خودآزمایی فصل اوّل

۱ـ جامعه و نمونه را تعریف کنید و مشخص کنید کدام یک از موارد زیر جامعه و کدام یک نمونه را توصیف می‌کنند.

الف) فارغ‌التحصیلان رشته‌ی مهندسی

ب) فارغ‌التحصیلان مهندسی نرم‌افزار در سال ۱۳۸۳

ج) دانش‌آموزان گروه سنی ۱۴ تا ۱۶ در سرشماری ۱۳۷۵

د) دانش‌آموزان ممتاز گروه سنی ۱۴ تا ۱۶ در سرشماری ۱۳۷۵

ه) رایانه‌های از رده خارج شده تا سال ۱۳۸۳

پاسخ: جامعه: جامعه از لحاظ آمار به مجموعه‌ای از افراد یا اشیایی که لااقل دارای یک صفت مشخص کننده باشند اطلاق می‌شود. جامعه‌ی آماری ممکن است متناهی (محدود) یا نامتناهی باشد.

نمونه: هر بخشی از جامعه‌ی آماری را نمونه گویند.

الف) جامعه ب) نمونه ج) جامعه د) نمونه ه) جامعه

۲ـ نشان دهید:

$\sum_{i=1}^{3} i^2 = 14$ $\sum_{i=1}^{3} i = 6$

$\sum_{i=1}^{7} (x_i + 2) = \sum_{i=1}^{7} x_i + 14$ $\sum_{i=1}^{3} i^3 = 36$

$\sum_{i=1}^{5} i = 15$ $\sum_{i=1}^{3} (2i - 2) = 4$

$\sum_{i=1}^{3} (i+1)^{i+1} = 76$ $\sum_{i=-1}^{3} i = 5$

$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$ $\sum_{i=1}^{n} \frac{i}{2} = \frac{n(n+1)}{4}$

یادآوری: عملگر جمع به صورت زیر تعریف می‌شود: $\sum_{i=1}^{n} X_i = X_1 + X_2 + \dots + X_n$

که تعدادی از خواص آن عبارتند از:

۱) $\sum_{i=1}^{n} k = k + k + \dots + k = nk$ وقتی که k ثابت است.

۲) $\sum_{i=1}^{n} (X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$

۳) $\sum_{i=1}^{n} (A + BX_i) = nA + B\sum_{i=1}^{n} X_i$ که در آن A و B ثابت هستند

پاسخ:

$$\sum_{i=1}^{3} i = 6 \rightarrow \sum_{i=1}^{3} i = 1+2+3 = 6$$

$$\sum_{i=1}^{3} i^2 = 14 \rightarrow \sum_{i=1}^{3} i^2 = (1)^2 + (2)^2 + (3)^2 = 1+4+9 = 14$$

$$\sum_{i=1}^{3} i^3 = 36 \rightarrow \sum_{i=1}^{3} i^3 = (1)^3 + (2)^3 + (3)^3 = 1+8+27 = 36$$

$$\sum_{i=1}^{7} (x_i + 2) = \sum_{i=1}^{7} x_i + 14 \rightarrow \sum_{i=1}^{7} (x_i + 2) = \xrightarrow{\text{با استفاده از خاصیت (۲)}} \sum_{i=1}^{7} x_i + \sum_{i=1}^{7} 2$$

$$= \xrightarrow{\text{با استفاده از خاصیت (۱)}} \sum_{i=1}^{7} x_i + (7 \times 2) = \sum_{i=1}^{7} x_i + 14$$

تذکر: در این قسمت $\sum_{i=1}^{3} (2i-2) = 6$ صحیح می‌باشد.

$$\sum_{i=1}^{3} (2i-2) = 6 \xrightarrow{\text{روش اول}} \sum_{i=1}^{3} (2i-2) = (2\times 1 - 2) + (2 \times 2 - 2) + (2 \times 3 - 2)$$

$$= 0 + 2 + 4 = 6$$

$$\sum_{i=1}^{3} (2i-2) = 6 \xrightarrow{\text{روش دوم}} \sum_{i=1}^{3} (2i-2) = \sum_{i=1}^{3} 2i - \sum_{i=1}^{3} 2 = (2+4+6) - (2 \times 3)$$

$$= 12 - 6 = 6$$

$$\sum_{i=1}^{5} i = 15 \rightarrow \sum_{i=1}^{5} i = 0+1+2+3+4+5 = 15$$

$$\sum_{i=-1}^{3} i = 5 \rightarrow \sum_{i=-1}^{3} i = -1+0+1+2+3 = 5$$

تذکر: در این قسمت $\sum_{i=1}^{3} (i+1)^{i+1} = 287$ صحیح می‌باشد.

$$\sum_{i=1}^{3} (i+1)^{i+1} = 278 \rightarrow \sum_{i=1}^{3} (i+1)^{i+1} = (1+1)^{1+1} + (2+1)^{2+1} + (3+1)^{3+1}$$

$$= (2)^2 + (3)^3 + (4)^4 = 4 + 27 + 256 = 287$$

$$\sum_{i=1}^{n} \frac{i}{2} = \frac{n(n+1)}{4}$$

روش اول: برای ساده شدن محاسبات ابتدا با استفاده از خاصیت سیگما داریم:

$$\sum_{i=1}^{n} \frac{i}{2} = \frac{1}{2} \sum_{i=1}^{n} i = \frac{n(n+1)}{4} = \frac{1}{2} \times \frac{n(n+1)}{2}$$

ابتدا نشان می‌دهیم:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

(a) $\sum_{i=1}^{n} i = 1+2+3+\dots+n$

حال یک بار دیگر، مجموع بالا را می‌نویسیم ولی این بار از بزرگ به کوچک

(b) $\sum_{i=1}^{n} i = n+(n-1)+(n-2)+\dots+1$

با جمع دو رابطه‌ی (a) و (b) داریم:

$$2\sum_{i=1}^{n} i = (n+1)+(n+1)+\dots+(n+1)$$

$$2\sum_{i=1}^{n} i = n(n+1) \to \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

حال با به دست آمدن $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ داریم:

$$\sum_{i=1}^{n} \frac{i}{2} = \frac{1}{2}\sum_{i=1}^{n} i = \frac{1}{2}\times\frac{n(n+1)}{2} = \frac{n(n+1)}{4}$$

یادآوری: استقرای ریاضی: روشی دیگر برای اثبات یک حکم ریاضی روش استقرای ریاضی می‌باشد که در ذیل به معرفی اصل استقرای ریاضی می‌پردازیم:

اصل استقرای ریاضی: اگر $P(n)$ یک گزاره‌ی مربوط به عدد طبیعی n باشد به طوری که:

الف) $P(1)$ درست باشد.

ب) برای $k\in N$ داشته باشیم $P(k) \Rightarrow P(k+1)$

آن گاه $P(n)$ برای هر عدد طبیعی n درست است.

