqrt{20}$ ۲) 20 ۳) 40 ۴) 400

۱۴ ـ در جدول زیر، نماد «مد» کدام است؟

x_i	4	8	5	10	6
f_i	6	9	6	5	8

۱) 6 ۲) 8 ۳) 9 ۴) 10

۱۵ ـ اگر در توزیع نرمال $\int_0^1 = 0.3413$ باشد، مقدار $\int_{-\infty}^{-1}$ کدام است؟

۱) 0.1587 ۲) 0.3413 ۳) 0.6587 ۴) 0.6826

۱۶ ـ از بین 5 کارمند حسابداری و 4 کارمند کارگزینی به چند طریق می‌شود چهار نفر، شامل 2 کارمند حسابداری و 2 کارمند کارگزینی را برای شرکت در یک سمینار انتخاب کرد؟

۱) 60 ✓ ۲) 126 ۳) 240 ۴) 400

۱۷ ـ اگر واریانس اعداد x_1 ، x_2 ، ... ، x_n مساوی a باشد، انحراف معیار اعداد $2x_1 - 1$ ، $2x_2 - 1$ ، ... ، $2x_n - 1$ کدام است؟

۱) $\sqrt{a}$ ۲) $2\sqrt{a-1}$ ۳) $2\sqrt{a}$ ✓ ۴) $4\sqrt{a}$

۱۸ ـ اگر گشتاور مرتبه چهارم حول میانگین مساوی 162 و واریانس برابر 9 باشد، ضریب کشیدگی کدام است؟

۱) 54 ۲) 18 ۳) 2.5 ۴) 2 ✓

۱۹ ـ معدل یک دانشجو در ده درس 14.2 بوده است. اگر نمره یکی از درس‌های او که 16 بوده، حذف شود، معدل او در 9 درس بقیه کدام است؟

۱) 13.8 ۲) 13.9 ۳) 14 ✓ ۴) 14.1

۲۰ ـ اگر بخواهیم ثابت کنیم که مجموع مساحت هیستوگرام مساوی یک (واحد) می‌شود، برای رسم آن از کدام پارامتر باید استفاده کنیم؟

۱) چگالی فراوانی‌های مطلق

۲) چگالی فراوانی‌های نسبی

۳) درصد فراوانی‌های نسبی

۴) فراوانی‌های نسبی

۲۱ ـ سه فروشگاه یک شرکت به ترتیب 20%، 30% و 50% فروش آن را صورت می‌دهند و برگشت از فروش آن‌ها به ترتیب 7%، 2% و 3% محموله‌های فروش بوده؛ احتمال برگشتی بودن یکی از محموله‌ها به طور تصادفی، چقدر است؟

۱) 0.022 ۲) 0.035 ۳) 0.12 ۴) 0.126

۲۲ ـ در پرتاب دو سکه سالم رو شدن یک H دارای یک امتیاز و رو شدن دو H دارای دو امتیاز بوده؛ رو نشدن H «چند امتیاز منفی» داشته باشد، تا امید ریاضی این بازی برابر صفر گردد؟

۱) 3 ۲) 4 ۳) 5 ۴) 6

۲۳ ـ در بررسی اندازه‌های دو متغیر X و Y انحراف معیار X برابر 3، انحراف معیار Y برابر 5 و ضریب همبستگی بین X و Y برابر 0.8 شده؛ «کوواریانس» آن‌ها، چقدر است؟

۱) 6.4 ۲) 10 ۳) 12 ۴) 18.75

۲۴ ـ از فارغ‌التحصیلان کارشناسی ارشد دانشکده‌ای در یک دوره، 6 نفره شاغل و 4 نفره سر کار نرفته‌اند؛ در صورت تماس تصادفی با 7 نفر آنان، «چند درصد» احتمال دارد بتوان از 2 نفرشان دعوت به کار نمود؟

۱) 30 ۲) 50 ۳) 40 ۴) 20

۲۵ ـ در آزمون فرض‌ها؛ «خطای نوع دوم» کدام است؟

۱) رد کردن H_0 وقتی H_1 درست است.

۲) رد نکردن H_0 وقتی اعتبار H_1 نشان داده نشده است.

۳) رد کردن H_0 وقتی H_0 درست است.

۴) رد نکردن H_0 وقتی H_1 درست است.

۲۶ ـ افزایش حجم نمونه با «چه ضریبی»، موجب کاهش خطای معیار $\overline{X}$ به $\frac{1}{3}$ مقدار اولیه‌اش می‌گردد؟

۱) 3 ۲) 9 ۳) 6 ۴) 12

۲۷ ـ نمرات دو داوطلب در آزمون درس آمار 14 , 12 و برحسب واحدهای استاندارد به ترتیب 0.25 و 0.25 - شده؛ «واریانس نمرات این آزمون» چقدر است؟

۱) 4 ۲) 8 ۳) 13 ۴) 16

۲۸ ـ «هر خصوصیت عددی از توزیع جامعه» چه نامیده می‌شود؟

۱) آماره ۲) پارامتر ۳) آزمایش ۴) فرض آماری

۲۹ـ برای رسم «چه نموداری» نسبت گروه‌های مختلف یک جامعه در 360 ضرب می‌شود تا زاویهٔ قطاع سهم هر گروه مشخص گردد؟

۱) کلوچه‌ای ۲) فراوانی نسبی ۳) چند ضلعی ۴) خطی فراوانی نسبی

۳۰ـ انحراف‌های مقادیر مشاهده شده از میانگین در 6 مورد از یک نمونهٔ 7تائی به صورت اعداد 5- , 4- , 1- , 2- , 1- , 3 , 4 محاسبه شده؛ «انحراف معیار نمونه» چقدر است؟

۱) 2 ۲) 3 ۳) 4 ۴) 5

۳۱ـ در نمونه $x_1, \dots, x_{10}$ اگر $\sum_{i=1}^{10} x_i = 60$ ، $\sum_{i=1}^{10} x_i^2 = 396$ باشد، واریانس کدام است؟

۱) 2 ۲) 3 ۳) 4 ۴) 5

۳۲ـ اگر $Y = 4 + 2U$ و $S_U^2 = 1$ ، S_Y^2 کدام است؟

۱) 1 ۲) 2 ۳) 3 ۴) 4

۳۳ـ در کارخانه‌ای دو خط تولید وجود دارد که به ترتیب 40% و 60% کل تولید را به عهده دارند. می‌دانیم که به ترتیب 3% و 4% محصولات دو خط تولید معیوب است. اگر کالایی به تصادف انتخاب و معیوب باشد، احتمال آن‌که از خط تولید دوم باشد، چقدر است؟

۱) $\frac{1}{3}$ ۲) $\frac{3}{4}$ ۳) $\frac{2}{3}$ ۴) $\frac{1}{2}$

۳۴ـ اگر $P(A) = 0.4$ ، $P(B) = 0.1$ و $P(A \cap B) = 0.08$ باشد، A و B چگونه‌اند؟

۱) وابسته ۲) مکمل ۳) مستقل ۴) ناسازگار

۳۵ـ اگر X با توزیع $f(x) = \left\{\frac{x}{6} + \frac{1}{12} ; 0 \le x \le 3\right.$ متغیر تصادفی پیوسته باشد، امید ریاضی X کدام است؟

۱) $\frac{9}{8}$ ۲) $\frac{3}{2}$ ۳) $\frac{15}{8}$ ۴) 1

۳۶ـ سکه سالمی 20 بار پرتاب می‌شود، اگر متغیر تصادفی X تعداد یک روی ظاهر شده از سکه باشد، مطلوبست امید ریاضی و واریانس X؟

۱) $15 , \sqrt{15}$ ۲) $20 , 5$ ۳) $10 , \sqrt{5}$ ۴) $10 , 5$

۳۷ـ اندازه قد دانش‌آموزان کلاس اول دارای توزیع نرمال با میانگین 100 و انحراف معیار 5 سانتی‌متر است. متغیر توزیع نرمال استاندارد برای قد دانش‌آموزان بین 110 و 115 سانتیمتر، در کدام فاصله است؟

۱) $1 \le Z \le 2$ ۲) $2 \le Z \le 3$ ۳) $1.5 \le Z \le 2$ ۴) $2 \le Z \le 2.5$

۳۸ـ از یک نمونه‌گیری اطلاعات زیر در دست است:

$r_{XY} = 0.9 , \sigma_X = \sqrt{2} , \sigma_Y = \sqrt{8} , \bar{y} = 6 , \bar{x} = 3$

