آمار و احتمال — فرمولهای فصل اول

n : تعداد داده‌ها ، حجم جامعهٔ آماری K : تعداد رده‌ها ، تعداد طبقات

C : طول رده ، فاصله طبقات F_i : فراوانی مطلق R : دامنهٔ تغییرات

f_i : فراوانی نسبی Fc_i : فراوانی تراکمی یا تجمعی کمتر از یا فراوانی تجمعی بیشتر از

۱) $R = x_{max} - x_{min}$ $\Rightarrow$ دامنه تغییرات = مقدار بزرگترین داده − مقدار کوچکترین داده

۲) $n = 2^K$ ۳) $k = 1 + 3.3 \log(n)$ ۴) $C = \frac{R}{K}$ $\Rightarrow$ طول رده = دامنه تغییرات / تعداد رده

۵) $\sum F_i = n$ ۶) $\sum f_i = 1$ ۷) فراوانی نسبی = فراوانی مطلق / حجم جامعه آماری $\Rightarrow f_i = \frac{F_i}{\sum F_i} = \frac{F_i}{n}$

شاخص‌های گرایش مرکزی :

۱- میانگین $\bar{X}$: فرمول میانگین حسابی در جوامع نمونه (داده) (عدد متوسط جامعه)

۸) $\bar{X} = \frac{\sum^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$

فرمول " " در جدول : m_i معدل هر رده و F_i فراوانی مطلق هر رده

۹) $\bar{X} = \frac{\sum F_i m_i}{n}$

۲- میانه : (عدد وسط جامعه) (Md یا $\tilde{X}$)

الف - روش داده‌های معمولی برای محاسبهٔ میانه : اول داده‌ها را به ترتیب نزولی یا صعودی مرتب می‌کنیم. اگر تعداد داده‌ها فرد باشد $n=2K+1$ یک عدد در وسط قرار می‌گیرد عدد $(K+1)$ ام میانه خواهد بود. K عدد در چپ و K عدد در راست قرار می‌گیرد.
ولی اگر تعداد داده‌ها زوج باشد $(n=2K)$ معدل دو عدد وسط میانه خواهد بود. یعنی معدل عدد K ام و $(K+1)$ ام

ب - روش محاسبهٔ میانه از روی جدول : ستون فراوانی تجمعی را پیدا می‌کنیم (FC_i) - $\frac{N}{2}$ را حساب می‌کنیم → در ستون FC_i عددی که بلافاصله بعد از $\frac{N}{2}$ قرار دارد پیدا کرده و آن را طبقه ~~ضبی~~ میانه‌دار می‌نامیم.

L_i : کرانهٔ پایین طبقه میانه‌دار

۱۰) $Md = L_i + \frac{\frac{N}{2} - FC_{i-1}}{F_i} \times C$

نکته : اگر داده‌ها پیوسته باشند در L_i همان کرانهٔ پایین طبقه را می‌نویسیم ولی درصورتیکه داده‌ها ناپیوسته باشند، از کرانهٔ پایین طبقه ۰/۵ کم می‌کنیم $L_i = L_i - 0.5$

۳- مُد MO : بیشترین تکرار در جامعه آماری الف- روش محاسبهٔ مد از روی داده‌ها :

۸-۴-۲-۳-۴-۵ $\Rightarrow MO = 4$

۸-۴-۳-۳-۴-۵ $\Rightarrow MO = 4, 3$

ب - روش محاسبهٔ مد از روی جدول :

d_1 : اختلاف فراوانی طبقه مددار با طبقه بالاتر

d_2 : اختلاف فراوانی طبقه مددار با طبقه پایین‌تر

۱۱) $MO = L_i + \frac{d_1}{d_1 + d_2} \times C$

انواع نمودارها: ۱- نمودار چند ضلعی (پلی گون) ۲- نمودار میله‌ای ۳- نمودار ستونی (هیستوگرام - بارگراف) ۴- نمودار دایره‌ای

هیستوگرام: پیوسته - برای واحدهای پیوسته یا شلیل های که بصورت پیوسته‌اند.

بارگراف: ناپیوسته - جداولی که ستون اول آن ناپیوسته نوشته شده باشد. شلیل ها دارای یک واحد فاصله اند.

نکته: برای رسم نمودار از دو بعد استفاده می کنیم. یکی نمودار افقی و دیگری نمودار عمودی. در محور افقی اگر x_i داشته باشیم آن را می نویسیم و اگر نداشته باشیم، CL را می نویسیم. در محور عمودی F_i را می نویسیم. در نمودار شلیلی بجای x_i در محور افقی CL می نویسیم.

نمودار دایره‌ای: کل حجم جامعه را یک دایره فرض می کنیم.

$$۳۵)\quad S=\frac{۳۶۰}{N}\times F_i$$

توزیع های فراوانی می تواند هر شکل و وضعی داشته باشند:

۱- توزیع زنگ شکل متقارن ۲- توزیع با چولگی منفی یا چولگی به چپ ۳- توزیع با چولگی مثبت یا چولگی به راست

۴- توزیع J معکوس شکل ۵- توزیع U شکل

ضریب چولگی پیرسن: $\tilde{x}$ میانه ، $\bar{x}$ میانگین

$$۳۶)\quad SK=\frac{۳(\bar{x}-\tilde{x})}{S}$$

میزان کشیدگی یا برجستگی منحنی: S انحراف معیار نمونه ، $M_۴$ گشتاور مرکزی رتبه چهارم

$$۳۷)\quad K=\frac{M_۴}{S^۴}-۳$$

گشتاور مرکزی رتبه چهارم:

$$۳۸)\quad M_۴=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^۴$$

جدولهای دو متغیره: مشاهدات حقیقی مربوط به دو صفت از جامعه که بصورت زیر نمونه گیری شده اند را در جدولهای دو بعدی مرتب می کنیم:

(x_1,y_1) و (x_2,y_2) و ... و (x_n,y_n)

معیارهای توصیفی برای دو متغیر:

$$۳۹)\quad \bar{x}=\frac{1}{n}\sum_{j=1}^{t}\sum_{i=1}^{s}n_{ij}x_i=\frac{1}{n}\sum_{i=1}^{s}n_{i0}x_i$$

$$۴۰)\quad \bar{y}=\frac{1}{n}\sum_{j=1}^{t}\sum_{i=1}^{s}n_{ij}y_j=\frac{1}{n}\sum_{j=1}^{t}n_{0j}y_j$$

$$۴۱)\quad S_x^2=\frac{1}{n}\sum_{i=1}^{s}\sum_{j=1}^{t}n_{ij}(x_i-\bar{x})^2=\frac{1}{n}\sum_{i=1}^{s}n_{i0}(x_i-\bar{x})^2$$

$$۴۲)\quad S_y^2=\frac{1}{n}\sum_{i=1}^{s}\sum_{j=1}^{t}n_{ij}(y_i-\bar{y})^2=\frac{1}{n}\sum_{j=1}^{t}n_{0j}(y_i-\bar{y})^2$$