روش دوم: با استفاده از استقرای ریاضی:

$$\sum_{i=1}^{n} \frac{i}{2} = \frac{1}{2}\sum_{i=1}^{n} i = \frac{1}{2}\times\frac{n(n+1)}{2}$$

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

$$1+2+3+\dots+n = \frac{n(n+1)}{2}$$

جمله‌ی اول $P(1): 1 = \frac{1(1+1)}{2} \Rightarrow 1=1$

فرض استقرا $P(k): 1+2+\dots+k = \frac{k(k+1)}{2}$

حکم استقرا $P(k+1): 1+2+\dots+k+(k+1) = \frac{(k+1)((k+1)+1)}{2}$

اثبات $= \frac{k(k+1)}{2} + (k+1) = \frac{k(k+1)+2(k+1)}{2} = \frac{(k+1)(k+2)}{2}$

حکم برقرار است.

* $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

با استفاده از استقرای ریاضی:

$$1^2+2^2+\dots+n^2 = \frac{n(n+1)(2n+1)}{6}$$

جمله‌ی اول $P(1): 1^2 = \frac{1(1+1)(2\times1+1)}{6} = \frac{1(2)(3)}{6} \Rightarrow 1=1$

فرض استقرا $P(k): 1^2+2^2+\dots+k^2 = \frac{k(k+1)(2k+1)}{6}$

حکم استقرا $P(k+1): 1^2+2^2+\dots+k^2+(k+1)^2 = \frac{(k+1)(k+2)(2k+3)}{6}$

اثبات $= \frac{k(k+1)(2k+1)}{6} + (k+1)^2 = \frac{k(k+1)(2k+1)+6(k+1)^2}{6}$

$$= \frac{(k+1)[2k^2+k+6k+6]}{6} = \frac{(k+1)[2k^2+7k+6]}{6}$$

$$= \frac{(k+1)(k+2)(2k+3)}{6}$$

حکم برقرار است.

۳ـ دستمزد ۱۵ کارگر در ادارات مختلف عبارتند از:

۱۱/۶۳	۸/۲۲	۱۲/۵۴	۲۹/۲۳	۱۸/۲۳	۱۱/۴۹	۱۱/۳۰	۱۷/۰۰
۹/۱۶	۸/۶۴	۲۷/۵۶	۸/۲۳	۱۹/۷۷	۱۲/۸۱		

میانگین حسابی، هندسی و هارمونیک دستمزد کارگران را به دست آورید.

پاسخ: میانگین حسابی: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1+x_2+\dots+x_n)$

$$\bar{x} = \frac{1}{15}\sum_{i=1}^{15} x_i = \frac{1}{15}(11.63+8.22+12.54+\dots+19.77+12.81) = \frac{217.95}{15} = 14.53$$

میانگین هندسی: $G = \sqrt[n]{x_1.x_2.\dots.x_n} = \sqrt[n]{\prod_{i=1}^{n} x_i}$

تذکر: چون تعداد اعداد برای محاسبه‌ی میانگین هندسی زیاد است برای سهولت در محاسبه‌ی این میانگین می‌توان از فرمول زیر استفاده کرد:

$$G = 10^{\frac{1}{n}\sum_{i=1}^{n} \log x_i}$$

ابتدا باید $\log x_i$ها را برای ۱۵ داده‌ی بالا به دست آورد.

$\log 11.63=1.066$	$\log 8.22=0.915$	$\log 12.54=1.098$
$\log 12.14=1.084$	$\log 29.23=1.466$	$\log 18.23=1.261$
$\log 11.49=1.060$	$\log 11.30=1.053$	$\log 17.00=1.230$
$\log 9.16=0.962$	$\log 8.64=0.937$	$\log 27.56=1.440$
$\log 8.23=0.915$	$\log 19.77=1.296$	$\log 12.81=1.108$

سپس مقدار توان میانگین هندسی را به دست می‌آوریم:

$$\frac{1}{n}\sum_{i=1}^{n}\log x_i=\frac{1}{15}[1/066+0/915+1/098+\ldots+1/296+1/108]=\frac{16/891}{15}=1/126$$

$$G=10^{\frac{1}{n}\sum_{i=1}^{n}\log x_i}\rightarrow G=10^{1/126}=13/366$$

میانگین هارمونیک:

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

$$H=\frac{15}{\left[\frac{1}{11/63}+\frac{1}{8/22}+\frac{1}{12/54}+\ldots+\frac{1}{19/77}+\frac{1}{12/81}\right]}$$

$$=\frac{15}{[0/086+0/122+0/080+\ldots+0/051+0/078]}=\frac{15}{1/205}=12/448$$

همان طور که مشاهده می‌شود:

میانگین حسابی ≤ میانگین هندسی ≤ میانگین هارمونیک

$$12/448\le 13/366\le 14/53$$

۴ـ جدول زیر مصرف گاز مایع و رشد سالانه‌ی آن را از سال ۱۳۶۳ تا سال ۱۳۶۷ در کشور نشان می‌دهد. رشد متوسط را طی این دوره محاسبه کنید.

سال	۱۳۶۳	۱۳۶۴	۱۳۶۵	۱۳۶۶	۱۳۶۷
مصرف	**۸۹۵**	**۱۰۵۴**	**۱۰۸۳**	**۱۰۷۶**	**۱۱۹۷**
رشد سالانه	**—**	**۱/۱۸**	**۱/۰۳**	**۰/۹۹**	**۱/۱۱**
درصد رشد سالانه	—	۱۱۸	۱۰۳	۹۹	۱۱۱

(رشد سالانه از تقسیم مصرف هر سال به سال قبل به دست می‌آید)

پاسخ: روش اول:

$$G=\sqrt[n]{x_1.x_2.\ldots.x_n}=\sqrt[n]{\prod_{i=1}^{n}x_i}$$

برای به دست آوردن میانگین هندسی، مقادیر به دست آمده از رشد سالانه را در ۱۰۰ ضرب می‌کنیم تا به صورت درصد باشد پس داریم:

$$G=\sqrt[4]{(1/18\times 100)\times(1/03\times 100)\times(0/99\times 100)\times(1/11\times 100)}$$

$$G=\sqrt[4]{(1/18\times 1/03\times 0/99\times 1/11)\times 100^4}=\sqrt[4]{1/336\times 100^4}=1/075\times 100=107/5$$

روش دوم: همچنین می‌توان نشان داد که رشد متوسط اخیر از رابطه‌ی زیر نیز قابل محاسبه است:

$$G=\sqrt[n-1]{\frac{x_n}{x_1}}\times 100$$

$$G=\sqrt[4]{\frac{1197}{895}}\times 100=\sqrt[4]{1/337}\times 100=1/075\times 100=107/5$$

۵ـ اگر $\overline{x}$ میانگین حسابی، G میانگین هندسی و H میانگین هارمونیک برای یک نمونه به حجم n باشند. نشان دهید که رابطه‌ی زیر همواره برقرار است.

$$H\le G\le\overline{x}$$

در چه حالت $H=G=\overline{x}$ است؟

پاسخ: برای اثبات این نامساوی باید نشان داد که میانگین ریشه‌ای رتبه‌ی r یعنی:

$$M_r=\left[\sum_{i=1}^{k}\frac{f_i x_i^r}{n}\right]^{\frac{1}{r}}$$

تابعی صعودی از r است. سپس با قرار دادن r برابر ۱-، ۰، ۱ در M_r نامساوی بالا به دست می‌آید.