معادله رگرسیون کدام است؟

۱) $\hat{y} = 0.6 + 1.8x$ ۲) $\hat{y} = 1.2 + 1.8x$ ۳) $\hat{y} = -0.6 + 1.8x$ ۴) $\hat{y} = 1.2 - 1.8x$

۳۹ـ در آزمون فرض یک دامنه با سطح اطمینان 95% و $z = 1.96$ ، z متناظر با داده‌های نمونه کدام باشد تا فرضیه H_1 رد نشود؟

۱) $|z| < +1.96$ ۲) $z < +1.96$ ۳) $z > 1.96$ ۴) $|z| > 1.96$

۵۰ـ اگر تکرار رقم مجاز نباشد، چند عدد سه رقمی بخش‌پذیر بر 5 یافت می‌شود؟

۱) 136 ۲) 256 ۳) 320 ۴) 328

۵۱ـ جعبه‌ای محتوی 12 مهره سفید و 8 مهره سیاه است. متوالیاً دو مهره از آن بیرون می‌آوریم و مجدداً به جعبه برمی‌گردانیم. احتمال این که هر دو مهره سفید باشد، کدام است؟

۱) $\frac{3}{95}$ ۲) $\frac{6}{95}$ ۳) $\frac{33}{95}$ ✓ ۴) $\frac{9}{25}$

۵۲ـ چند عدد 5 رقمی یا 6 رقمی می‌توان با رقم‌های 1 , 2 , 2 , 2 , 3 , 4 درست کرد؟

۱) 120 ۲) 240 ۳) 720 ۴) 1440

۵۳ـ در یک بازی سرگرمی تاسی (مکعب شش وجهی منتظم) را می‌ریزیم و معادل عددی که تاس نشان می‌دهد با واحد تومان جایزه می‌گیریم. برای پرتاب هر بار تاس چند تومان باید بپردازیم تا بازی عادلانه باشد؟ (جمع جبری امید ریاضی برد و باخت مساوی صفر شود.)

۱) 3.5 ۲) 3 ۳) 6 ۴) 7

۵۴ـ اگر در تحلیل واریانس برای تفاوت بین میانگین‌های سه گروه جمعاً شامل 30 عضو، مجموع مجذورات بین گروه‌ها "SS_b" مساوی 8 و مجموع مجذورات داخل گروه‌ها "SS_w" مساوی 540 باشد مقدار آماره F (نسبت واریانس بین گروه‌های به واریانس داخل گروه‌ها) کدام است؟

۱) 0.1 ۲) $\frac{1}{45}$ ۳) 0.2 ۴) $\frac{2}{9}$

۵۵ـ Var (ax + by) برای دو متغیر تصادفی x و y همواره برابر است با:

۱) $a^2 \operatorname{var} x + b^2 \operatorname{Var} y + 2\operatorname{COV}(x - y)$
۲) $a^2 \operatorname{Var} x + b^2 \operatorname{Var} y + ab\operatorname{COV}(x , y)$
۳) $a^2 \operatorname{Var} x + b^2 \operatorname{Var} y$
۴) $a^2 \operatorname{Var} x + b^2 \operatorname{Var} y + 2ab\operatorname{COV}(x , y)$

۵۶ـ اگر فوتبالیستی 80% پنالتی‌هایش را وارد دروازه کند، چقدر احتمال دارد که چهارمین پنالتی او اولین موفقیتش باشد؟

۱) $\frac{64}{625}$ ۲) $\frac{4}{625}$ ۳) $\frac{4}{125}$ ✓ ۴) $\frac{64}{125}$

۵۷ـ در یک کارخانه سه نوبت (شیفت) صبح، عصر و شب به ترتیب 50%، 30% و 20% سهم کل تولید را به عهده دارند که نسبت کالاهای معیوب تولید شده در این سه نوبت به ترتیب 0.02 و 0.03 و 0.01 است. اگر کالای معیوبی از تولیدات این کارخانه به دستمان برسد، احتمال این که از شیفت شب باشد چقدر است؟

۱) $\frac{2}{21}$ ۲) $\frac{1}{3}$ ۳) $\frac{2}{1000}$ ۴) $\frac{21}{1000}$

۵۸ـ اگر در یک توزیع هنجاری $x = \int_{-\infty}^{a}$ باشد، $\int_{-a}^{+a}$ کدام است؟

۱) x - 0.5 ۲) 2- 2x ۳) 2x - 1 ۴) 2x - 0.5

۵۹ـ اگر در یک نمونه تصادفی 9 عضوی از یک جامعه آماری نسبتاً بزرگ میانگین نمونه مساوی 20 و انحراف معیار نمونه برابر 6 باشد. میانگین این جامعه در چه فاصله‌ای قرار دارد؟ اگر t برای هشت درجه و نه درجه آزادی به ترتیب 3.355 و 3.250 باشد.

۱) 23.250 , 16.750
۲) 36.71 , 13.29
۳) 24.5 , 15.5
۴) 23.355 , 16.645

۶۰ـ در یک آزمون بزرگ (کنکور) میانگین نمرات 60 و انحراف معیار نمرات 20 است. اگر 10% شرکت کنندگان بتوانند قبول شوند، حداقل نمره قبولی چقدر خواهد بود اگر بدانیم که $\int_{-\infty}^{1.28} = 0.9$

۱) 75.6
۲) 85
۳) 75
۴) 85.6

۶۱ـ سکه‌ای را طوری ساخته‌اند که احتمال آمدن روی آن دو برابر احتمال آمدن پشت آن است اگر این سکه را سه بار پرتاب کنیم احتمال این که یک بار روی آن و دوبار پشت آن ظاهر شود کدام است؟

۱) $\frac{2}{9}$
۲) $\frac{2}{3}$
۳) $\frac{4}{9}$
۴) $\frac{4}{81}$

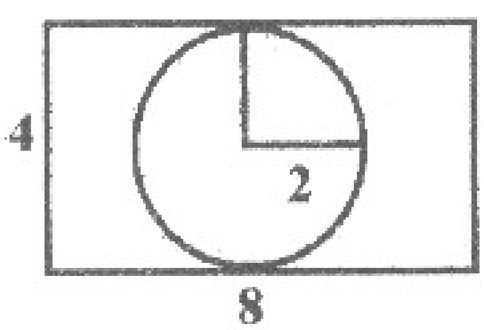
۶۲ـ اگر واریانس چند عدد مساوی 10 باشد و به هر یک از داده‌ها 20% اضافه شود، واریانس چقدر خواهد شد؟

۱) 11.2
۲) 11.44
۳) 14.4
۴) 10

۶۳ـ تاسی را آن قدر می‌ریزیم تا بالاخره عدد یک بالا قرار گیرد. احتمال آن‌که در سومین نوبت ریختن تاس، برای اولین دفعه عدد یک بالا قرار گیرد، کدام است؟

۱) $\frac{1}{216}$
۲) $\frac{1}{36}$
۳) $\frac{1}{6}$
۴) $\frac{25}{216}$

۶۴ـ تیراندازی به تصادف تیری به سمت شکل مقابل شلیک می‌کند. احتمال آن‌که تیر با دایره برخورد کند کدام است؟



۱) $\frac{\pi}{32}$
۲) $\frac{\pi}{8}$
۳) $\frac{\pi}{4}$
۴) $\frac{\pi}{2}$

۶۵ـ با ارقام صفر و 2 و 4 و 6 چند عدد سه رقمی بزرگتر از 400 و بدون تکرار ارقام می‌شود نوشت؟

۱) 6
۲) 12
۳) 18
۴) 24

۶۶ـ خانواده‌ای می‌خواهند چهار فرزند داشته باشند. احتمال آن‌که تعداد پسرهای این خانواده مساوی امید ریاضی تعداد پسرها باشد، کدام است؟

۱) $\frac{1}{4}$
۲) $\frac{3}{8}$
۳) $\frac{3}{4}$
۴) $\frac{1}{2}$

۶۷- در یک بررسی از یک توزیع نرمال $\mu = 100$ و $\delta = 15$ مشاهده شده است. چه نسبتی از این افراد 130 یا بالاتر از 130 هستند؟

$\int_{-2}^{+2} f_z(z)d_z = 0.9544$

۱) 0.0228 ✓ ۲) 0.0456 ۳) 0.0912 ۴) 0.4772

۶۸- اگر نمرات استاندارد شده دو نمره 10 و 5 به ترتیب 1 و 1- باشند زوج مرتب (μ , δ^2) در این جامعه کدام است؟

۱) (2.5 , 7.5) ۲) (7.5 , 2.5) ۳) (7.5 , 6.25) ۴) (7.5 , 5)

۶۹- 9 اسباب بازی را به چند طریق می‌توان به طور مساوی بین 3 بچه توزیع کرد؟

۱) 84 ۲) 504 ۳) 729 ۴) 1680

۷۰- ده سکه همتراز را با هم پرتاب می‌کنیم، اگر x را تعداد دفعات ظاهر شدن شیر در نظر بگیریم به طور متوسط انتظار داریم چند شیر ظاهر شود.