ترکیب کمیتی که می تواند تاثیرات دو متغیر بر یکدیگر را اندازه گیری کند کوواریانس x و y می نامند.

$$۴۳)\quad S_{xy}=\frac{1}{n}\sum_{j=1}^{t}\sum_{i=1}^{s}n_{ij}(y_i-\bar{y})(x_i-\bar{x}) \Rightarrow S_{xy}=\frac{1}{n}\sum_{j=1}^{t}\sum_{i=1}^{s}n_{ij}(x_iy_j-\bar{x}\bar{y})$$

اگر دو متغیر همسو باشند مقدار کوواریانس مثبت می شود و اگر دو متغیر همسو نباشند مقدار کوواریانس منفی می شود.

$$۴۴)\quad S_{ax+b,\,cy+d}=ac\,S_{xy}$$

ضریب همبستگی خطی: سنجش میزان وابستگی دو متغیر

$$۴۵)\quad r_{xy}=r=\frac{S_{xy}}{S_xS_y}$$

نکته: ۱- $r_{ax+b,\,cy+d}=\pm r_{xy}$

۲- ضریب همبستگی خطی r در فاصله $-1<r\leqslant 1$ تغییر می کند.

۳- $r=\pm1$ ضریب همبستگی کامل ، $r<0$ ضریب همبستگی معکوس ، $r>0$ ضریب همبستگی مستقیم ⇐ بنابراین هر چه r به 1+ نزدیکتر، همبستگی خطی شدیدتر و هرچه به صفر نزدیک‌تر شود همبستگی ضعیف تر می شود. (مثالها در جزوه)

فرمول محاسبه میانگین به روش کدگذاری

$$(24)\quad \bar{x} = m_0 + \frac{\sum_{i=1}^{k} u_i f_i}{n}\, c$$

که در آن m_0 نماینده رده‌ای است که در کدگذاری نماینده‌ی جدید ۰ را برای آن در نظر گرفته‌ایم و c طول مشترک رده‌هاست.

فرمول محاسبه انحراف معیار به روش کدگذاری

$$(25)\quad S = c\sqrt{\frac{n\left(\sum_{i=1}^{k} u_i^2 f_i\right) - \left[\sum_{i=1}^{k} u_i f_i\right]^2}{n(n-1)}}$$

کاربردهای انحراف معیار:

۱- ضریب تغییر: از تقسیم انحراف معیار بر میانگین بدست می‌آید و معمولاً در عدد ۱۰۰ ضرب می‌شود.

$$(26)\quad V = \frac{S}{\bar{x}} \times 100$$

۲- نمره استاندارد: $x_1, x_2, \dots, x_n$ داده‌ها و $\bar{x}$ میانگین این داده‌ها و $n, \dots$ و ...

$$(27)\quad z_i = \frac{x_i - \bar{x}}{S}$$

چنانچه بخواهیم دو مشاهده از دو نمونه‌ی مختلف را با هم مقایسه کنیم، استاندارد شده آن کار را آسان می‌کند.

$$z_1, z_2, \dots, z_n \Rightarrow \bar{z} = 0 \quad \text{و} \quad S_z = 1$$

۳- قضیه چبیشف: چنانچه پراکندگی دو نمونه از یک جمعیت را با یکدیگر مقایسه کنیم می‌گوییم نمونه‌ای که دارای انحراف معیار بزرگتری است پراکندگی آن بیشتر است. چنانچه داده‌ها به میانگین نزدیک‌تر شود انحراف معیار به صفر نزدیک‌تر می‌شود.

قضیه: برای هر مجموعه از داده‌ها و هر عدد ثابت $k \geq 1$، حداقل $(1-\frac{1}{k})\times 100$ درصد داده باید به فاصله k برابر انحراف معیار در دو طرف میانگین قرار گیرند.

قضیه چبیشف برای حالتهای خاص k بصورت زیر است:

۱- ممکن است تعداد خیلی کم داده‌ها به بازه $(\bar{x}-S \text{ و } \bar{x}+S)$ تعلق داشته باشند.

۲- حداقل ۷۵ درصد داده‌ها به بازه $(\bar{x}-2S \text{ و } \bar{x}+2S)$ تعلق خواهند داشت.

دستور تجربی: برای داده‌هایی که منحنی فراوانی آنها زنگی شکل است داریم:

الف - تقریباً ۶۸ درصد داده‌ها به بازه $(\bar{x}-S \text{ و } \bar{x}+S)$ تعلق دارد.

ب - " ۹۵ " " " $(\bar{x}-2S, \bar{x}+2S)$ " " .

ج - " ۹۹٫۷ " " " $(\bar{x}-3S, \bar{x}+3S)$ " " .

چارک‌ها:

چارک اول (Q_1) نقطه‌ای است که ۲۵ درصد حجم جامعه آماری در پایین آن قرار دارد و چارک دوم (Q_2) (میانه) نقطه‌ای است که ۵۰٪ داده‌ها در پایین آن قرار دارد و چارک سوم (Q_3) نقطه است که ۷۵٪ داده‌ها در پایین آن قرار دارد.

$$(28)\quad Q = \frac{Q_3 - Q_1}{2}$$ انحراف چارکی

$$(29)\quad R = Q_3 - Q_1$$ دامنه تغییرات بین چارکها

$$(30)\quad Q_1 = L_i + \frac{\frac{N}{4} - FC_{i-1}}{F_i} \times C$$ چارک اول

$$(31)\quad Q_3 = L_i + \frac{\frac{3N}{4} - FC_{i-1}}{F_i} \times C$$ چارک سوم

$$(32)\quad D_6 = L_i + \frac{\frac{6N}{10} - FC_{i-1}}{F_i} \times C$$ دهک ششم

$$(33)\quad P_{70} = L_i + \frac{\frac{70N}{100} - FC_{i-1}}{F_i} \times C$$ صدک هفتادم

$$(34)\quad \text{ضریب تغییر چارکی} = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

۴- میانگین همساز (هارمونیک) $(\bar{X}_H)$: معکوس میانگین حسابی معکوس های داده هاست ،

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

فرمول میانگین همساز برای زمانی که داده ها در اختیار باشند.

$$۱۲)\ \bar{X}_H = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$$

فرمول میانگین همساز برای زمانی که جدول فراوانی (داده های رده بندی شده) در اختیار باشند.

$$۱۳)\ \bar{X}_H = \frac{n}{\sum_{i=1}^{n}\frac{1}{m_i}f_i}$$

m_i : معدل رده یا طبقه

نکته : اگر داده ها دارای واحد اندازه گیری ترکیبی باشد مثل متر بر ثانیه و ... یا داده های مربوط به آهنگ تغییرات میانگین همساز مفیدتر از حسابی است.

۵- میانگین هندسی : میانگین هندسی n عدد مثبت ، ریشه ی n ام حاصلضرب آنهاست .