همچنین با توجه به رابطه‌ی $\frac{1}{n}\sum_{i=1}^{k}f_i x_i\ge\sqrt[n]{x_1^{f_1}.x_2^{f_2}.\ldots.x_k^{f_k}}\ge\frac{n}{\sum_{i=1}^{k}\frac{f_i}{x_i}}$ رابطه‌ی $H\le G\le\overline{x}$ برقرار است.

رابطه‌ی $H\le G\le\overline{x}$ هنگامی به رابطه‌ی $H=G=\overline{x}$ تبدیل می‌شود که تمامی داده‌ها دارای مقدار یکسانی باشند.

۶ـ اگر در سفر چهار ساعته از تهران به سمنان با سرعت ۵۰km/h در ساعت اول، ۶۵km/h در ساعت دوم، ۸۰km/h در ساعت سوم و ۵۵km/h در ساعت چهارم طی مسیر شده باشد میانگین هارمونیک را به دست آورید.

پاسخ: میانگین هارمونیک:

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

$$H=\frac{4}{\sum_{i=1}^{4}\frac{1}{x_i}}=\frac{4}{\frac{1}{50}+\frac{1}{65}+\frac{1}{80}+\frac{1}{55}}=\frac{4}{0/020+0/015+0/013+0/018}=\frac{4}{0/066}=60/606$$

۷ـ میانگین حسابی و هندسی دو عدد به ترتیب ۱۳ و ۱۲ است. آن دو عدد را پیدا کنید و میانگین هارمونیک آن‌ها را به دست آورید.

پاسخ: اگر دو عدد را x_1 و x_2 در نظر بگیریم، میانگین حسابی دو داده $\overline{x}=13$ و میانگین هندسی آن $G=12$ می‌باشد.

رابطه‌ی (۱) $$\overline{x}=\frac{x_1+x_2}{2}\rightarrow 13=\frac{x_1+x_2}{2}\rightarrow x_1+x_2=26$$

رابطه‌ی (۲) $$G=\sqrt[2]{x_1.x_2}\rightarrow G=(x_1.x_2)^{\frac{1}{2}}\rightarrow 12=(x_1.x_2)^{\frac{1}{2}}$$

هر دو طرف رابطه‌ی (۲) را به توان ۲ می‌رسانیم دو معادله به صورت زیر ساده می‌شود:

$$\begin{cases}x_1+x_2=26 & \text{(1)}\\ x_1.x_2=144 & \text{(2)}\end{cases}$$

رابطه‌ی (۱)
رابطه‌ی (۲)

برای به دست آوردن مقادیر $x_۱$ و $x_۲$ و حل معادلات، معادله را به صورت یک مجهولی در آورده و به صورت زیر می‌نویسیم:

$x_۱ + x_۲ = ۲۶ \rightarrow x_۱ = ۲۶ - x_۲$ رابطه‌ی *

رابطه‌ی * را در رابطه‌ی (۲) جایگزین می‌کنیم پس داریم:

$x_۱ . x_۲ = ۱۴۴$ رابطه‌ی (۲)

$\rightarrow (۲۶ - x_۲) . x_۲ = ۱۴۴$ با استفاده از رابطه‌ی *

$\rightarrow ۲۶x_۲ - x_۲^۲ = ۱۴۴ \rightarrow x_۲^۲ - ۲۶x_۲ + ۱۴۴ = ۰$

معادله‌ی بالا را به روش دلتا محاسبه می‌کنیم تا ریشه‌های معادله به دست آید.

$b = -۲۶ \qquad a = ۱ \qquad c = ۱۴۴$

$\Delta = b^۲ - ۴ac = (-۲۶)^۲ - ۴(۱)(۱۴۴) = ۶۷۶ - ۵۷۶ = ۱۰۰$

چون مقدار $\Delta > ۰$ است پس معادله دو ریشه‌ی حقیقی دارد.

$$x_۱, x_۲ = \frac{-b \pm \sqrt{\Delta}}{۲a} = \frac{+۲۶ \pm ۱۰}{۲} \begin{cases} x_۱ = \frac{۲۶ + ۱۰}{۲} = ۱۸ \\ x_۲ = \frac{۲۶ - ۱۰}{۲} = ۸ \end{cases}$$

که اگر $x_۱$ و $x_۲$ را در رابطه‌های (۱) و (۲) قرار دهیم در معادلات صدق می‌کند.

با مشخص شدن $x_۱ = ۱۸$ و $x_۲ = ۸$ می‌توان میانگین هارمونیک دو عدد را به دست آورد.

$$H = \frac{n}{\sum_{i=۱}^{n} \frac{۱}{x_i}} \rightarrow H = \frac{۲}{\sum_{i=۱}^{۲} \frac{۱}{x_i}} = \frac{۲}{\frac{۱}{۱۸} + \frac{۱}{۸}} = \frac{۲}{۰/۰۵۶ + ۰/۱۲۵} = \frac{۲}{۰/۱۸۱} = ۱۱/۰۵۰$$

یادآوری: برای حل معادلات درجه‌ی دوم $ax^۲ + bx + c = ۰$ به روش دلتا به صورت زیر عمل می‌کنیم:

ابتدا مقدار Δ را از فرمول زیر به دست می‌آوریم:

$\Delta = b^۲ - ۴ac$

الف) اگر $\Delta > ۰$ باشد معادله دارای دو ریشه‌ی حقیقی است و از فرمول زیر به دست می‌آید:

$$x_۱, x_۲ = \frac{-b \pm \sqrt{\Delta}}{۲a}$$

ب) اگر $\Delta = ۰$ باشد معادله دارای یک ریشه‌ی مضاعف می‌باشد:

$$x_۱ = x_۲ = \frac{-b}{۲a}$$

ج) اگر $\Delta < ۰$ باشد معادله ریشه‌ی حقیقی ندارد.

۸ـ میانه‌ی سری مشاهدات زیر را به دست آورید:

سری اول:

۱ ۲ ۳ ۴ ۶ ۷ ۸ ۹ ۵

۲/۳ ۳/۵ ۱۴/۱ ۱۶/۶ ۱۸/۷ ۲۱/۳ ۲۹/۷ ۳۶/۶ ۱۷/۲

سری دوم:

۱ ۳ ۴ ۶ ۷ ۵ ۲

۶۵ ۶۹ ۷۴ ۸۴ ۹۳ ۱۰۲ ۱۰۵ ۱۰۲ ۹۸ ۸۸ ۷۴ ۶۶

سری سوم:

۱۰ ۲ ۱۷ ۶ ۸ ۳ ۱۰ ۲ ۹ ۱۰ ۱ ۱۳ ۹ ۵ ۱۵

یادآوری: برای محاسبه‌ی میانه برای داده‌های گسسته، ابتدا مشاهدات را به ترتیب صعودی مرتب می‌کنیم و دو حالت زیر را در نظر می‌گیریم:

الف) اگر حجم مشاهدات فرد باشد عدد میانی، میانه‌ی مشاهدات خواهد بود.