۱) 5 ۲) 10 ۳) 100 ۴) 1024

۷۱- در داده‌های زیر چارک اول کدام است؟

X_i:	10	5	7	11	14	15
f_i:	3	2	1	7	4	2

۱) 5 ۲) 7 ۳) 9 ۴) 10

۷۲- اگر حجم نمونه 100 و میانگین نمونه 12 و برآورد واریانس داده‌های نمونه 256 باشد خطای معیار میانگین نمونه کدام است؟

۱) 16% ۲) 1.6 ۳) 2.56 ۴) 25.6

۷۳- اگر $f(x) = cx \quad x = 1 , 2, 3 , 4 , 5$ یک تابع احتمال برمتغیر تصادفی x باشد در این صورت c برابر کدام است؟

۱) $\frac{1}{15}$ ✓ ۲) $\frac{1}{5}$ ۳) $\frac{1}{4}$ ۴) 15

۷۴- اگر H ، G ، X به ترتیب میانگین‌های حسابی و هندسی و همساز (هارمونیک) چند عدد مختلف باشند، کدام رابطه همواره بین آن‌ها درست است؟

۱) $X < G < H$ ۲) $G < H < X$ ۳) $H < G < X$ ۴) $H < X < G$

۷۵- در بسط $(2x_1 + x_2 + x_3)^6$ ضریب $x_1^3 x_2 x_3^2$ کدام است؟

۱) 60 ۲) 180 ۳) 240 ۴) 480

۷۶- یک فروشگاه شامل دو قسمت عمده فروشی و خرده فروشی است و 70% فروش این فروشگاه در قسمت خرده فروشی است. حسابدار این فروشگاه می‌داند که 5% از صورت حساب‌های خرده فروشی و 2% از صورت حساب‌های عمده فروشی اشتباه ثبت می‌شود. اگر او به یک مورد اشتباه برخورد کند چقدر احتمال دارد که این اشتباه مربوط به بخش خرده فروشی باشد؟

۱) $\frac{85}{100}$ ۲) $\frac{14}{100}$ ۳) $\frac{35}{41}$ ۴) $\frac{43}{70}$

۷۷ـ به چند طریق می‌توان از بین اعضای 12 نفره تیمی، 3 نفر را جهت مقام‌های اول تا سوم انتخاب کرد؟

۱) 110 ۲) 220 ۳) 1100 ۴) 1320

۷۸ـ اگر میانگین داده‌ها $x_{10}, \ldots, x_3, x_2, x_1$ برابر 5 باشد، میانگین 16 و $x_{10}, \ldots, x_2, x_1$ کدام است؟

۱) 5 ۲) 6 ۳) 6.6 ۴) 10.5

۷۹ـ از کیسه‌ای که محتوی 2 مهره قرمز و 3 مهره سبز است. دو مهره با هم و به طور تصادفی بیرون می‌آوریم احتمال این‌که دو مهره هم رنگ نباشند، کدام است؟

۱) $\frac{6}{25}$ ۲) $\frac{3}{10}$ ۳) $\frac{2}{5}$ ۴) $\frac{3}{5}$

۸۰ـ برای دو پیشامد A و B داریم $P(A)=\frac{1}{9}$ و $P(B|A)=\frac{1}{2}$ آن‌گاه $P(A\cap B)$ برابر است با:

۱) $\frac{1}{18}$ ۲) $\frac{3}{18}$ ۳) $\frac{7}{18}$ ۴) $\frac{11}{18}$

۸۱ـ در یک شرکت بازرگانی 20% حساب‌ها را حسابدار اول 40% را حسابدار دوم و 40% را حسابدار سوم تنظیم می‌نمایند. به تجربه ثابت شده که 2%، 3% و 5% از حساب‌ها به ترتیب توسط حسابدارهای اول و دوم و سوم دارای اشتباه می‌باشد. یکی از حساب‌های تنظیم شده در این شرکت را به طور تصادفی انتخاب می‌کنیم اگر حساب دارای اشتباه باشد احتمال این‌که توسط حسابدار دوم تنظیم شده باشد چقدر است؟

۱) 0.05 ۲) 0.036 ۳) 0.33 ۴) 0.83

۸۲ـ در یک توزیع دو جمله $\sigma = 5$ و $n = 100$ مقدار P کدام است؟

۱) $\frac{1}{2}$ ۲) $\frac{1}{3}$ ۳) 0.25 ۴) 0.75

۸۳ـ در یک نوار خاص به طور متوسط یک عیب در هر 220 متر وجود دارد. احتمال این‌که 2 عیب در یک بسته 880 متری وجود داشته باشد کدام است؟

۱) e^{-1} ۲) $4e^{-1}$ ۳) $4e^{-4}$ ۴) $8e^{-4}$

۸۴ـ اگر میانگین یک متغیر تصادفی X برابر با 10 و انحراف معیار آن برابر با 2 باشد آن‌گاه مقدار استاندارد شده برای $X = 20$ کدام است؟

۱) $\frac{1}{2}$ ۲) 2 ۳) 5 ۴) 10

۸۵ـ طول عمر یک نوع یخچال تقریباً دارای توزیع نرمال با میانگین 2 سال و انحراف معیار 1 سال است اگر این نوع یخچال برای یک سال تضمین شده باشد چه نسبتی از آن‌ها طی سال اول احتیاج به تعمیر پیدا می‌کند؟ ($P(z<1)=0.8413$)

۱) 0.3413 ۲) 0.8413 ۳) 0.1562 ۴) 0.1587

۸۶ـ اگر $SP_{xy} = 300$ و $SS_x = 200$ و $\bar{x} = 2$ و $\bar{y} = 3$ باشد معادله خط رگرسیون کدام است؟

۱) $y = 1.5x + 2$ ۲) $y = 1.5x$ ۳) $y = \frac{2}{3}x$ ۴) $y = \frac{2}{3}x + 2$

۸۷- تقاضای کالای x با توزیع نرمال برای 30 روز دارای میانگین $\bar{x} = 30$ و انحراف معیار $s = 5$ می‌باشد. فاصله اطمینان 99% برای میانگین جامعه μ برابر کدام است؟ $(Z_{0.005} = 2.58)$

۱) $30 \pm 2.58 \frac{5}{\sqrt{30}}$ ✓
۲) $30 \pm 2.58 \frac{\sqrt{5}}{\sqrt{30}}$
۳) $30 \pm 2.58 \times 5$
۴) $30 \pm 2.58 \frac{5}{30}$

۸۸- تاسی را آن‌قدر می‌ریزیم تا بالاخره عدد یک بالا قرار گیرد. احتمال آن‌که در سومین نوبت ریختن تاس، برای اولین دفعه عدد یک بالا قرار گیرد، کدام است؟

۱) $\frac{1}{6}$
۲) $\frac{1}{36}$
۳) $\frac{1}{216}$
۴) $\frac{25}{216}$

۸۹- می‌خواهیم یک کمیته 5 نفری از بین 5 مرد و 3 زن تشکیل دهیم. احتمال این‌که در این کمیته 2 زن و 3 مرد، باشند، برابر است با:

۱) $\frac{5}{28}$
۲) $\frac{15}{28}$
۳) $\frac{1}{56}$
۴) $\frac{12}{56}$

۹۰- 9 اسباب بازی را به چند طریق می‌توان به طور مساوی بین 3 بچه توزیع کرد؟

۱) 1680
۲) 729
۳) 504
۴) 84

۹۱- اگر X یک متغیر تصادفی با واریانس K باشد $Var(aX + b)$ کدام است؟

۱) aK
۲) $aK + b$
۳) a^2K ✓
۴) $a^2K + b$

۹۲- اگر در یک جدول توزیع فراوانی حجم جامعه برابر 40 و فراوانی مطلق طبقه سوم آن برابر 5 باشد، درصد فراوانی نسبی آن چند درصد است؟