اگر داده ها در اختیار باشند ⇐

$$۱۴)\ \bar{x}_G = \sqrt[n]{x_1x_2\dots x_n} = (x_1x_2\dots x_n)^{\frac{1}{n}} \qquad ۱۵)\ \log \bar{x}_G = \frac{1}{n}\sum_{i=1}^{n}\log(x_i)$$

اگر جدول در اختیار باشد ⇐ m_i : مرکز رده iام و f_i فراوانی این رده

$$۱۶)\ \log(\bar{x}_G) = \frac{1}{n}\sum_{i=1}^{k}f_i\log(m_i)$$

نکته ۱: میانگین هندسی در پیدا کردن آهنگ متوسط تغییرات ، مثلا در جمعیت و یا محاسبه میانگین نسبت تغییرات مانند افزایش قیمت و عنوان ، مفید است .

نکته ۲ : در داده های مثبت رابطه بین میانگین ها به صورت زیر است :

$$۱۷)\ \underset{\text{حسابی}}{\bar{x}} \geq \underset{\text{هندسی}}{\bar{x}_G} \geq \underset{\text{همساز}}{\bar{x}_H}$$

شاخص های پراکندگی :

۱- دامنه : ساده ترین معیار سنجش پراکندگی

دامنه تغییرات = مقدار بزرگترین داده − مقدار کوچکترین داده

عیب عمده دامنه آن است که فقط بزرگترین و کوچکترین مقدار داده ها را در نظر می گیرد.

۲- انحراف معیار : مفیدترین معیار سنجش پراکندگی

نکته : در هر جامعه آماری مجموع انحراف از میانگین صفر است . برای رفع این مشکل ، انحراف داده ها از میانگین را داخل قدر مطلق قرار می دهیم.

$$۱۸)\ \sum_{i=1}^{n}(x_i - \underset{\text{میانگین}}{\bar{x}}) = 0$$

فرمول

$$۱۹)\ E = \frac{\sum|x_i - \bar{x}|}{n}$$

(انحراف میانگین) میانگین قدر مطلق انحرافات

فرمول انحراف معیار و واریانس درصورتیکه داده ها در اختیار باشند.

$$۲۰)\ S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \qquad ۲۱)\ S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

فرمول 〃 〃 〃 〃 داده های رده بندی شده در اختیار باشند.

$$۲۲)\ S^2 = \frac{\sum_{i=1}^{k}(m_i - \bar{x})^2 f_i}{n-1} \qquad ۲۳)\ S = \sqrt{\frac{\sum_{i=1}^{k}(m_i - \bar{x})^2 f_i}{n-1}}$$

واریانس و میانگین به روش کدگذاری :

در این روش نماینده یکی از رده ها (ترجیحاً رده ی وسطی یا یکی از رده های واقع در اواسط جدول) را صفر (۰) اختیار می کنیم . سپس نماینده های رده های بالای این رده را اعداد صحیح منفی ⇐ و واقع در زیر این رده را اعداد صحیح مثبت ، یعنی ۱، ۲، ... اختیار می کنیم و این نماینده های جدید رده ها را با u_1 ، u_2 ، ... ، u_n نشان می دهیم.

تعمیم اصل اول شمارش (اصل جمع): اگر در انجام عملی بتوانیم یکی از k عمل را انتخاب کنیم به طوری که عمل اول به n_1 راه و عمل دوم به n_2 راه و... کل عمل به $n_1+n_2+\dots+n_k$ طریق انجام می‌شود.

تعمیم اصل دوم شمارش (اصل ضرب): اگر عملی مرکب از k مرحله باشد به طوری که مرحله اول به n_1 طریق و برای هر یک از این راهها مرحله دوم به n_2 طریق و بدین ترتیب و آنگاه کل عمل به $n_1 n_2 \dots n_k$ طریق صورت می‌پذیرد.

جایگشت: صورتهای مختلف قرار گرفتن n شئ مختلف را (n به n)، پهلوی هم، جایگشت می‌گویند.

نکته: تعداد جایگشت‌های n شئ متمایز که روی یک دایره مرتب شده‌اند برابر است با:

$$1)\ n! = n(n-1)\dots\times 2\times 1 \Rightarrow 1)\ P_n = n!$$

$$2)\ P_n = (n-1)!$$

ترتیب: تعداد ترتیب r تایی از n شئ متمایز برای $r=0,\dots,n$ عبارت است از:

$$3)\ {}_nP_r = \frac{n!}{(n-r)!}$$

نکته: ترتیب قرار گرفتن اشیا پهلوی هم مهم می‌باشد.

فرمول ترتیب یا جایگشت با تکرار بصورت n^r می‌باشد.

$$4)\ n^r$$

ترکیب: اگر n شئ متمایز داشته باشیم و بخواهیم از n شئ فقط r شئ انتخاب کنیم به طوریکه ترتیب r شئ مهم نباشد برابر است با:

$$5)\ {}_nC_r = \binom{n}{r} = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

فرمول ترکیب با تکرار بصورت مقابل است:

$$6)\ \binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1!\, n_2! \dots n_k!}$$

نکته:

۱- ترکیب هر عدد با یک خود آن عدد است. $\binom{n}{1} = n$

۲- ترکیب هر عدد با خودش یک است. $\binom{n}{n} = 1$

۳- $0! = 1$

۴- ${}_nP_0 = 1$ 　 ${}_nP_0 = \frac{n!}{(n-0)!} = 1$

ضرایب دوجمله‌ای:

برای هر عدد صحیح مثبت n داریم:

$$7)\ (x+y)^n = \sum_{i=0}^{n} \binom{n}{r} x^{n-r} y^n$$

$$8)\ \binom{n}{r} = \binom{n}{n-r} \quad r=0,1,2,\dots,n \quad \xRightarrow{\text{اثبات}} \binom{n}{n-r} = \frac{n!}{(n-r)!\,(n-(n-r))!} = \frac{n!}{r!\,(n-r)!} = \binom{n}{r}$$

$$9)\ \binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

$$10)\ \sum_{r=0}^{k} \binom{m}{r}\binom{n}{k-r} = \binom{m+n}{k}$$

k، m و n عدد صحیح مثبت

نکته: ضریب چندجمله‌ای $x_1^{r_1} x_2^{r_2} \dots x_k^{r_k}$ در بسط $(x_1+x_2+\dots+x_n)^n$ برابر است با:

$$11)\ \binom{n}{r_1, r_2, \dots, r_k} = \frac{n!}{r_1!\, r_2! \dots r_k!}$$

آمار و احتمال

آزمایش تصادفی: یک آزمایش تصادفی (یا آماری) آزمایشی است که در آن فضای نمونه (S) مشخص باشد، نتایجی که از قبل قطعی نیست (پیشامد یا برآمد) بدست آید و تحت شرایط یکسان تکرار شود.