ب) اگر حجم مشاهدات زوج باشد میانه، میانگین دو عدد میانی خواهد بود.

پاسخ: سری اول: ابتدا داده‌ها را به صورت صعودی مرتب می‌کنیم:

۲/۳ ۳/۵ ۱۴/۱ ۱۶/۶ ۱۷/۲ ۱۸/۷ ۲۱/۳ ۲۹/۷ ۳۶/۶

میانه $= ۱۷/۲ \quad , \quad n = ۹$

سری دوم: ابتدا داده‌ها را به صورت صعودی مرتب می‌کنیم:

۶۵ ۶۶ ۶۹ ۷۴ ۷۴ ۸۴ ۸۸ ۹۳ ۹۸ ۱۰۲ ۱۰۲

۱۰۵

میانه $= \frac{۸۴ + ۸۸}{۲} = ۸۶ \quad , \quad n = ۱۲$

سری سوم: ابتدا داده‌ها را به صورت صعودی مرتب می‌کنیم:

۱ ۲ ۲ ۳ ۵ ۶ ۸ ۹ ۹ ۱۰ ۱۰ ۱۰ ۱۳ ۱۵ ۱۷

میانه $= ۹ \quad , \quad n = ۱۵$

۹ـ نما یا مد سری مشاهدات زیر را به دست آورید:

سری اول: ۳ ۹ ۵ ۴ ۵ ۸ ۵ ۳

سری دوم: ۲۴ ۲۷ ۱۰ ۲۴ ۲۷ ۱۰

سری سوم: ۹ ۱۱ ۱۰ ۷ ۹ ۵ ۸ ۷ ۵

یادآوری: نمای یک مجموعه، عددی است که در آن مجموعه بیش از بقیه تکرار شده باشد. یک مجموعه از مشاهدات ممکن است یک یا چند نما داشته باشد و ممکن است فاقد نما باشد.

پاسخ: سری اول: ۳ ۹ ۵ ۴ ۵ ۸ ۵ ۳

نما = ۵

سری دوم: ۲۴ ۲۷ ۱۰ ۲۴ ۲۷ ۱۰

مجموعه فاقد نما می‌باشد چون تمام داده‌ها به یک اندازه تکرار شده‌اند.

سری سوم: ۹ ۱۱ ۱۰ ۷ ۹ ۵ ۸ ۷ ۵

نما = ۵ و ۷ و ۹

۱۰ـ چارک اول، دوم و سوم مشاهدات زیر را به دست آورید.

۱۷/۷۷ ۲۲/۸۷ ۳۰/۱۷ ۲۱/۳۶ ۴۶/۵۹ ۵۲/۲۲

یادآوری: مراحل محاسبه‌ی چارک‌ها، به ترتیب مراحل زیر را مورد استفاده قرار می‌دهیم:

۱) داده‌ها را به طور صعودی مرتب می‌کنیم.

۲) داده‌های مرتب شده را از ۱ تا n کدگذاری می‌کنیم.

۳) جایگاه چارک a ام را توسط رابطه‌ی زیر به دست می‌آوریم:

$$CQ_a = \frac{n.a}{۴} + \frac{۱}{۲} \qquad a = ۱, ۲, ۳$$

۴) با استفاده از جایگاه چارک، مقدار چارک را تعیین می‌کنیم.

پاسخ: ۱۷/۷۷ ۲۲/۸۷ ۳۰/۱۷ ۲۱/۳۶ ۴۶/۵۹ ۵۲/۲۲

داده‌ها را به صورت صعودی مرتب کرده و کدگذاری می‌کنیم:

	۱۷/۷۷	۲۱/۳۶	۲۲/۸۷	۳۰/۱۷	۴۶/۵۹	۵۲/۲۲
کدگذاری	۱	۲	۳	۴	۵	۶

جایگاه چارک اول $CQ_۱ = \frac{n.a}{۴} + \frac{۱}{۲} \qquad a = ۱ \quad , \quad n = ۶$

$$CQ_۱ = \frac{۶}{۴} + \frac{۱}{۲} = \frac{۸}{۴} = ۲$$

دومین داده، مقدار چارک اول می‌باشد. چارک اول $Q_۱ = ۲۱/۳۶$

جایگاه چارک دوم (میانه) $CQ_۲ = \frac{n.a}{۴} + \frac{۱}{۲} \qquad a = ۲ \quad , \quad n = ۶$

$$CQ_۲ = \frac{۱۲}{۴} + \frac{۱}{۲} = \frac{۱۴}{۴} = ۳/۵$$

میانگین دو داده‌ی سوم و چهارم مقدار چارک دوم می‌باشد.

چارک دوم $Q_۲ = \frac{۲۲/۸۷ + ۳۰/۱۷}{۲} = ۲۶/۵۲$

جایگاه چارک سوم $CQ_۳ = \frac{n.a}{۴} + \frac{۱}{۲} \qquad a = ۳ \quad , \quad n = ۶$

$$CQ_۳ = \frac{۱۸}{۴} + \frac{۱}{۲} = \frac{۲۰}{۴} = ۵$$

پنجمین داده، مقدار چارک سوم می‌باشد. چارک سوم $Q_۳ = ۴۶/۵۹$

۱۱ـ جمعیت ۱۸ ناحیه عبارت است از:

۱۱۳ ۱۰۰ ۹۵ ۷۵ ۸۰ ۱۰۷ ۵۷ ۶۸ ۸۳ ۷۶ ۷۷

۷۴ ۴۶ ۸۷ ۸۳ ۱۲۱ ۱۲۱ ۱۱۹

دامنه، میانگین و واریانس جمعیت نواحی را به دست آورید.