۱) 0.05
۲) 0.125
۳) 8
۴) 12.5

۹۳- ده سکه همتراز را با هم پرتاب می‌کنیم، اگر X را تعداد دفعات ظاهر شدن شیر در نظر بگیریم به طور متوسط انتظار داریم چند شیر ظاهر شود؟

۱) 5
۲) 10
۳) 100
۴) 1024

۹۴- اگر حجم نمونه 100 و میانگین نمونه 12 و برآورد واریانس داده‌های نمونه 256 باشد، خطای معیار میانگین نمونه کدام است؟

۱) 0.16
۲) 1.6
۳) 2.56
۴) 25.6

۹۵- به چند طریق می‌توان از بین اعضای 12 نفره تیمی، 3 نفر را جهت مقام‌های اول تا سوم انتخاب کرد؟

۱) 110
۲) 220
۳) 1100
۴) 1320

۹۶- با ارقام 5 ، 4 ، 3 ، 2 ، 1 ، 0 چند عدد چهار رقمی زوج بدون تکرار می‌توان نوشت؟

۱) 148
۲) 152
۳) 156
۴) 180

۹۷- اگر S انحراف معیار داده‌های X_1 , X_2 , ... , X_{20} باشد، انحراف معیار $(2X_1+5)$, $(2X_2+5)$, ... , $(2X_{20}+5)$ کدام است؟

۱) $2S$ ۲) $2S+5$ ۳) $2S+10$ ۴) $2S+100$

۹۸- در یک میهمانی، شش زوج ازدواج کرده، شامل 6 مرد و همسران آن‌ها شرکت دارند. اگر به طور تصادفی دو نفر از بین آن‌ها انتخاب کنیم، احتمال آن‌که این دو نفر زن و شوهر باشند، کدام است؟

۱) $\frac{1}{12}$ ۲) $\frac{1}{11}$ ۳) $\frac{2}{11}$ ۴) $\frac{1}{6}$

۹۹- اگر در یک آزمایش تصادفی احتمال موفقیت را p و احتمال عدم موفقیت را q بدانیم، مقدار ماکزیمم pq کدام است؟

۱) 0.16 ۲) 0.25 ۳) 0.5 ۴) 1

۱۰۰- معمولاً $\frac{1}{100}$ مسافران هواپیماها به موقع به پرواز نمی‌رسند. احتمال آن‌که از 200 مسافر یک پرواز سه نفر به موقع نرسند، کدام است؟ $(e=2.7)$

۱) 0.000001 ۲) 0.01 ۳) 0.1 ۴) 0.18

۱۰۱- کدام خاصیت جزء خواص واریانس نیست؟

۱) $Var(X+a)=Var(X)$ ۲) $Var(bX+a)=b^2 Var(X)$
۳) واریانس X نمی تواند منفی باشد. ۴) $Var(bX)=b\,Var(X)$

۱۰۲- در آزمون‌های فرض، زمانی خطای نوع اول رخ می‌دهد که:

۱) فرض صفر رد نشود، در صورتیکه فرض صفر درست است.
۲) فرض صفر رد نشود، در صورتیکه فرض صفر غلط است.
۳) فرض صفر (H_0) رد شود در صورتیکه فرض صفر درست است.
۴) فرض صفر رد شود، در صورتیکه فرض صفر غلط است.

۱۰۳- اگر میانگین اعداد X_1 , X_2 , ... , X_n مساوی 20 باشد، مقدار $\sum_{i=1}^{n}(x_i-20)$ کدام است؟

۱) صفر ۲) n ۳) n^2-20 ۴) 20

۱۰۴- تاسی را آن‌قدر می‌ریزیم تا سرانجام عدد شش ظاهر شود. واریانس تعداد دفعاتی که باید منتظر باشیم تا سرانجام عدد شش ظاهر شود، کدام است؟

۱) $\frac{6}{5}$ ۲) $\frac{25}{6}$ ۳) 6 ۴) 30

۱۰۵- در یک بررسی میانگین و انحراف معیار بهره هوشی (I. Q.) با توزیع نرمال به ترتیب 100 و 15 بوده است. چه نسبتی از نمرات بهره هوشی بالاتر از 130 بوده است؟ $\left(\int_{-2}^{+2}=0.9544\right)$

۱) 0.0228 ۲) 0.0456 ۳) 0.0912 ۴) 0.4772

۱۰۶ـ درآمد حدود 95% رانندگان تاکسی در روز در فاصله 1000 تا 5000 تومان است. با فرض نرمال بودن توزیع درآمد، انحراف معیار درآمد این صنف کدام است؟

۱) 1000 ۲) 1500 ۳) 2000 ۴) 666.6

۱۰۷ـ اگر A و B دو پیشامد ناسازگار باشند، آن‌گاه $P(A \mid B)$ کدام است؟

۱) $P_{(A)}$ ۲) $P_{(B)}$ ۳) صفر ۴) یک

۱۰۸ـ در مجموعه اعداد 40 ، 80 ، 60 ، 50 ، 60 عدد 60 کدام پارامتر است؟

۱) میانه ۲) میانه و نما ۳) میانگین ۴) نما و میانگین

۱۰۹ـ اتومبیلی با سرعت 240 کیلومتر در ساعت از تهران به کرج رفته و با سرعت 160 کیلومتر در ساعت همین مسیر را بر می‌گردد، متوسط سرعت او در رفت و برگشت چند کیلومتر در ساعت است؟

۱) 192 ۲) 198 ۳) 200 ۴) 208

۱۱۰ـ به چند طریق می‌توان از بین ده سوال آمار و پنج سؤال ریاضی به 8 سؤال آمار و 2 سؤال ریاضی پاسخ داد؟

۱) 55 ۲) 450 ۳) 900 ۴) 1800

۱۱۱ـ اگر در جدول توزیع فراوانی‌ها، فراوانی مطلق طبقه سوم 12 و فراوانی نسبی همان طبقه 0.48 باشد، فراوانی تجمعی طبقه آخر کدام است؟

۱) 25 ۲) 50 ۳) 96 ۴) 100

۱۱۲ـ اگر انحراف معیار اعداد $x_1 , x_2 , \dots , x_n$ مساوی 8 باشد، واریانس اعداد $2x_1 - 1 , 2x_2 - 1 , \dots , 2x_n - 1$ کدام است؟

۱) 16 ۲) 128 ۳) 255 ۴) 256

۱۱۳ـ میانگن ده عدد مساوی 12 شده است. اگر یک عدد را کنارگذاریم میانگین 9 عدد باقیمانده مساوی 11 می‌شود. عددی که کنار گذاشته شده، کدام است؟

۱) 11 ۲) 12 ۳) 20 ۴) 21

۱۱۴ـ اگر معادله خط رگرسیون y نسبت به x به صورت $y = -2x + b$ و $SS_y = 4SS_x$ باشد، ضریب همبستگی بین x و y کدام است؟

۱) 1 ۲) 0.9 ۳) -1 ۴) -0.9

۱۱۵ـ نمرات احمد و محمود در درس ریاضی 18 و 12 و نمرات استاندارد شده آن‌ها به ترتیب 1 و 1 - می‌باشد، میانگین و انحراف معیار نمرات ریاضی کل دانشجویان به ترتیب کدامند؟

۱) 2 , 15 ۲) 3 , 15 ۴) 2 , 16 ۴) 3 , 16

۱۱۶ـ اگر میانگین و انحراف معیار یک نمونه چند عضوی به ترتیب 20 و 5 و مقدار z مساوی 2 باشد، **برآورد** فاصله‌ای میانگین جامعه کدام است؟

۱) 24 , 16 ۲) 22 , 18 ۳) 21 , 19 ۴) 10.5 , 19.5

۱۱۷ـ واریانس اعداد 3 , 7 , 11 , 15 و 19 کدام است؟

۱) $4\sqrt{2}$ ۲) 16 ۳) 32 ۴) 64

۱۱۸ـ برای یک جامعه $\sum x_i = 100$ و $\sum x_i^2 = 580$ می‌باشد اگر حجم جامعه $N = 20$ باشد ضریب تغییرات چند درصد است؟

۱) 80 ۲) 70 ۳) 50 ۴) 40

۱۱۹ـ احتمال برد یک تیم در بازی $\frac{2}{5}$ است. اگر این تیم 4 بازی انجام دهد احتمال این‌که بیش از نصف بازی‌ها را ببرد کدام است؟

۱) $\frac{4}{5}$ ۲) $\frac{8}{125}$ ۳) $\frac{60}{625}$ ۴) $\frac{112}{625}$

۱۲۰ـ چند عدد سه رقمی زوج بدون تکرار با ارقام 1 , 2 , 5 , 8 , 9 می‌توان نوشت؟

۱) 12 ۲) 24 ۳) 36 ۴) 60

۱۲۱ـ اگر A و B دو پیشامد ناسازگار باشند و احتمال آن دو به ترتیب a و b باشد مقدار $P(A^c \cap B^c)$ کدام است؟ (A^c متمم A است.)