پیشامد: یک پیشامد زیرمجموعه‌ای از S است و با A و B و C نشان داده می‌شود. گوییم پیشامد A رخ می‌دهد هرگاه برآمد آزمایش عضوی از آن باشد.

نکته ۱: اگر پیشامد A تنها یک عضو داشته باشد ⇐ پیشامد ساده — اگر $A = S$ ⇐ پیشامد حتمی

اگر پیشامد A تهی باشد ⇐ پیشامد محال — اگر هر عضو از A در B نیز باشد ⇐ $A \subset B$ زیرپیشامد

نکته ۲: اگر $B \subset A$ و $A \subset B$ ⇐ $A = B$ پیشامدهای برابر

نکته ۳: $A \cup B$ اجتماع دو پیشامد از تمام عضوهای A یا B یا هر دو تشکیل می‌شود. زمانی رخ می‌دهد که لااقل یکی از A و B رخ دهد.

نکته ۴: اگر A و B دو پیشامد از S باشند و $A \cap B = \phi$ ⇐ A و B را دو پیشامد جدا از هم یا ناسازگار می‌نامند.

نکته ۵: تفاضل دو پیشامد A از B را با $A - B$ نشان می‌دهد که برابر است با تمام عناصری که در A باشد در B نباشد.

نکته ۶: اگر $A \subset B$ باشد $A - B = \phi$ و به $B - A$ یک تفاضل واقعی گویند، زیرا B تمام A را در بر می‌گیرد.

نکته ۷: تفاضل پیشامد A از فضای نمونه یعنی $S - A$ را متمم پیشامد A می‌گویند و آن را با A' نشان می‌دهند.

برای محاسبه متمم احتمال، احتمال محاسبه شده را از یک کم می‌کنیم.

1) $A \cup A' = S$ ، $A \cap A' = \phi$ 10) $P(A') = 1 - P(A)$

نکته: تفاضل متقارن دو پیشامد A و B عناصر آن یا عضوی از A، یا عضوی از B ولی نه عضو هر دو تشکیل می‌شود و آن را با $A \Delta B$ نشان می‌دهند.

2) $A \Delta B = (A-B) \cup (B-A) = (A \cup B) - (A \cap B)$

دو پیشامد سازگار و ناسازگار: هرگاه در وقوع دو پیشامد اشتراک آنها تهی باشد، ناسازگار و اگر غیر تهی باشد ⇐ سازگار

$A \cap B = \phi$ ⇐ ناسازگار

$A \cap B \neq \phi$ ⇐ سازگار

احتمال اجتماع دو پیشامد:

3) $A \cap B = \phi$ ⇐ ناسازگار ⇒ $P_{A \cup B} = P_A + P_B$

4) $A \cap B \neq \phi$ ⇐ سازگار ⇒ $P_{A \cup B} = P_A + P_B - P_{A \cap B}$

احتمال: هر موقع از وقوع قطعی یک حادثه نتوانیم با اطمینان حرف بزنیم از کلمهٔ احتمال استفاده می‌کنیم، و با P نشان می‌دهیم.

مجموعه حالات مساعد، مجموعه‌ای است که حالات دلخواه از انجام یک آزمایش را نشان می‌دهد.

$P(A) = \dfrac{\text{تعداد مجموعهٔ حالات مساعد}}{\text{تعداد مجموعهٔ فضای نمونه}}$ ⇐ 5) $P(A) = \dfrac{n(A)}{n(S)}$

6) $P(\phi) = 0$ تصادفی $0 < P(A) < 1$

$P(A) = 1$ حتمی الوقوع $P(A) = 0$ غیرممکن

احتمال یک پیشامد عدد حقیقی نامنفی است

7) $P(S) = 1$ 8) $0 \leqslant P(A) \leqslant 1$

اگر $A_1, A_2, A_3, \ldots$ دنباله متناهی یا نامتناهی از پیشامدهای دو به دو ناسازگار در S باشند، آنگاه ⇐

9) $P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$

فرمول تعداد عضوهای فضای نمونه : ۱۱) $n(S) = (\text{تعداد پیشامد در یک آزمایش})^{\text{تعداد آزمایش}}$

نکته : در سکه پایه ۲ است و در تاس پایه ۶ است.

مثال : سه بار پرتاب تاس $\Rightarrow$ ۱۳) $n(s)=6^3$ سه بار پرتاب سکه $\Rightarrow$ ۱۲) $n(s)=2^3$

نکته : در احتمال اگر (و) تفسیر شود، احتمال ضعیف‌تر می‌شود $\Rightarrow$ در هم ضرب می‌کنیم
" " " (یا) " " " قوی‌تر " $\Rightarrow$ باهم جمع می‌کنیم

قضیه : برای هر دو پیشامد ناسازگار از S مانند A و B داریم: $P(A\cup B)=P(A)+P(B)$

نکته : اگر A یک پیشامد و A' متمم آن باشد داریم: $S=A\cup A'$

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

قضیه : اگر A و B دو پیشامد از فضای نمونه S باشند و $A\subset B$ آنگاه ۱۴) $P(A)\leqslant P(B)$

قضیه : اگر A و B دو پیشامد از فضای نمونه S باشند آنگاه $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

قضیه : اگر A و B و C سه پیشامد دلخواه از فضای نمونه‌ی S باشند آنگاه

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

احتمال شرطی : هرگاه وقوع یک احتمالی را مشروط بر وقوع احتمال دیگر کنیم، آن را احتمال شرطی می‌نامند و با $P(A/B)$ نشان می‌دهند

۱۶) $P(A|B)=\dfrac{P(A\cap B)}{P(B)}$ $\xrightarrow{\text{قاعده ضرب}}$ $P(A)\neq 0 \Rightarrow$ ۱۷) $P(A\cap B)=P(A)\,P(B|A)$

$\xrightarrow{\text{تعمیم قاعده ضرب}}$ $P(A)\neq 0$, $P(A\cap B)\neq 0 \Rightarrow$ ۱۸) $P(A\cap B\cap C)=P(A).P(B|A).P(C|A\cap B)$

طبق جزوه : فرمول احتمال شرطی $\Rightarrow$ $P(A|B)=P(A)\times P(B)$ معمولاً حرف ربط (و) علامت احتمال شرطی است. یا اینکه جمله را بتوان طوری تفسیر کرد که از (و) استفاده نشود.

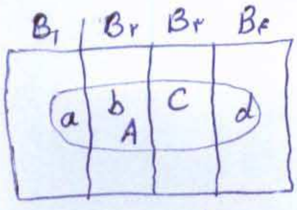
دو پیشامد مستقل و غیرمستقل : اگر وقوع یک پیشامدی بر احتمال وقوع پیشامد دیگر هیچ تأثیری نداشته باشد آن دو پیشامد را مستقل و اگر نتیجه آنها برهم دیگر تأثیر داشته باشد، آن دو پیشامد را غیرمستقل می‌نامند.