پاسخ:

دامنه : $R = x_{max} - x_{min}$

$$R = ۱۲۱ - ۴۶ = ۷۵$$

میانگین: $\bar{x} = \frac{۱}{n} \sum_{i=۱}^{n} x_i$

$$\bar{x} = \frac{۱}{۱۸}[۷۷ + ۷۶ + ۸۳ + ... + ۴۶ + ۷۴] = \frac{۱۵۸۲}{۱۸} = ۸۷/۸۸۹$$

واریانس : $S^۲ = \frac{۱}{n} \sum_{i=۱}^{n} (x_i - \bar{x})^۲$

$$S^۲ = \frac{۱}{۱۸}\left[(۷۷ - ۸۷/۸۸۹)^۲ + (۷۶ - ۸۷/۸۸۹)^۲ + ... + (۷۴ - ۸۷/۸۸۹)^۲\right] = ۴۵۴/۸۸۴$$

تذکر: اگر برای محاسبه‌ی واریانس از فرمول $S^۲ = \frac{۱}{n-۱} \sum_{i=۱}^{n} (x_i - \bar{x})^۲$ استفاده کنیم مقدار واریانس به صورت زیر است:

$$S^۲ = \frac{۱}{۱۷} \sum_{i=۱}^{۱۸} (x_i - ۸۷/۸۸۹)^۲ = ۴۸۱/۶۲۷$$

۱۲ـ توزیع سنی ۱۳۰ مورد در اولین ازدواج در جدول فراوانی زیر داده شه است. نما، میانگین، واریانس و انحراف معیار را محاسبه کنید.

سن	۱۸	۱۹	۲۰	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶	۲۷	۲۸	۲۹
تعداد	۲	۱	۴	۸	۱۰	۱۲	۱۷	۱۹	۱۸	۱۴	۱۳	۱۲

پاسخ: در داده‌های طبقه‌بندی شده در جدول، نما را طبقه‌ای در نظر می‌گیریم که بیش‌ترین فراوانی را دارند. در این مثال، نما، گروه سنی ۲۵ سال می‌باشد که فراوانی آن از همه بیش‌تر است.

سن x_i	تعداد f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^۲$	$f_i (x_i - \bar{x})^۲$
۱۸	۲	۳۶	۶/۹۲-	۴۷/۸۹	۹۵/۸۷
۱۹	۱	۱۹	۵/۹۲-	۳۵/۰۵	۳۵/۰۵
۲۰	۴	۸۰	۴/۹۲-	۲۴/۲۱	۹۶/۸۴
۲۱	۸	۱۶۸	۳/۹۲-	۱۵/۳۷	۱۲۲/۹۶
۲۲	۱۰	۲۲۰	۲/۹۲-	۸/۵۳	۸۵/۳۰
۲۳	۱۲	۲۷۶	۱/۹۲-	۳/۶۹	۴۴/۲۸
۲۴	۱۷	۴۰۸	۰/۹۲-	۰/۸۵	۱۴/۴۵
۲۵	۱۹	۴۷۵	۰/۰۸	۰/۰۱	۰/۱۹
۲۶	۱۸	۴۶۸	۱/۰۸	۱/۱۷	۲۱/۰۶
۲۷	۱۴	۳۷۸	۲/۰۸	۴/۳۳	۶۰/۶۲
۲۸	۱۳	۳۶۴	۳/۰۸	۹/۴۹	۱۲۳/۳۷
۲۹	۱۲	۳۴۸	۴/۰۸	۱۶/۶۵	۱۹۹/۸۰
	۱۳۰	۳۲۴۰	—	—	۸۹۹/۷۰

میانگین $\bar{x} = \frac{۱}{n}\sum_{i=۱}^{k} f_i x_i$

$$\bar{x} = \frac{۱}{۱۳۰}\sum_{i=۱}^{۱۲} f_i x_i = \frac{۳۲۴۰}{۱۳۰} = ۲۴/۹۲$$

واریانس $S^۲ = \frac{۱}{n}\sum_{i=۱}^{k} f_i (x_i - \bar{x})^۲$

$$S^۲ = \frac{۱}{۱۳۰}\sum_{i=۱}^{۱۲} f_i (x_i - ۲۴/۹۲)^۲ = \frac{۸۹۹/۷۰}{۱۳۰} = ۶/۹۲$$

انحراف معیار $S = \sqrt{S^۲} = \sqrt{۶/۹۲} = ۲/۶۳$

۱۳ـ قد ۹ سرباز بر حسب اینچ عبارت است از:

۷۲ ۶۸ ۶۵ ۶۳ ۶۴ ۶۹ ۶۷ ۶۶ ۶۹

قد این ۹ سرباز را استاندارد نمایید.

پاسخ: $n = ۹ \quad , \quad \bar{x} = \frac{۱}{۹}\sum_{i=۱}^{۹} x_i = ۶۷$

$$S^۲ = \frac{۱}{۹}\sum_{i=۱}^{۹}(x_i - ۶۷)^۲ = ۷/۱۱ \quad , \quad S = \sqrt{S^۲} = \sqrt{۷/۱۱} = ۲/۶۶۷$$

$$z_i = \frac{x_i - \bar{x}}{S} = \frac{x_i - ۶۷}{۲/۶۶۷} \quad , \quad i = ۱, ۲, ..., ۹$$

$Z_۱$	$Z_۲$	$Z_۳$	$Z_۴$	$Z_۵$	$Z_۶$	$Z_۷$	$Z_۸$	$Z_۹$
۰/۷۵	-۰/۳۷۵	۰	۰/۷۵	-۱/۱۲۵	-۱/۵	-۰/۷۵	۰/۳۷۵	۱/۸۷۵

Z_i های به دست آمده فاقد واحد اندازه‌گیری‌اند و متغیرهای استاندارد نامیده می‌شوند.

۱۴ـ نمره‌ی دانش‌آموزان یک کلاس در امتحان ریاضی دارای میانگین ۷۲ و انحراف معیار ۱۵ و در امتحان فیزیک دارای میانگین ۵۰ و انحراف معیار ۲۰ می‌باشد. اگر نمره‌ی دانش‌آموزی در ریاضی ۶۰ و در فیزیک ۳۵ باشد این دانش‌آموز در کدام کلاس مستعدتر است؟

پاسخ: برای این که نمره‌ی دانش‌آموز در درس ریاضی و فیزیک با هم مقایسه کنیم باید نمرات را در دو درس استاندارد کنیم.

اگر میانگین و انحراف معیار را در درس ریاضی با $\bar{x}_R = ۷۲$ و $S_R = ۱۵$ نشان دهیم نمره‌ی ریاضی دانش‌آموز بعد از استاندارد شدن به صورت:

ریاضی $z_R = \frac{۶۰ - ۷۲}{۱۵} = -۰/۸$

همچنین برای درس فیزیک با $\bar{x}_F = ۵۰$ و $S_F = ۲۰$ نشان داده و نمره‌ی فیزیک دانش‌آموز بعد از استاندارد شدن به صورت:

فیزیک $z_F = \frac{۳۵ - ۵۰}{۲۰} = -۰/۷۵$

چون $z_F > z_R \Leftarrow -۰/۸ < -۰/۷۵$ پس این دانش‌آموز در درس فیزیک مستعدتر می‌باشد.

۱۵ـ جدول زیر مقدار تولید شیر خشک غیرچربی در ۱۲ ماه گذشته نشان می‌دهد:

۶۷/۱ ۴۹/۳ ۵۰/۳ ۵۳/۲ ۷۶/۲ ۹۸/۷ ۱۲۶/۵ ۱۳۱/۴ ۱۱۱/۵ ۹۵/۸ ۸۱/۶ ۸۳/۵

دامنه، میانگین، واریانس و ضریب تغییرات مقدار تولیدات را به دست آورد.