۱) $b - a$ ۲) $1 - ab$ ۳) $1 - a - b$ ۴) $1 + a + b$

۱۲۲ـ با حروف کلمه «انقلاب اسلامی» چند کلمه چهار حرفی می‌توان ساخت در صورتی که تکرار حروف جایز نباشد؟

۱) 1380 ۲) 1680 ۳) 1830 ۴) 1860

۱۲۳ـ در پرتاب دو مکعب، احتمال این که مجموع دو عدد آمده 4 باشد به شرط این‌که دو عدد آمده یکسان نباشند کدام است؟

۱) $\frac{1}{15}$ ۲) $\frac{2}{36}$ ۳) $\frac{1}{5}$ ۴) $\frac{1}{3}$

۱۲۴ـ انحراف معیار داده‌ها $x_1 , x_2 , \ldots , x_n$ برابر a می‌باشد انحراف معیار داده‌های $-2x_1 + 3 , \ldots , -2x_n + 3$ چقدر است؟

۱) $-2a + 3$ ۲) $-2a$ ۳) $2a$ ۴) $4a$

۱۲۵ـ از اعداد طبیعی مضرب 7 کوچکتر از 451 سه عدد یکی یکی و بدون جایگذاری انتخاب می‌کنیم احتمال این‌که هر سه عدد مضرب 5 باشد چقدر است؟

۱) $\frac{(3^{12})}{(3^{64})}$ ۲) $\frac{12^3}{64^3}$ ۳) $\frac{12 \times 11 \times 10}{64 \times 63 \times 62}$ ۴) $\frac{10 \times 9 \times 8}{64 \times 63 \times 62}$

۱۲۶ـ یک قفل رمزدار دارای یک رمز سه رقمی فرد با ارقام 9 , ... , 2 , 1 می‌باشد اگر رمز قفل را ندانیم و برای پیدا کردن هر رمز 2 دقیقه طول بکشد حداکثر چند ساعت طول می‌کشد تا قفل باز شود؟

۱) 12 ۲) 12.5 ۳) 13 ۴) 13.5

۱۲۷ـ عددی به تصادف از فضای نمونه $\{1 , 2 , 3 , \ldots , 9\}$ انتخاب می‌کنیم احتمال آن‌که عدد انتخاب شده زوج یا مضرب 3 باشد کدام است؟

۱) $\frac{1}{3}$ ۲) $\frac{2}{3}$ ۳) $\frac{2}{9}$ ۴) $\frac{5}{9}$

۱۲۸ـ واریانس داده‌های 5 , 4 , 4 , 3 کدام است؟

۱) 0.25 ۲) 0.75 ۳) 0.5 ۴) 1.5

۱۲۹ـ در نمودار دایره‌ای 60 داده آماری، کمانی به اندازه 30 درجه به یک طبقه تعلق دارد، فراوانی مطلق آن طبقه کدام است؟

۱) 7 ۲) 5 ۳) 8 ۴) 10

۱۳۰ـ چارک اول داده‌های جدول زیر کدام است؟

حدود طبقات	1 - 5	5 - 9	9 - 13	13 - 17
فراوانی	9	8	12	14

۱) $\frac{47}{8}$ ۲) $\frac{43}{4}$ ۳) $\frac{23}{8}$ ۴) $\frac{23}{4}$

۱۳۱ـ کدامیک از موارد زیر از خواص توزیع نرمال است؟

۱) صفر = میانگین و 1 = انحراف معیار
۲) 1 = میانگین و صفر = انحراف معیار
۳) صفر = میانگین و 1 = میانه و نما
۴) میانه با میانگین منطبق است.

۱۳۲ـ در کدام حالت ضریب همبستگی (R) مستقیم و کامل است؟

۱) $R = 0$ ۲) $R = -1 < +1$ ۳) $R = -1$ ۴) $R = +1$

۱۳۳ـ وقتی که یک متغیر تصادفی مانند (X) بر طبق قانون توزیع نرمال با واریانس نامعلوم باشد، برای آزمون آن کدامیک از توابع آزمون کننده زیر مناسب است؟

۱) F فیشر ۲) t استیودنت ۳) t , z استیودنت ۴) X^2 پیرسون

۱۳۴ـ در صورتی که $Z = 1.96$ جدول و $-3 = $ (Z محاسبه شده) باشد چه قضاوتی می‌توان درباره H_0 و H_1 داشت؟

۱) H_0 در ناحیه بحرانی قرار نمی‌گیرد.
۲) H_0 قبول می‌شود.
۳) H_0 رد می‌شود.
۴) H_1 رد می‌شود.

۱۳۵ـ یک مؤسسه برای چهار نفر دعوتنامه می‌فرستد، به چند طریق ممکن است نامه هیچکس به دست خودش نرسد؟

۱) 17 ۲) 13 ۳) 12 ۴) 9

۱۳۶ـ توزیع احتمال متغیر x به صورت جدول زیر است. احتمال این‌که حداقل x برابر 2 باشد کدام است؟

x	0	1	2	3
P	$\frac{1}{16}$	$\frac{1}{4}$	a	$\frac{1}{4}$

۱) $\frac{1}{8}$
۲) $\frac{5}{8}$
۳) $\frac{8}{11}$
۴) $\frac{11}{16}$

۱۳۷ ـ توزیع احتمال متغیر x به صورت زیر است، میانگین جامعه کدام است؟

x	0	1	2	3
P	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

۱) 1.5
۲) 1.85 ✓
۳) 2.5
۴) 2.85

۱۳۸ ـ میانگین 10 عدد برابر 12 می‌باشد. دو عدد را اشتباهی به جای 8 و 4 برابر 18 و 14 گرفته‌ایم. میانگین صحیح چقدر می‌شود؟

۱) 8　۲) 14　۳) 10　۴) 12

۱۳۹ ـ میانه n عدد نمایانگر چیست؟ (وقتی که داده‌ها از کوچک به بزرگ مرتب شده باشند)

۱) وسط ترین عدد را نشان می‌دهد.
۲) تعداد اعدادی را نشان می‌دهد که از میانگین بزرگترند.
۳) تعداد اعدادی را نشان می‌دهد که از میانگین کوچکترند.
۴) میانگین تغییرات بین داده‌ها را نشان می دهد.

۱۴۰ ـ در یک جعبه 4 مهره سفید، 4 مهره سیاه و 4 مهره قرمز وجود دارد، از این جعبه 1 مهره بیرون می کشیم و می بینیم سیاه نیست، احتمال این که سفید باشد برابر کدام است؟

۱) کمتر از $\frac{1}{2}$　۲) $\frac{1}{2}$　۳) بیشتر از $\frac{1}{2}$　۴) نزدیک به 1

۱۴۱ ـ به چند طریق 7 نفر می‌توانند دور هم بنشینند در حالی که یک فرد همواره در جای ثابتی باشد؟

۱) $5!$　۲) $7!$　۳) $7 \times 5!$　۴) $6!$

۱۴۲ ـ واریانس داده‌های آماری صفر است، کدام نتیجه گیری در این مورد درست می باشد؟

۱) میانگین داده صفر است.　۲) تمام داده‌ها برابرند. ✓
۳) داده‌ها متقارن هستند.　۴) میانگین بیشتر از هر داده

۱۴۳ ـ در آزمون کاردانی علمی کاربردی از چهار صد هزار شرکت کننده، یکصد هزار نفر دیپلم فنی حرفه‌ای، 150 هزار نفر دیپلم تجربی و 50 هزار نفر دیپلم انسانی بقیه دیپلم ریاضی هستند. در نمودار دایره‌ای، چند درصد دیپلم انسانی می‌باشند و زاویه متناظر آن کدام است؟

۱) 12.5 درصد ـ 30 درجه　۲) 12.5 درصد ـ 45 درجه
۳) 10 درصد ـ 25 درجه　۴) 10 درصد ـ 30 درجه