(بدون جایگذاری $\Rightarrow$ غیرمستقل (وابسته) و با جایگذاری $\Rightarrow$ مستقل (غیروابسته))

دو پیشامد A و B را مستقل گویند اگر و فقط اگر $P(A\cap B)=P(A).P(B)$

اگر A و B دو پیشامد مستقل باشند آنگاه A و B' ، A' و B ، A' و B' نیز مستقل هستند.

قاعده بیز :



$P(A)=P(B_1)P(A|B_1)+\dots+P(B_4)P(A|B_4)$ $\left\{ \begin{array}{l} i=1,2,\dots,k \\ P(B_i)\neq 0 \end{array}\right.$

$a=P(A|B_1)$ $b=P(A|B_2)$ $c=P(A|B_3)$ $d=P(A|B_4)$

۱۹) $P(A)=\sum_{i=1}^{k} P(B_i)P(A|B_i)$

۲۰) $P(B_r|A)=\dfrac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$ $\left\{ \begin{array}{l} i=1,2,\dots,k \\ P(B_i)\neq 0 \\ P(A)\neq 0 \\ r=1,2,\dots,k \end{array}\right.$

فرمولهای فصل چهارم

آمار و احتمال

متغیر تصادفی X : متغیری است که مقدار آن پس از انجام آزمایش مشخص می‌شود. (تابعی است عضوی که روی S تعریف می‌شود.)

متغیر تصادفی گسسته (جدا) : متغیری است که مجموعه برد آن مجموعه‌ای با تعداد اعضای متناهی یا نامتناهی ولی شمارش پذیر باشد.

متغیر تصادفی پیوسته : متغیری است که مجموعه برد آن مجموعه‌ای با تعداد اعضای نامتناهی و شمارش ناپذیر باشد.

مراحل حل مسأله :

۱- شناخت کامل آزمایش تصادفی　　۲- تعریف فضای نمونه آزمایش (به طور مناسب)

۳- تعریف متغیر تصادفی مناسب　　۴- تعیین برد متغیر تصادفی

۵- احتمال برابر متغیر تصادفی را با هر کدام از اعضای برد متغیر تصادفی محاسبه می‌کنیم.

۶- محاسبه احتمال مجهول با استفاده از مقادیر تابع احتمال

تابع احتمال (توزیع احتمال) : تابعی است که دامنه آن مقادیر ممکن متغیر تصادفی و حوزه آن احتمال مربوط به هر مقدار متغیر تصادفی است.

تابع احتمال گسسته (توزیع احتمال گسسته) : $f(x)$

x_i	x_1	x_2	....	x_n
$f(x)$	$f(x_1)$	$f(x_2)$		$f(x_n)$

۱- شرایط تابع احتمال گسسته ⇐

$1)\ \forall x_i \in X \quad f(x_i) \geq 0$　　$2)\ \sum_{i=1}^{n} f(x_i) = 1$

تابع احتمال پیوسته (تابع چگالی متغیر تصادفی پیوسته) : اگر متغیر تصادفی X پیوسته ⇐ ~~نمودار~~ نمودار توزیع احتمال پیوسته

۲- شرایط تابع احتمال پیوسته :

$1)\ f(x) \geq 0$　　$2)\ \int_{-\infty}^{+\infty} f(x)\,dx = 1$

$3)\ x = a \Rightarrow P(x=a) = 0$　　$4)\ a<x<b \Rightarrow P(a<x<b) = \int_a^b f(x)\,dx$

$5)\ P(a<x<b) = P(a \leq x < b) = P(a < x \leq b) = P(a \leq x \leq b)$

تابع توزیع (تابع احتمال تجمعی) : تابعی است که به ازاء جمیع مقادیر ممکن متغیر تصادفی X، احتمال وقوع مقدار کوچکتر یا مساوی با x را نشان می‌دهد. با $P(X \leq x)$ یا $F(x)$ نشان می‌دهند.

۳- اگر متغیر تصادفی X گسسته باشد ⇐ تابع توزیع ⇐

$$F_x(x) = P(X \leq x) = \sum_{t \leq x} f_x(t)$$

۴- اگر متغیر تصادفی X پیوسته باشد ⇐ تابع توزیع ⇐

$$F_x(x) = P(X \leq x) = \int_{-\infty}^{x} f_x(t)$$

۵- شرایط تابع توزیع :

$1)\ F(x) \geq 0$　　$2)\ 0 \leq F(x) \leq 1$　　$3)\ F(x)$ تابع غیر نزولی

$4)\ a<b \Rightarrow F(a) \leq F(b)$　　$5)\ F(x)$ از راست پیوسته است　　$6)\ F(+\infty)=1,\ F(-\infty)=0$

$7)\ P(a \leq X \leq b) = F(b^{+}) - F(a^{-})$　　۸) اگر تابع توزیع پیوسته باشد ⇐

$$\begin{cases} f(x) = \dfrac{dF(x)}{dx} \\ P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt \end{cases}$$

آمار و احتمال

فرمولهای فصل چهارم

۹) درصورتیکه تابع توزیع نقاط گسستگی داشته باشد:

$$f(x)=\begin{cases}(F(x))' & \text{در فواصل}\\ |F(x^+)-F(x^-)| & \text{در نقاط مرزی}\end{cases}$$

۱۰) محاسبه احتمال با استفاده از تابع توزیع $\Rightarrow$

$$\begin{cases}P(X=a)=|F(a^+)-F(a^-)| & \text{در نقاط گسسته } a\\ P(X=a)=0 & \text{درصورت پیوسته بودن}\end{cases}$$

نکته: اگر از تابع چگالی مشتق بگیریم و برابر صفر قرار دهیم، مد یا نما بدست می آید.

۶) مد: $f'(x)=0 \Rightarrow X=MO$

$F''(x)=0 \Rightarrow X=MO$

نکته: اگر تابع توزیع را برابر $\frac{1}{2}$ قرار دهیم، مقدار میانه بدست می آید.

۷) میانه: $\int_{-\infty}^{x} f(t)\,dt=\frac{1}{2} \Rightarrow x=$ میانه

نکته: اگر تابع توزیع را برابر $\frac{a}{4}$ قرار دهیم. $(a=1,2,3)$ چارک ها بدست می آید.

۸) چارک: $\int_{-\infty}^{x} f(t)\,dt=\frac{a}{4} \Rightarrow (X=Q_a$ چارک$)$

توزیع احتمال توام: تابع چگالی احتمال توام (مشترک):

اگر X و Y دو متغیر تصادفی باشند احتمال پیشامد همزمان X و Y را توزیع احتمال توام X و Y گویند و با $f(x,y)$ نشان می‌دهند.