پاسخ: دامنه: $R = x_{max} - x_{min}$

$$R = ۱۳۱/۴ - ۴۹/۳ = ۸۲/۱$$

میانگین: $\bar{x} = \frac{۱}{n}\sum_{i=۱}^{n} x_i \to \bar{x} = \frac{۱}{۱۲}\sum_{i=۱}^{۱۲} x_i = ۸۵/۴۲۵$

واریانس: $S^۲ = \frac{۱}{n}\sum_{i=۱}^{n}(x_i - \bar{x})^۲ \to S^۲ = \frac{۱}{۱۲}\sum_{i=۱}^{۱۲}(x_i - ۸۵/۴۲۵)^۲ = ۷۳۱/۷۵۹$

انحراف معیار: $S = \sqrt{S^۲} = \sqrt{۷۳۱/۷۵۹} = ۲۷/۰۵۱$

ضریب تغییرات: $CV = \frac{S}{\bar{x}} \quad \to \quad CV = \frac{۲۷/۰۵۱}{۸۵/۴۲۵} = ۰/۳۱۷$

معمولاً CV را به صورت درصد بیان می‌کنند پس

$\% \; CV = ۳۱/۷$

۱۶ـ نرخ سود سپرده ۱۰۰ حساب بانک ملت دارای میانگین ۳/۵ و انحراف معیار ۳/۴۸ است و در بانک ملی دارای میانگین ۳/۲۵ و انحراف معیار ۳/۵۲ است. ضریب تغییر هر بانک را محاسبه کنید. برای پس‌انداز، کدام بانک را ترجیح می‌دهید.

پاسخ: بانک ملت $\bar{x} = ۳/۵ \quad , \quad S = ۳/۴۸$

بانک ملت $CV = \frac{S}{\bar{x}} = \frac{۳/۴۸}{۳/۵} = ۰/۹۹۴$

بانک ملی $\bar{x} = ۳/۲۵ \quad , \quad S = ۳/۵۲$

بانک ملی $CV = \frac{S}{\bar{x}} = \frac{۳/۵۲}{۳/۲۵} = ۱/۰۸۳$

چون مقدار CV برای بانک ملت کم‌تر است پس پس‌انداز در بانک ملت مطلوب‌تر است.

۱۷ـ اگر x_iها را دو به دو مقایسه کنیم و معدل مجموع $n^۲$ پرانتز $(x_i - x_j)^۲$ را در نظر بگیریم معیاری به صورت $\sigma_*^۲ = \frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{j=۱}^{n}(x_i - x_j)^۲$ برای پراکندگی به دست می‌آید نشان دهید که $\sigma_*^۲ = ۲S^۲$ که $S^۲$ واریانس نمونه است.

پاسخ: می‌خواهیم نشان دهیم $\sigma_*^۲ = ۲S^۲$، از طرف اول رابطه شروع می‌کنیم. رابطه را ساده می‌کنیم تا به طرف دوم برسیم.

$$\sigma_*^۲ = \frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{j=۱}^{n}(x_i - x_j)^۲$$

راهنمایی: داخل پرانتز $\bar{x}$ را اضافه و کم می‌کنیم.

$$=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-x_j+\bar{x}-\bar{x})^2=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(x_i-\bar{x})-(x_j-\bar{x})\right]^2$$

با توجه به اتحاد $(a-b)^2=a^2+b^2-2ab$ اتحاد داخل کروشه را باز می‌کنیم که در آن:

$$b=(x_j-\bar{x}) \quad , \quad a=(x_i-\bar{x})$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(x_i-\bar{x})^2+(x_j-\bar{x})^2-2(x_i-\bar{x})(x_j-\bar{x})\right]$$

با استفاده از خواص سیگما، سیگما را در تک تک جمله‌ها داخل کروشه تأثیر می‌دهیم:

$$=\frac{1}{n^2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-\bar{x})^2+\sum_{i=1}^{n}\sum_{j=1}^{n}(x_j-\bar{x})^2-2\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-\bar{x})(x_j-\bar{x})\right]$$

$$=\frac{1}{n^2}\left[n\sum_{i=1}^{n}(x_i-\bar{x})^2+n\sum_{j=1}^{n}(x_j-\bar{x})^2-2\sum_{i=1}^{n}(x_i-\bar{x})\sum_{j=1}^{n}(x_j-\bar{x})\right]$$

با توجه به خواص میانگین حسابی می‌دانیم که مجموع انحرافات از میانگین داده‌ها برابر صفر است یا به عبارتی:

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

پس مقدار جمله‌ی سوم داخل کروشه برابر صفر است:

$$2\sum_{i=1}^{n}(x_i-\bar{x})\sum_{j=1}^{n}(x_j-\bar{x})=0$$

حال $\frac{1}{n^2}$ پشت کروشه را در دو جمله‌ی اول باقی‌مانده تأثیر می‌دهیم و داریم:

$$=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2+\frac{1}{n}\sum_{j=1}^{n}(x_j-\bar{x})^2=S^2+S^2=2S^2$$

رابطه‌ی بالا ثابت شد.

۱۸ـ داده‌های آماری زیر که عمر ۸ رایانه را نشان می‌دهد:

۱۳   ۸   ۱۰   ۱۵   ۷   ۴   ۲   ۵

مطلوب است:

الف) گشتاورهای مرتبه‌ی اول تا چهارم حول نقطه‌ی صفر

ب) گشتاورهای مرتبه‌ی اول تا چهارم حول میانگین نمونه

ج) ضریب تغییر

پاسخ: الف)

$$m'_r=\frac{1}{n}\sum_{i=1}^{n}(x_i-a)^r \qquad r=1,2,3,4$$

چون مقدار $a=0$ است پس:

$$m'_r=\frac{1}{n}\sum_{i=1}^{n}x_i^r$$

گشتاور غیرمرکزی مرتبه‌ی اول: $m'_1=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{8}\sum_{i=1}^{8}x_i=\frac{1}{8}[5+2+\dots+13]=\frac{64}{8}=8$

که m'_1 حول نقطه‌ی صفر، میانگین نمونه می‌باشد.

$$m'_1=\bar{x}=8$$

گشتاور غیرمرکزی مرتبه‌ی دوم: $m'_2=\frac{1}{8}\sum_{i=1}^{8}x_i^2=\frac{1}{8}[25+4+16+\dots+169]=\frac{652}{8}=81/5$

گشتاور غیرمرکزی مرتبه‌ی سوم: $m'_3=\frac{1}{8}\sum_{i=1}^{8}x_i^3=\frac{1}{8}[125+8+64+\dots+2197]=\frac{7624}{8}=953$