۱۴۴ ـ در داده‌های آماری، داده‌ای که بیشترین فراوانی را دارد، کدام است؟

۱) میانه　۲) میانگین　۳) چارک اول　۴) مد یا نما

۱۴۵ ـ میانگین هندسی اعداد 1 , 2 , 32 برابر کدام است؟

۱) 2　۲) 4　۳) 8　۴) $\frac{35}{3}$

۱۴۶ـ از دوازده عدد تخم‌مرغ که سه عدد آن شکسته و بقیه سالم‌اند سه عدد تخم مرغ را به تصادف انتخاب می‌کنیم، احتمال آن که هر سه عدد سالم باشند، کدام است؟

۱) $\frac{3}{4}$ ۲) $\frac{3}{12}$ ۳) $\frac{23}{45}$ ۴) $\frac{21}{55}$

۱۴۷ـ میانگین چهار عدد 80 است. اگر به این چهار عدد، عدد 30 اضافه شود، میانگین 5 عدد حاصل کدام است؟

۱) 55 ۲) 60 ۳) 70 ۴) 75

۱۴۸ـ در جدول داده‌های روبرو مقدار میانگین کدام است؟

x_i	1	3	5	7	9
f_i	2	4	8	5	1

۱) 4.7
۲) 4.9 ✓
۳) 5.1
۴) 5.2

۱۴۹ـ از بین 8 نفر به چند طریق می‌توان یک تیم حداقل 2 نفره انتخاب کرد؟

۱) 247 ۲) 216 ۳) 164 ۴) 112

۱۵۰ـ هرگاه X یک متغیر تصادفی باشد که در آن $P(X = C) = 1$، C مقدار ثابت آن‌گاه $E(X)$ و $Var(X)$ به ترتیب کدام است؟

۱) $0 , 1$ ۲) $0 , C$ ✓ ۳) $C , 1$ ۴) C^2 , C

۱۵۱ـ یک جفت تاس را 144 بار پرتاب می‌کنیم، انتظار دارید چند بار هر دو عدد رو شده، مضرب 3 باشند؟

۱) 36 ۲) 16 ۳) 18 ۴) 24

۱۵۲ـ در یک جدول توزیع فراوانی که شامل 50 داده آماری است و به صورت صعودی مرتب شده است، فراوانی تجمعی دسته ماقبل آخر برابر 42 می‌باشد. درصد فراوانی نسبی طبقه آخر کدام است؟

۱) 84 ۲) 42 ۳) 16 ۴) 12

۱۵۳ـ هرگاه x یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f(x) = \frac{3}{2}x^2 + \frac{1}{2}k$ وقتی $0 < x < 1$ و در جای دیگر برابر صفر باشد، مقدار k چقدر است؟

۱) 1 ✓ ۲) 2 ۳) $\frac{2}{3}$ ۴) $\frac{3}{2}$

۱۵۴ـ 5 دانشجوی حسابداری و 3 دانشجوی کامپیوتر در یک صف ایستاده‌اند احتمال این که اول و آخر صف دانشجوی حسابداری باشد چقدر است؟

۱) $\frac{3}{14}$ ۲) $\frac{5}{14}$ ۳) $\frac{9}{28}$ ۴) $\frac{9}{14}$

۱۵۵ ـ اگر X یک متغیر تصادفی با جدول توزیع احتمال زیر باشد: $E(3X - 2E(X))$ چقدر است؟

x	-1	1	2
$P(X=x)$	$\frac{1}{5}$	$1-a$	$\frac{1}{5}$

۱) $-\frac{1}{5}$ ۲) $\frac{1}{5}$ ۳) $\frac{2}{5}$ ۴) $\frac{4}{5}$

۱۵۶ـ یک نمونه 9 تایی از حلب‌های 5 کیلویی تولیدی یک کارخانه که توزیع نرمال دارند انتخاب کرده‌ایم در این نمونه میانگین وزن حلب‌ها 5 کیلو و انحراف معیار وزن حلب‌ها 0.2 کیلوگرم است یک فاصله اطمینان با ضریب 95 درصد برای میانگین وزن تمام حلب‌ها کدام است؟

$\left(t_{0.975}(8) = 2.25\ ,\ z_{0.975} = 1.95\right)$

۱) $(4.75\ ,\ 5.25)$ ۲) $(4.85\ ,\ 5.15)$

۳) $(4.87\ ,\ 5.13)$ ۴) $(4.77\ ,\ 5.23)$

۱۵۷ـ هرگاه واریانس 5 , 4 , 3 , 2 , 1 برابر a^2 باشد. اعداد 210 , 208 , 206 , 204 , 202 کدام است؟

۱) $4a^2$ ۲) $2a^2$

۳) $2a^2 + 10$ ۴) $4a^2 + 10$

۱۵۸ـ 10 عدد معین را در 4 ضرب کرده و 20 را از آن‌ها کم می‌کنیم. میانگین و انحراف معیار اعداد جدید به ترتیب 4 , 2 است. میانگین و انحراف معیار اولیه کدام هستند؟

۱) $\sigma_x = 1, \mu_x = 4$ ۲) $\sigma_x = 1, \mu_x = 5$

۳) $\sigma_x = \frac{1}{2}, \mu_x = 6$ ۴) $\sigma_x = 1, \mu_x = 4$

۱۵۹ـ اگر X تعداد پرتاب های منجر به گل یک بسکتبال باشد و میانگین آن‌ها 20 و واریانس آن $\frac{20}{3}$ باشد آن‌گاه احتمال این‌که در 5 پرتاب حداقل 1 پرتاب گل شود چقدر است؟

۱) $\frac{1}{243}$ ۲) $\frac{80}{243}$ ۳) $\frac{239}{243}$ ۴) $\frac{242}{243}$

۱۶۰ـ احتمال این‌که معادله زیر دو جواب حقیقی داشته باشد چقدر است؟ $x^2 - ax + 1 = 0$ که در آن a تعداد شیرها در 3 پرتاب سکه است.

۱) $\frac{1}{2}$ ۲) $\frac{1}{4}$ ۳) $\frac{1}{8}$ ۴) $\frac{3}{8}$

۱۶۱ـ اگر تابع احتمال X به صورت زیر باشد:

$$P_x(x)=\begin{cases}\frac{1}{5} & x=0\\ \frac{2}{5} & x=2\\ \frac{2}{5} & x=3\\ 0 & \text{سایر مقادیر}\end{cases}$$

چارک اول و میانه چقدر می‌شود؟

۱) $t_{25}=2\ ,\ t_{50}=2$
۲) $t_{25}=0\ ,\ t_{50}=2$
۳) $t_{25}=0\ ,\ t_{50}=3$
۴) $t_{25}=2\ ,\ t_{50}=3$

۱۶۲ـ اگر n زوج باشد آن‌گاه مقدار $C_n^0-C_n^1+C_n^2+-\dots+C_n^n$ چه مقدار است؟

۱) ✓ 0
۲) 1
۳) 2^n
۴) 2^{-n}

۱۶۳ـ کدام یک از فرمول‌ها ذیل، درست است؟

۱) ✓ $(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$

۲) $(x+y)^n=\sum_{k=1}^{n}\binom{n}{k}x^k y^{n-k}$

۳) $(x+y)^n=\sum_{k=1}^{n}\binom{n}{k}x^{k-1} y^{n-k-1}$

۴) $(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^{k-1} y^{n-k-1}$

۱۶۴ـ یک جفت تاس را، یک بار می‌ریزیم و متوجه می‌شویم که دو عددی که آمده‌اند، یکسان نمی‌باشند کدامیک از موارد ذیل، احتمال آن است که مجموع 7 باشد؟

۱) $\frac{2}{36}$
۲) $\frac{1}{5}$
۳) $\frac{1}{6}$
۴) $\frac{2}{5}$

۱۶۵ـ فرض کنید X، یک متغیر تصادفی باشد که مقادیر $2\ ,\ 2^2\ ,\ 2^3\ ,\dots ,\ 2^k\ ,\dots$ را با احتمال $P(X_k)=P(2^k)=\frac{1}{2^k}$، اختیار نماید، در این صورت کدامیک از گزاره‌های ذیل درست است؟

۱) میانگین X، برابر با 1 است.
۲) واریانس X، برابر با 1 است.
۳) ✓ میانگین قابل محاسبه نیست.
۴) انحراف معیار برابر با $\sqrt{2}$ است.