خواص توزیع احتمال توام:

۹) $f(x,y)\geq 0 \Rightarrow \forall(x,y)$

۱۰) $\sum_x \sum_y f(x,y)=1$ (گسسته)

۱۱) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy$ (پیوسته)

توزیع حاشیه‌ای X و Y:

۱۲) اگر X و Y گسسته باشند. $\Rightarrow$

$$\begin{cases}f(x)=\sum_y f(x,y) & \text{توزیع حاشیه‌ای } X\\ f(y)=\sum_x f(x,y) & \text{توزیع حاشیه‌ای } Y\end{cases}$$

۱۳) اگر X و Y پیوسته باشند $\Rightarrow$

$$\begin{cases}f(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy & \text{توزیع حاشیه‌ای } X\\ f(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx & \text{توزیع حاشیه‌ای } Y\end{cases}$$

نکته: X و Y را مستقل می گویند اگر و فقط اگر

۱۴) $f(x,y)=f(x)\,f(y)$

توزیع احتمال شرطی:

توزیع احتمال شرطی متغیر تصادفی X درصورتیکه $Y=y$ چنین است:

۱۵) $f(x/y)=\dfrac{f(x,y)}{f_Y(y)}\ ,\quad f_Y(y)>0$

توزیع احتمال شرطی متغیر تصادفی Y درصورتیکه $X=x$ چنین است:

۱۶) $f(y/x)=\dfrac{f(x,y)}{f_X(x)}\ ,\quad f_X(x)>0$

امید ریاضی E :

اگر X یک متغیر تصادفی گسسته و f تابع احتمال آن، آنگاه :

$$1)\ E(X) = \sum_{x} x f(x)$$

اگر X یک متغیر تصادفی پیوسته و f تابع چگالی آن، آنگاه :

$$2)\ E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

قضیه ۱:
اگر X متغیر تصادفی گسسته و $f(x)$ مقدار توزیع احتمال آن به ازای X باشد، آنگاه :

$$3)\ E(g(x)) = \sum_{x} g(x) f(x)$$

اگر X متغیر تصادفی پیوسته و $f(x)$ تابع چگالی احتمال آن به ازای X باشد، آنگاه :

$$4)\ E(g(x)) = \int_{-\infty}^{+\infty} g(x) f(x)\, dx$$

قضیه ۲: اگر a و b مقادیر ثابتی باشند، آنگاه :

$$5)\ E(aX+b) = aE(X) + b$$

قضیه ۳: اگر C_1، C_2، ...، C_n مقادیری ثابت باشند آنگاه :

$$6)\ E\left[\sum_{i=1}^{n} c_i g_i(X)\right] = \sum_{i=1}^{n} c_i E[g_i(X)]$$

نکته :

$$7)\ E[(aX+b)^n] = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^{i} E(X^{n-i})$$

قضیه ۴: اگر X و Y متغیرهای تصادفی گسسته بوده و $f(x,y)$ مقدار توزیع احتمال توأم آنها در (x,y) باشند، مقدار امید ریاضی متغیر تصادفی $g(X,Y)$ برابر است با :

$$8)\ E[g(X,Y)] = \sum_{x}\sum_{y} g(x,y) f(x,y)$$

اگر X و Y متغیرهای تصادفی پیوسته بوده و $f(x,y)$ مقدار چگالی توأم آن در (x,y) باشند، مقدار امید ریاضی متغیر تصادفی $g(X,Y)$ برابر است با :

$$9)\ E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f(x,y)\, dx\, dy$$

قضیه ۵ : اگر C_1 و C_2 و ... و C_n مقادیری ثابت باشند آنگاه :

$$10)\ E\left[\sum c_i g_i(X_1, X_2, \dots, X_k)\right] = \sum_{i=1}^{n} c_i E[g_i(X_1, X_2, \dots, X_k)]$$

گشتاورها :

تعریف : rامین گشتاور حول مبدأ متغیر تصادفی X که با μ'_r نشان داده می‌شود امید ریاضی متغیر تصادفی X^r است که $r = 0, 1, 2, \dots$

اگر X متغیر تصادفی گسسته باشد آنگاه :

$$11)\ \mu'_r = E(X^r) = \sum_{x} x^r f(x)$$

اگر X متغیر تصادفی پیوسته باشد آنگاه :

$$12)\ \mu'_r = E(X^r) = \int_{-\infty}^{+\infty} x^r f(x)\, dx$$

نکته: μ'_1 میانگین توزیع X یا صرفاً میانگین X نامیده می‌شود و با μ نمایش می‌دهیم.

تعریف: گشتاور rام حول میانگین متغیر تصادفی X که آن را با μ_r نشان می‌دهیم، مقدار امید ریاضی $(X-\mu)^r$ است برای $r=0,1,2,\dots$

اگر X متغیر تصادفی گسسته باشد، آنگاه:

$$۱۳)\ \mu_r = E[(X-\mu)^r] = \sum_x (x-\mu)^r f(x)$$

اگر X متغیر تصادفی پیوسته باشد، آنگاه:

$$۱۴)\ \mu_r = E[(X-\mu)^r] = \int_{-\infty}^{+\infty} (x-\mu)^r f(x)\,dx$$

نکته: $\mu_0 = 1$ و $\mu_1 = 0$

نکته: μ_2 را واریانس توزیع X گویند و آن را با σ_X^2 یا $Var(X)$ یا $V(X)$ نشان می‌دهند.

قضیه ۶:

$$۱۵)\ \sigma^2 = \mu'_2 - \mu^2$$

اثبات: μ_2 برابر واریانس X است لذا:

$$\sigma^2 = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E(X^2) - 2E(\mu X) + \mu^2$$

$$\sigma^2 = E(X^2) - 2\mu E(X) + \mu^2 = \mu'_2 - 2\mu\mu + \mu^2 = \mu'_2 - \mu^2$$

قضیه ۷: اگر واریانس X برابر σ^2 باشد آنگاه:

$$۱۶)\ Var(aX+b) = a^2 Var(X)$$

قضیه ۸: (قضیه چبیشف)

اگر μ و σ به ترتیب میانگین و انحراف معیار متغیر تصادفی X باشند، آنگاه برای هر ثابت مثبت k احتمال اینکه X مقداری با فاصله کمتر از $k\sigma$ از میانگین اختیار کند حداقل $1-\frac{1}{k^2}$ است.

$$۱۷)\ P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

تعریف: تابع مولد گشتاورهای متغیر تصادفی X درصورت وجود اگر X گسسته باشد، آنگاه

t: عدد حقیقی

$$۱۸)\ M_X(t) = E(e^{tX}) = \sum_x e^{tx} f(x)$$

و اگر X پیوسته باشد، آنگاه، با تابع مقابل نمایش داده می‌شود.

$$۱۹)\ M_X(t) = E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tx} f(x)\,dx$$

$$۲۰)\ M_X(t) = \sum_x \left[1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \dots + \frac{t^rx^r}{r!} + \dots\right] f(x)$$

$$= \sum_x f(x) + t\sum_x x f(x) + \frac{t^2}{2!}\sum_x x^2 f(x) + \dots + \frac{t^r}{r!}\sum_x x^r f(x) + \dots$$

$$= 1 + \mu t + \mu'_2 \frac{t^2}{2!} + \dots + \mu'_r \frac{t^r}{r!} + \dots$$

پس ضریب $\frac{t^r}{r!}$ همان μ'_r است.