گشتاور غیرمرکزی مرتبه‌ی چهارم: $m'_4=\frac{1}{8}\sum_{i=1}^{8}x_i^4=\frac{1}{8}[625+16+256+\dots+28561]$

$$=\frac{96580}{8}=12072/5$$

ب)

$$m_r=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^r \qquad r=1,2,3,4$$

ابتدا باید مقدار $\bar{x}$ را به دست آورد، چون در قسمت الف مقدار $\bar{x}$ یا همان m'_1 محاسبه شده است پس داریم:

$$\bar{x}=8$$

گشتاور مرکزی مرتبه‌ی اول: $m_1=\frac{1}{8}\sum_{i=1}^{8}(x_i-8)=\frac{1}{8}\left[(5-8)+(2-8)+\dots+(13-8)\right]=0$

نکته: برای $r=1$ ، $m_1=0$

گشتاور مرکزی مرتبه‌ی دوم: $m_2=\frac{1}{8}\sum_{i=1}^{8}(x_i-8)^2=\frac{1}{8}[9+36+16+1+49+4+0+25]$

$$=\frac{140}{8}=17/5\rightarrow S^2=m_2=17/5$$

نکته: برای $r=2$ ، $m_2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=S^2$

گشتاور مرکزی مرتبه‌ی سوم: $m_3=\frac{1}{8}\sum_{i=1}^{8}(x_i-8)^3$

$$=\frac{1}{8}[-27-216-64-1+343+8+0+125]=\frac{168}{8}=21$$

گشتاور مرکزی مرتبه‌ی چهارم: $m_4=\frac{1}{8}\sum_{i=1}^{8}(x_i-8)^4$

$$=\frac{1}{8}[81+1296+256+1+2401+16+0+625]$$

$$=\frac{4676}{8}=584/5$$

ج) ضریب تغییر : $CV=\frac{S}{\bar{x}}$

که مقدار S^2 و $\bar{x}$ در قسمت الف و ب به دست آمد.

$$\bar{x}=8 \quad , \quad S^2=17/5 \quad , \quad S=4/183$$

پس داریم:

$$CV=\frac{4/183}{8}=0/523\rightarrow \%CV=52/3$$

۱۹ـ جدول توزیع فراوانی زیر را در نظر بگیرید.

کلاس	۰-۱	۱-۲	۲-۳	۳-۴	۴-۵	۵-۶	۶-۷	۷-۸	۸-۹	۹-۱۰
فراوانی	۳	۹	۲	۸	۱۱	۱۳	۱۲	۸	۹	۵

الف) کلاس‌ها را به صورت کراندار بنویسید.

ب) حد متوسط یا نماینده‌ی کلاس‌ها را به دست آورید.

ج) طول هر کلاس چند است.

د) درصد فراوانی و درصد تجمعی را به دست آورید.

پاسخ:

شماره کلاس	کلاس	فراوانی	کلاس کراندار	حد متوسط	طول کلاس	درصد فراوانی	فراوانی تجمعی	درصد فراوانی تجمعی
۱	۰-۱	۳	۰-۱	۰/۵	۱	۳/۷۵	۳	۳/۷۵
۲	۱-۲	۹	۱-۲	۱/۵	۱	۱۱/۲۵	۱۲	۱۵
۳	۲-۳	۲	۲-۳	۲/۵	۱	۲/۵	۱۴	۱۷/۵
۴	۳-۴	۸	۳-۴	۳/۵	۱	۱۰	۲۲	۲۷/۵
۵	۴-۵	۱۱	۴-۵	۴/۵	۱	۱۳/۷۵	۳۳	۴۱/۲۵
۶	۵-۶	۱۳	۵-۶	۵/۵	۱	۱۶/۲۵	۴۶	۵۷/۵
۷	۶-۷	۱۲	۶-۷	۶/۵	۱	۱۵	۵۸	۷۲/۵
۸	۷-۸	۸	۷-۸	۷/۵	۱	۱۰	۶۶	۸۲/۵
۹	۸-۹	۹	۸-۹	۸/۵	۱	۱۱/۲۵	۷۵	۹۳/۷۵
۱۰	۹-۱۰	۵	۹-۱۰	۹/۵	۱	۶/۲۵	۸۰	۱۰۰
	—	۸۰	—	—	—	۱۰۰	—	—

الف) کران پایین کلاس کراندار iام را با a_i^* و کران بالای کلاس کراندار iام را با b_i^* نشان می‌دهیم و به صورت زیر به دست می‌آید:

$$a_i^* = \frac{a_i + b_{i-۱}}{۲} \quad , \quad b_i^* = \frac{a_{i+۱} - b_i}{۲}$$

ب) $\frac{a_i + b_i}{۲}$: حد متوسط یا نماینده‌ی کلاس iام $\qquad i = ۱, ۲, \ldots, k$

که در ستون پنجم جدول محاسبه شده است.

ج) طول هر کلاس از تفاضل حد پایین از حد بالا به دست می‌آید و آن را با l نمایش می‌دهند.

$l_i = b_i - a_i \qquad i = ۱ , ۲ , \ldots , k$

در این مثال طول کلاس‌ها برای تمام طبقات برابر می‌باشد و مقدار آن $l=۱$ است.

$\frac{f_i}{n} \times ۱۰۰$: درصد فراوانی کلاس iام $\qquad i = ۱ , ۲ , \ldots , k$

فراوانی کلاس‌های ماقبل + فراونی آن کلاس : فراوانی تجمعی کلاس iام (F_i)

$\frac{F_i}{n} \times ۱۰۰$: درصد فراوانی تجمعی کلاس iام $\qquad i = ۱ , ۲ , \ldots , k$

۲۰ـ برای داده‌های زیر، جدول توزیع فراوانی را کامل کنید و به سؤالات زیر پاسخ دهید:

۴۷/۵	۶۰/۹	۳۱/۳	۳۲/۳	۴۲/۳	۴۲/۲	۴۴/۵	۵۰/۰	۵۳/۹	۴۹/۲
۵۳/۴	۴۵/۷	۵۲/۳	۵۱/۷	۴۵/۳	۴۰/۲	۵۷/۶	۴۱/۱	۳۷/۹	۴۳/۵
۵۰/۳	۴۸/۶	۵۹/۶	۵۵/۱	۶۷/۲	۳۲/۵	۵۰/۰	۴۵/۹	۴۵/۷	۵۱/۰
۳۵/۸	۵۶/۳	۶۲/۹	۶۲/۲	۴۴/۰	۴۱/۵	۳۸/۳	۴۷/۴	۴۶/۸	۴۵/۱
۳۱/۸	۳۴/۳	۵۲/۱	۵۵/۲	۴۱/۳	۴۹/۶	۴۷/۴	۴۸/۵	۳۳/۵	۳۸/۳
۳۱/۶	۴۷/۹	۳۲/۸	۴۹/۲	۴۸/۴	۳۹/۸	۴۱/۲	۴۷/۰	۴۶/۰	۴۹/۳
۳۹/۱	۴۳/۴	۴۱/۵	۴۵/۱	۴۴/۴	۴۶/۲	۴۴/۵	۵۴/۱	۵۴/۵	۳۹/۱
۳۷/۴	۳۳/۸	۲۸/۷	۳۳/۶	۳۷/۲	۴۸/۸	۴۳/۷	۴۳/۱	۴۱/۶	۴۳/۵
۴۸/۸	۴۷/۳	۳۹/۰	۴۸/۵	۵۰/۶	۵۱/۹	۴۵/۱	۵۳/۰	۴۴/۲	۳۸/۲

الف) چند درصد از داده‌ها کم‌تر یا مساوی ۴۸ هستند؟

ب) بزرگ‌ترین و کوچک‌ترین کلاس را مشخص کنید.