۱۶۶ـ فرض کنید در کیسه‌ای، 4 توپ سفید و 8 توپ سیاه وجود داشته باشد، دو توپ را، بدون جایگذاری از کیسه خارج می‌کنیم، کدامیک از موارد ذیل، احتمال آن است که هر دو توپ سفید باشند؟

۱) $\frac{2}{11}$ ۲) $\frac{3}{12}$ ۳) $\frac{1}{11}$ ۴) $\frac{2}{12}$

۱۶۷ـ فرض کنید A و B دو پیشامد باشند، کدامیک از روابط ذیل، درست است؟

۱) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

۲) $P(A\cap B)=P(A)+P(B)+P(A\cup B)$

۳) $P(A\cap B)=P(A)+P(B)-P(A\mid B)$

۴) $P(A\cup B)=P(A)+P(B)-P(A\mid B)$

۱۶۸ـ فرض کنید T، یک متغیر تصادفی از نوع نمائی باشد، کدام یک از روابط ذیل برقرار است؟

۱) $P(P>a+b\mid T>a)=P(T>b)$

۲) $P(T>a+b\mid T>b)=P(T>a)$

۳) $P(T>a-b\mid T>b)=P(T>b)$

۴) $P(T>a+b\mid T>a)=P(T>a)$

۱۶۹ـ فرض کنید A_n , ... , A_2 , A_1، پیشامدهایی باشند که دو به دو از هم جدا و فضای پیشامدی را بسازند، در این صورت کدام یک از فرمول های ذیل، درست است؟ فرض کنید A، یک پیشامد دلخواه باشد.

۱) $P[A]=P[A_1][A\mid A_1]+P[A_2]P[A_1\mid A_2]+\dots+P[A_n]P[A_{n-1}\mid A_n]$

۲) $P[A]=P[A_1][A\mid A_1]+P[A_2]P[A\mid A_2]+\dots+P[A_n]P[A\mid A_n]$

۳) $P[A]=P[A_1][A\mid A_1]+P[A_2]P[A_2\mid A_1]+\dots+P[A_n]P[A_n\mid A_{n-1}]$

۴) $P[A]=P[A_1][A\mid A_1]+P[A_2]P[A_2\mid A_1]+\dots+P[A_n][A_n\mid A_{n-1}]$

۱۷۰ـ فرض کنید X، یک متغیر تصادفی گسسته و h(X) یک تابع تعریف شده بر روی X باشد و X بتواند مقادیر x_1 , x_2 , ... را اختیار نماید. در اینصورت کدامیک از موارد ذیل، درست است؟

۱) $E[h(X)]=h(x_1).P(x_1)+h(x_2).P(x_2)+\dots$

۲) $E[h(X)]=h(x_1).P(x_1\mid x_2)+h(x_2).P(x_2\mid x_3)+\dots$

۳) $E[h(X)]=h(x_1).P(x_2\mid x_1)+h(x_2).P(x_3\mid x_2)+\dots$

۴) $E[h(X)]=\int_{-\infty}^{\infty}h(x).P(x_i).dx_i$

۱۷۱ـ یک عدد 3 رقمی را به تصادف انتخاب می‌کنیم، کدامیک از موارد زیر، احتمال آن است که این عدد حداقل یک رقم 1، داشته باشد؟

۱) 0.72 ۲) 0.029 ۳) 0.28 ۴) 0.69

۱۷۲ـ فرض کنید A ، B و C سه پیشامد بر روی فضای احتمالاتی مفروض باشند. کدام یک از گزاره‌های ذیل درست است؟

۱) $P[A \cup B \cup B \cup C] = P[A] + P[B] + P[C]$

۲) $P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B]$

۳) $P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C] + P[A \cap B \cap C]$

۴) $P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C]$

۱۷۳ـ فرض کنید n، زوج باشد، کدامیک از موارد ذیل، برابر با $C_n^1 + C_n^3 + \dots + C_n^{n-1}$ می‌باشد.

۱) $C_n^0 + C_n^2 + \dots + C_n^{n-2}$

۲) $C_n^0 + C_n^2 + \dots + C_n^n$

۳) $C_n^2 + C_n^4 + \dots + C_n^n$

۴) $C_n^2 + C_n^4 + \dots + C_n^{n-2}$

۱۷۴ـ فرض کنید در یک جعبه، دو لامپ قرمز، سه لامپ آبی و چهار لامپ سبز وجود داشته باشد، به چند طریق می‌توان سه لامپ از این جعبه را، انتخاب کرد؟

۱) C_9^3 ۲) $C_2^1 + C_3^1 + C_4^1$ ۳) $\frac{3!}{6!}$ ۴) $\frac{3!}{9!}$

۱۷۵ـ در جعبه‌ای، 4 توپ سفید و 8 توپ سیاه قرار دارد. دو توپ به تصادف و بدون جای‌گذاری از آن خارج می‌کنیم. کدامیک از موارد ذیل، احتمال آن را نشان می‌دهد که توپ دوم، سفید باشد؟

۱) $\frac{1}{1}$ ۲) $\frac{1}{12}$ ۳) $\frac{1}{3}$ ۴) $\frac{2}{11}$

۱۷۶ـ فرض کنید X، یک متغیر تصادفی باشد، کدامیک از روابط ذیل، درست است؟

۱) $var[X] = E[X^2] - (E[X])^2$ ✓

۲) $var[X] = E[X^2] - (E[X-\mu])^2$

۳) $var[X] = E[(X-\mu)^2] - (E[X])^2$

۴) $var[X] = E[X] - (E[X-\mu])^2$

پاسخنامه

۱ـ گزینه ۴ صحیح می‌باشد.

۲ـ گزینه ۱ صحیح می‌باشد.

۳ـ گزینه ۳ صحیح می‌باشد.

۴ـ گزینه ۲ صحیح می‌باشد.

۵ـ گزینه ۴ صحیح می‌باشد.

۶ـ گزینه ۳ صحیح می‌باشد.

۷ـ گزینه ۲ صحیح می‌باشد.

۸ـ گزینه ۳ صحیح می‌باشد.

۹ـ گزینه ۱ صحیح می‌باشد.

۱۰ـ گزینه ۳ صحیح می‌باشد.

۱۱ـ گزینه ۱ صحیح می‌باشد.

۱۲ـ گزینه ۳ صحیح می‌باشد.

۱۳ـ گزینه ۲ صحیح می‌باشد.

۱۴ـ گزینه ۲ صحیح می‌باشد.

۱۵ـ گزینه ۱ صحیح می‌باشد.

۱۶ـ گزینه ۱ صحیح می‌باشد.

۱۷ـ گزینه ۳ صحیح می‌باشد.

۱۸ـ گزینه ۴ صحیح می‌باشد.

۱۹ـ گزینه ۳ صحیح می‌باشد.

۲۰ـ گزینه ۲ صحیح می‌باشد.

۲۱ـ گزینه ۲ صحیح می‌باشد.

۲۲ـ گزینه ۲ صحیح می‌باشد.

۲۳ـ گزینه ۳ صحیح می‌باشد.

۲۴ـ گزینه ۴ صحیح می‌باشد.

۲۵ـ گزینه ۴ صحیح می‌باشد.

۲۶ـ گزینه ۲ صحیح می‌باشد.

۲۷- گزینه ۴ صحیح می‌باشد.

۲۸- گزینه ۱ صحیح می‌باشد.

۲۹- گزینه ۱ صحیح می‌باشد.

۳۰- گزینه ۳ صحیح می‌باشد.

۳۱- گزینه ۳ صحیح می‌باشد.

۳۲- گزینه ۴ صحیح می‌باشد.

۳۳- گزینه ۲ صحیح می‌باشد.

۳۴- گزینه ۱ صحیح می‌باشد.

۳۵- گزینه ۳ صحیح می‌باشد.

۳۶- گزینه ۴ صحیح می‌باشد.

۳۷- گزینه ۲ صحیح می‌باشد.

۳۸- گزینه ۱ صحیح می‌باشد.

۳۹- گزینه ۳ صحیح می‌باشد.

۴۰- گزینه ۴ صحیح می‌باشد.

۴۱- گزینه ۲ صحیح می‌باشد.

۴۲- گزینه ۴ صحیح می‌باشد.

۴۳- گزینه ۳ صحیح می‌باشد.

۴۴- گزینه ۲ صحیح می‌باشد.

۴۵- گزینه ۱ صحیح می‌باشد.

۴۶- گزینه ۲ صحیح می‌باشد.

۴۷- گزینه ۴ صحیح می‌باشد.

۴۸- گزینه ۲ صحیح می‌باشد.

۴۹- گزینه ۲ صحیح می‌باشد.

۵۰- گزینه ۱ صحیح می‌باشد.

۵۱- گزینه ۴ صحیح می‌باشد.

۵۲- گزینه ۲ صحیح می‌باشد.