قضیه ۹:

$$۲۱)\ \left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0} = \mu'_r$$

اگر تابعی به صورت سری توانی برحسب t بسط داده شود،

ضریب $\frac{t^r}{r!}$، مشتق مرتبه‌ی rام تابع نسبت به t به ازای $t=0$ است.

قضیه ۱۰: اگر a و b دو مقدار ثابت باشند آنگاه:

۲۲) $M_{X+a}(t) = E\left[e^{(X+a)t}\right] = e^{at} M_X(t)$

۲۳) $M_{bX}(t) = E\left(e^{bXt}\right) = M_X(bt)$

۲۴) $M_{\frac{X+a}{b}}(t) = E\left[e^{\left(\frac{X+a}{b}\right)t}\right] = e^{\left(\frac{a}{b}\right)t} M_X\left(\frac{t}{b}\right)$

نکته: قسمت اول این قضیه، (۱۰) وقتی $a=-\mu$، اهمیت خاصی دارد و قسمت سوم (۲۴) وقتی $a=-\mu$ و $b=\sigma$ نیز دارای اهمیت خاصی است. در این حالت:

۲۵) $M_{\frac{X-\mu}{\sigma}}(t) = e^{\frac{-\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right)$

تعریف: rامین و sامین گشتاور حاصلضرب حول مبدأ متغیرهای تصادفی X و Y که با $\mu'_{r,s}$ نشان داده می‌شود، مقدار مورد انتظار X^rY^s است؛ به صورت نمادی، برای $r=0,1,2,3,\ldots$ و $s=0,1,2,3,\ldots$

با X و Y گسسته

۲۶) $\mu'_{r,s} = E\left(X^rY^s\right) = \sum_x \sum_y x^r y^s f(x,y)$

و با X و Y پیوسته

۲۷) $\mu'_{r,s} = E\left(X^rY^s\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^r y^s f(x,y)\,dx\,dy$

نکته:

۲۸) $\mu'_{0,1} = E(Y) = \mu_Y$ ۲۹) $\mu'_{1,0} = E(X) = \mu_X$

تعریف: rامین و sامین گشتاور حاصلضرب حول مبدأ متغیر تصادفی X و Y که با $\mu_{r,s}$ نشان داده می‌شود، مقدار مورد انتظار $(X-\mu_X)^r(Y-\mu_Y)^s$ است که $r=0,1,2,3,\ldots$ و $s=0,1,2,3,\ldots$ که

برای X و Y گسسته به صورت

۳۰) $\mu_{r,s} = E\left[(X-\mu_X)^r(Y-\mu_Y)^s\right] = \sum_x\sum_y (X-\mu_X)^r(Y-\mu_Y)^s f(x,y)$

و برای X و Y پیوسته به صورت

۳۱) $\mu_{r,s} = E\left[(X-\mu_X)^r(Y-\mu_Y)^s\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (X-\mu)^r(Y-\mu)^s f(x,y)\,dx\,dy$

تعریف: $\mu_{1,1}$ را کوواریانس X و Y می‌نامند و آن را با σ_{XY}، $\operatorname{cov}(X,Y)$ یا $C(X,Y)$ نشان می‌دهند.

قضیه ۱۱:

۳۲) $\sigma_{XY} = \mu'_{1,1} - \mu_X\mu_Y$

اثبات:

$\sigma_{XY} = E\left[(X-\mu_X)(Y-\mu_Y)\right] = E\left[XY - \mu_X Y - X\mu_Y + \mu_X\mu_Y\right]$

$= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X\mu_Y$

$= E[XY] - \cancel{\mu_X\mu_Y} - \mu_Y\mu_X + \cancel{\mu_X\mu_Y} = \mu'_{1,1} - \mu_X\mu_Y$

قضیه ۱۲ : اگر X و Y مستقل باشند، آنگاه

$$۳۳)\begin{cases} E(XY)=E(X)E(Y) \\ \sigma_{XY}=0 \end{cases}$$

اثبات :

بنابه تعریف برای حالت گسسته داریم : $E(XY)=\sum_x\sum_y xy\, f(x,y) \xrightarrow[\text{مستقل اند}]{\text{چون } X \text{ و } Y} f(x,y)=g(x)\,h(y)$

که در آن $g(x)$ و $h(y)$ به ترتیب مقادیر توزیع های حاشیه ای X و Y هستند.

$$E(XY)=\sum_x\sum_y xy\, g(x)h(y)=\Big[\sum_x x\,g(x)\Big]\Big[\sum_y y\,h(y)\Big]=E(X)E(Y)$$

بنابراین

$$\sigma_{XY}=\mu'_{1,1}-\mu_X\mu_Y=E(XY)-E(X)E(Y)=0$$

نکته : استقلال دو متغیر تصادفی، صفر بودن کوواریانس را ایجاب می کند، اما صفر بودن کوواریانس الزاماً استقلال آنها را نتیجه نمی دهد

قضیه ۱۳ : اگر X_1، X_2، ...، X_n مستقل باشند، آنگاه

$$۳۴)\ E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n)$$

قضیه ۱۴ : اگر X_1، X_2، ...، X_n متغیرهای تصادفی و $Y=\sum_{i=1}^{n}a_iX_i$ باشند که در آن a_1، a_2، ...، a_n مقادیر ثابت اند آنگاه :

$$۳۵)\begin{cases} E(Y)=\sum_{i=1}^{n}a_i\,E(X_i) \\ var(Y)=\sum_{i=1}^{n}a_i^2\,var(X_i)+2\sum\sum_{i<j}a_ia_j\,cov(X_i,X_j) \end{cases}$$

که در آن مجموع یابی دوگانه روی تمام مقادیر i و j از ۱ تا n با شرط $i<j$ انجام می شود.

نتیجه : اگر متغیرهای تصادفی X_1، X_2، ...، X_n مستقل باشند و $Y=\sum_{i=1}^{n}a_iX_i$ آنگاه :

$$۳۶)\ var(Y)=\sum_{i=1}^{n}a_i^2\,var(X_i)$$

قضیه ۱۵ : اگر X_1، X_2، ...، X_n متغیرهای تصادفی باشند و

$Y_1=\sum_{i=1}^{n}a_iX_i$ و $Y_2=\sum_{i=1}^{n}b_iX_i$ که در آنها a_1، a_2، ...، a_n و b_1، b_2، ...، b_n مقادیر ثابت اند، آنگاه :

$$۳۷)\ cov(Y_1,Y_2)=\sum_{i=1}^{n}a_ib_i\,var(X_i)+\sum\sum_{i<j}(a_ib_j+a_jb_i)\,cov(X_i,X_j)$$

نتیجه : اگر متغیرهای تصادفی X_1، X_2، ...، X_n مستقل باشند، $Y_1=\sum_{i=1}^{n}a_iX_i$ و $Y_2=\sum_{i=1}^{n}b_iX_i$، آنگاه :

$$۳۸)\ cov(Y_1,Y_2)=\sum_{i=1}^{n}a_ib_i\,var(X_i)$$

فرمولهای فصل ششم

آمار و احتمال

توزیع یکنواخت گسسته :

اگر یک متغیر تصادفی بتواند k مقدار مختلف را با احتمال های برابر اختیار کند، گوییم که دارای توزیع یکنواخت گسسته است؛ اگر و تنها اگر توزیع احتمال آن، برای $x_i \neq x_j$ وقتی $i \neq j$ به صورت

۱) $f(x) = \dfrac{1}{k}$ $\quad x = x_1, x_2, \ldots, x_k$

توزیع برنولی :

اگر آزمایشی دو برآمد داشته باشد، «موفقیت» و «شکست»، و احتمال آنها به ترتیب θ و $1-\theta$ باشند، آنگاه متغیر تصادفی که منطبق با دو وضعیت، یا ۱ و ۰ را اختیار می کند، توزیع برنولی دارد اگر و فقط اگر توزیع احتمال آن به صورت زیر باشد :

۲) $f(x;\theta) = \theta^{x}(1-\theta)^{1-x}$ $\quad$ و $x = 0, 1$

توزیع دوجمله‌ای :

متغیر تصادفی X توزیع دوجمله ای دارد و به آن متغیر تصادفی دوجمله‌ای داده می شود، اگر و تنها اگر توزیع احتمال آن به صورت زیر باشد :

۳) $b(x;n,\theta) = \binom{n}{x}\theta^{x}(1-\theta)^{n-x}$ $\quad$ و $x = 0, 1, 2, \ldots, n$

که در آن n تعداد آزمایش ها و x تعداد موفقیت هاست، در واقع هدف X موفقیت در n امتحان است.

نکته : $B(x;n,\theta)$ به صورت زیر تعریف می شود :

۴) $B(x;n,\theta) = \sum_{k=0}^{n} b(k;n,\theta)$

قضیه ۱ :

۵) $b(x;n,\theta) = b(n-x;n,1-\theta)$

قضیه ۲ : میانگین و واریانس توزیع دوجمله ای برابرند با :

۶) $\sigma^2 = n\theta(1-\theta)$ $\quad$ ، ۷) $\mu = n\theta$

قضیه ۳ : اگر X توزیع دوجمله‌ای با پارامتر n و θ داشته باشد و $Y = \dfrac{X}{n}$ آنگاه :

۸) $E(Y) = \theta$ $\quad$ ، $\sigma_Y^2 = \dfrac{\theta(1-\theta)}{n}$

قضیه ۴ : تابع مولد گشتاور توزیع دوجمله ای به صورت :

۹) $M_X(t) = [1+\theta(e^t-1)]^n$

توزیع فوق هندسی :

جامعه با N عضو، را به دو قسمت k و $N-k$ عضوی تقسیم می کنیم و یک نمونه n تایی از آن انتخاب می کنیم.

X : تعداد نمونه هایی که متعلق به قسمت k تایی هستند.

۱۰) $h(x;n,k) = \dfrac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$

قضیه ۵ : میانگین و واریانس توزیع فوق هندسی عبارتست از :

۱۱) $\mu = \dfrac{nk}{N}$

۱۲) $\sigma^2 = \dfrac{nk(N-k)(N-n)}{N^2(N-1)}$

توزیع پواسن : پواسن آزمایشی است که تعداد وقایع (تعداد پیروزی ها) را در یک فاصله زمانی یا مکانی به دست می آورد.

x : تعداد اتفاقات در یک فاصله زمانی یا مکانی .

λ : متوسط تعداد اتفاقات در یک فاصله زمانی یا مکانی .

۱۳) $P(x,\lambda)=\dfrac{\lambda^x e^{-\lambda}}{x!}$; $x=0,1,2,\ldots$

قضیه ۶ : میانگین و واریانس توزیع پواسن عبارتند از :

۱۴) $\mu=E(x)=\lambda$ ۱۵) $\sigma^2=\lambda$

قضیه ۷ : تابع مولد گشتاورهای توزیع پواسن به صورت :

۱۶) $M_X(t)=e^{\lambda(e^t-1)}$

تعریف : متغیر تصادفی X دارای چگالی یکنواخت است اگر و فقط اگر چگالی احتمالی آن به صورت :

پارامترهای α و β ثابت های حقیقی اند و $\alpha<\beta$.

۱۷) $f(x)=\begin{cases}\dfrac{1}{\beta-\alpha} & \alpha<x<\beta\\ 0 & \text{سایر جاها}\end{cases}$

قضیه ۸ : میانگین و واریانس چگالی یکنواخت عبارتند از :

۱۸) $\mu=\dfrac{\alpha+\beta}{2}$ ۱۹) $\sigma^2=\dfrac{(\beta-\alpha)^2}{12}$

توزیع نرمال : توزیع نرمال یا توزیع زنگی به عنوان یک توزیع متقارن پیوسته بوده و دارای کاربردهای فراوانی در زمینه‌های مختلف از جمله موارد زیر می باشد.

الف- بسیاری از پدیده‌های طبیعی دارای توزیع نرمال هستند. ب- بسیاری از توزیع ها در شکل حدی دارای تقریب نرمال هستند.

در صورتی که متغیر تصادفی پیوسته x دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشد. تابع چگالی و مشخصات آن به صورت مقابل است :

۲۰) $n(x;\mu,\sigma)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ $-\infty<x<+\infty$

قضیه ۹ : تابع مولد گشتاورهای توزیع نرمال عبارت است از :

۲۱) $M_X(t)=e^{\mu t+\frac{1}{2}t^2\sigma^2}$ $-\infty<x<+\infty$

تقریب نرمال برای توزیع دوجمله ای : اگر X متغیر تصادفی با توزیع دوجمله ای با پارامترهای n و θ باشد آنگاه تابع مولد گشتاورهای وقتی $n\to+\infty$ به تابع مولد گشتاورهای توزیع نرمال استاندارد میل می کند.

۲۲) $Z=\dfrac{X-n\theta}{\sqrt{n\theta(1-\theta)}}$

توزیع نرمال دو متغیره :

جهت متغیرهای تصادفی X و Y دارای توزیع نرمال دو متغیره اند و به آن ها متغیرهای تصادفی که توأماً به صورت نرمال توزیع شده اند اطلاق می شوند اگر و فقط اگر چگالی احتمال توأم‌شان برای $-\infty<x<+\infty$ و $-\infty<y<+\infty$ به صورت زیر باشد

۲۳) $f(x,y)=\dfrac{e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$ $\sigma_1>0$ و $\sigma_2>0$ ، $-1<\rho<1$

قضیه ۱۰ : دو متغیر تصادفی که دارای توزیع نرمال دو متغیره اند مستقل اند اگر و فقط اگر $\rho=0$.