۵۳- گزینه ۱ صحیح می‌باشد.

۵۴- گزینه ۳ صحیح می‌باشد.

۵۵- گزینه ۴ صحیح می‌باشد.

۵۶- گزینه ۲ صحیح می‌باشد.

۵۷- گزینه ۱ صحیح می‌باشد.

۵۸- گزینه ۳ صحیح می‌باشد.

۵۹- گزینه ۲ صحیح می‌باشد.

۶۰- گزینه ۴ صحیح می‌باشد.

۶۱- گزینه ۱ صحیح می‌باشد.

۶۲- گزینه ۳ صحیح می‌باشد.

۶۳- گزینه ۴ صحیح می‌باشد.

۶۴- گزینه ۲ صحیح می‌باشد.

۶۵- گزینه ۲ صحیح می‌باشد.

۶۶- گزینه ۲ صحیح می‌باشد.

۶۷- گزینه ۱ صحیح می‌باشد.

۶۸- گزینه ۳ صحیح می‌باشد.

۶۹- گزینه ۴ صحیح می‌باشد.

۷۰- گزینه ۱ صحیح می‌باشد.

۷۱- گزینه ۴ صحیح می‌باشد.

۷۲- گزینه ۲ صحیح می‌باشد.

۷۳- گزینه ۱ صحیح می‌باشد.

۷۴- گزینه ۳ صحیح می‌باشد.

۷۵- گزینه ۴ صحیح می‌باشد.

۷۶- گزینه ۳ صحیح می‌باشد.

۷۷- گزینه ۴ صحیح می‌باشد.

۷۸- گزینه ۲ صحیح می‌باشد.

۷۹- گزینه ۴ صحیح می‌باشد.

۸۰- گزینه ۱ صحیح می‌باشد.

۸۱- گزینه ۳ صحیح می‌باشد.

۸۲- گزینه ۱ صحیح می‌باشد.

۸۳ـ گزینه ۴ صحیح می‌باشد.

۸۴ـ گزینه ۳ صحیح می‌باشد.

۸۵ـ گزینه ۴ صحیح می‌باشد.

۸۶ـ گزینه ۲ صحیح می‌باشد.

۸۷ـ گزینه ۱ صحیح می‌باشد.

۸۸ـ گزینه ۴ صحیح می‌باشد.

۸۹ـ گزینه ۲ صحیح می‌باشد.

۹۰ـ گزینه ۱ صحیح می‌باشد.

۹۱ـ گزینه ۳ صحیح می‌باشد.

۹۲ـ گزینه ۴ صحیح می‌باشد.

۹۳ـ گزینه ۱ صحیح می‌باشد.

۹۴ـ گزینه ۲ صحیح می‌باشد.

۹۵ـ گزینه ۴ صحیح می‌باشد.

۹۶ـ گزینه ۳ صحیح می‌باشد.

۹۷ـ گزینه ۱ صحیح می‌باشد.

۹۸ـ گزینه ۲ صحیح می‌باشد.

۹۹ـ گزینه ۲ صحیح می‌باشد.

۱۰۰ـ گزینه ۴ صحیح می‌باشد.

۱۰۱ـ گزینه ۴ صحیح می‌باشد.

۱۰۲ـ گزینه ۳ صحیح می‌باشد.

۱۰۳ـ گزینه ۱ صحیح می‌باشد.

۱۰۴ـ گزینه ۴ صحیح می‌باشد.

۱۰۵ـ گزینه ۱ صحیح می‌باشد.

۱۰۶ـ گزینه ۱ صحیح می‌باشد.

۱۰۷ـ گزینه ۳ صحیح می‌باشد.

۱۰۸ـ گزینه ۲ صحیح می‌باشد.

۱۰۹ـ گزینه ۱ صحیح می‌باشد.

۱۱۰ـ گزینه ۲ صحیح می‌باشد.

۱۱۱ - گزینه ۱ صحیح می‌باشد.

۱۱۲ - گزینه ۴ صحیح می‌باشد.

۱۱۳ - گزینه ۴ صحیح می‌باشد.

۱۱۴ - گزینه ۳ صحیح می‌باشد.

۱۱۵ - گزینه ۲ صحیح می‌باشد.

۱۱۶ - گزینه ۳ صحیح می‌باشد.

۱۱۷ - گزینه ۲ صحیح می‌باشد.

۱۱۸ - گزینه ۴ صحیح می‌باشد.

۱۱۹ - گزینه ۴ صحیح می‌باشد.

۱۲۰ - گزینه ۲ صحیح می‌باشد.

۱۲۱ - گزینه ۳ صحیح می‌باشد.

۱۲۲ - گزینه ۲ صحیح می‌باشد.

۱۲۳ - گزینه ۱ صحیح می‌باشد.

۱۲۴ - گزینه ۳ صحیح می‌باشد.

۱۲۵ - گزینه ۳ صحیح می‌باشد.

۱۲۶ - گزینه ۱ صحیح می‌باشد.

۱۲۷ - گزینه ۲ صحیح می‌باشد.

۱۲۸ - گزینه ۳ صحیح می‌باشد.

۱۲۹ - گزینه ۲ صحیح می‌باشد.

۱۳۰ - گزینه ۱ صحیح می‌باشد.

۱۳۱ - گزینه ۴ صحیح می‌باشد.

۱۳۲ - گزینه ۴ صحیح می‌باشد.

۱۳۳ - گزینه ۲ صحیح می‌باشد.

۱۳۴ - گزینه ۳ صحیح می‌باشد.

۱۳۵ - گزینه ۴ صحیح می‌باشد.

۱۳۶ - گزینه ۴ صحیح می‌باشد.

۱۳۷ - گزینه ۲ صحیح می‌باشد.

۱۳۸ - گزینه ۳ صحیح می‌باشد.

۱۳۹ـ گزینه ۱ صحیح می‌باشد.

۱۴۰ـ گزینه ۱ صحیح می‌باشد.

۱۴۱ـ گزینه ۳ صحیح می‌باشد.

۱۴۲ـ گزینه ۲ صحیح می‌باشد.

۱۴۳ـ گزینه ۲ صحیح می‌باشد.

۱۴۴ـ گزینه ۴ صحیح می‌باشد.

۱۴۵ـ گزینه ۲ صحیح می‌باشد.

۱۴۶ـ گزینه ۴ صحیح می‌باشد.

۱۴۷ـ گزینه ۳ صحیح می‌باشد.

۱۴۸ـ گزینه ۲ صحیح می‌باشد.

۱۴۹ـ گزینه ۱ صحیح می‌باشد.

۱۵۰ـ گزینه ۲ صحیح می‌باشد.

۱۵۱ـ گزینه ۲ صحیح می‌باشد.

۱۵۲ـ گزینه ۳ صحیح می‌باشد.

۱۵۳ـ گزینه ۱ صحیح می‌باشد.

۱۵۴ـ گزینه ۲ صحیح می‌باشد.

۱۵۵ـ گزینه ۴ صحیح می‌باشد.

۱۵۶ـ گزینه ۲ صحیح می‌باشد.

۱۵۷ـ گزینه ۱ صحیح می‌باشد.

۱۵۸ـ گزینه ۳ صحیح می‌باشد.

۱۵۹ـ گزینه ۴ صحیح می‌باشد.

۱۶۰ـ گزینه ۱ صحیح می‌باشد.

۱۶۱ـ گزینه ۱ صحیح می‌باشد.

۱۶۲ـ گزینه ۱ صحیح می‌باشد.

۱۶۳ـ گزینه ۱ صحیح می‌باشد.

۱۶۴ـ گزینه ۲ صحیح می‌باشد.

۱۶۵ـ گزینه ۳ صحیح می‌باشد.

۱۶۶ـ گزینه ۳ صحیح می‌باشد.

۱۶۷ـ گزینه ۱ صحیح می‌باشد.

۱۶۸ـ گزینه ۲ صحیح می‌باشد.

۱۶۹ـ گزینه ۲ صحیح می‌باشد.

۱۷۰ـ گزینه ۱ صحیح می‌باشد.

۱۷۱ـ گزینه ۳ صحیح می‌باشد.

۱۷۲ـ گزینه ۳ صحیح می‌باشد.

۱۷۳ـ گزینه ۲ صحیح می‌باشد.

۱۷۴ـ گزینه ۱ صحیح می‌باشد.

۱۷۵ـ گزینه ۳ صحیح می‌باشد.

۱۷۶ـ گزینه ۱ صحیح می‌باشد.