ج) چند درصد از داده‌ها بیش‌تر یا مساوی ۶۰ هستند؟

پاسخ: ابتدا باید مقدار k (تعداد طبقات) و l (طول کلاس‌ها) را مشخص کنیم.

$n=۹۰ \qquad k = ۱ + ۳/۳۲ \log n \rightarrow k = ۱ + ۳/۳۲ \log ۹۰ = ۷/۴۸۸ \sim ۷$

تعداد طبقات را $k=۷$ در نظر می‌گیریم:

$$R = x_{max} - x_{min} \rightarrow R = ۶۷/۲ - ۲۸/۷ = ۳۸/۵$$

$$l = \frac{R}{k} \rightarrow l = \frac{۳۸/۵}{۷} = ۵/۵$$

طول کلاس	درصد فراوانی تجمعی	فراوانی تجمعی	درصد فراوانی	حد متوسط	فراوانی	کلاس
۵/۵	۱۱/۱۱	۱۰	۱۱/۱۱	۳۱/۴۵	۱۰	۲۸/۷-۳۴/۲
۵/۵	۲۳/۳۲	۲۱	۱۲/۲۲	۳۶/۹۵	۱۱	۳۴/۲-۳۹/۷
۵/۵	۴۸/۸۹	۴۴	۲۵/۵۶	۴۲/۴۵	۲۳	۳۹/۷-۴۵/۲
۵/۵	۷۸/۸۹	۷۱	۳۰	۴۷/۹۵	۲۷	۴۵/۲-۵۰/۷
۵/۵	۹۲/۲۲	۸۳	۱۳/۳۳	۵۳/۴۵	۱۲	۵۰/۷-۵۶/۲
۵/۵	۹۶/۶۷	۸۷	۴/۴۴	۵۸/۹۵	۴	۵۶/۲-۶۱/۷
۵/۵	۱۰۰	۹۰	۳/۳۳	۶۴/۴۵	۳	۶۱/۷-۶۷/۲
—	—	—	۱۰۰	—	۹۰	

الف) باید فراوانی کلاس‌های اول، دوم، سوم و کلاس چهارم تا عدد ۴۸ را به دست آوریم. در کلاس ۲۸/۷–۳۴/۲، فراوانی ۱۰، در کلاس ۳۴/۲–۳۹/۷ فراوانی ۱۱، در کلاس ۳۹/۷–۴۵/۲ فراوانی ۲۳ و در کلاس ۴۵/۲–۴۸، ۱۳ عدد وجود دارد. بنابراین:

$$۱۰ + ۱۱ + ۲۳ + ۱۳ = ۵۷$$

درصد داده‌های کمتر از ۴۸: $\frac{۵۷}{۹۰} \times ۱۰۰ = ۶۳/۳۳$

ب) بزرگ‌ترین کلاس (بیش‌ترین فراوانی) کلاس ۴۵/۲–۵۰/۷ با ۲۷ فراوانی و کوچک‌ترین کلاس (کم‌ترین فراوانی) کلاس ۶۱/۷–۶۷/۲ با فراوانی ۳ می‌باشد.

ج) در کلاس ۵۶/۲–۶۱/۷، یک عدد بیش‌تر یا مساوی ۶۰ وجود دارد و در کلاس ۶۱/۷–۶۷/۲ فراوانی ۳ است پس:

$$۱ + ۳ = ۴$$

درصد داده‌های بیش‌تر یا مساوی ۶۰ : $\frac{۴}{۹۰} \times ۱۰۰ = ۴/۴۴$

۲۱ـ جدول توزیع فراوانی زیر میزان منواکسیدکربن خارج شده از ۷۹۴ اتومبیل را نشان می‌دهد که بر اساس میزان منواکسیدکربن ایجاد شده توسط هر اتومبیل به گروه‌های به طول ۲۴ گرم در هر مایل گروه‌بندی شده‌اند.

فاصله	۰–۲۴	۲۴–۴۸	۴۸–۷۲	۷۲–۹۶	۹۶–۱۲۰	۱۲۰–۱۴۴	۱۴۴–۱۶۸
حد متوسط	۱۲	۳۶	۶۰	۸۴	۱۰۸	۱۳۲	۱۵۶
فراوانی	۱۳	۹۸	۱۶۱	۱۸۹	۱۴۸	۸۵	۴۵

فاصله	۱۶۸–۱۹۲	۱۹۲–۲۱۶	۲۱۶–۲۴۰	۲۴۰–۲۶۴	۲۶۴–۲۸۸	۲۸۸–۳۱۲	۳۱۲–۳۳۶	۳۳۶–۳۶۰
حدمتوسط	۱۸۰	۲۰۴	۲۲۸	۲۵۲	۲۷۶	۳۰۰	۳۲۴	۳۴۸
فراوانی	۳۰	۱۰	۵	۵	۱	۲	۱	۱

الف) درصد فراوانی، فراوانی تجمعی و درصد فراوانی تجمعی را به دست آورید.

ب) میانگین، واریانس و ضریب تغییر را حساب کنید.

ج) میانه، نما، چارک اول و سوم را حساب کنید.

پاسخ:

شماره کلاس	کلاس	حد متوسط x_i	فراوانی f_i	درصد فراوانی	فراوانی تجمعی F_i	درصد فراوانی تجمعی	$f_i x_i$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
۱	۰–۲۴	۱۲	۱۳	۱/۶۴	۱۳	۱/۶۴	۱۵۶	۶۷۶۵/۰۶	۸۷۹۴۵/۸۱
۲	۲۴–۴۸	۳۶	۹۸	۱۲/۳۴	۱۱۱	۱۳/۹۸	۳۵۲۸	۳۳۹۳/۰۶	۳۳۲۵۲۰/۱۳
۳	۴۸–۷۲	۶۰	۱۶۱	۲۰/۲۸	۲۷۲	۳۴/۲۶	۹۶۶۰	۱۱۷۳/۰۶	۱۸۸۸۶۳/۰۶
۴	۷۲–۹۶	۸۴	۱۸۹	۲۳/۸۰	۴۶۱	۵۸/۰۶	۱۵۸۷۶	۱۰۵/۰۶	۱۹۸۵۶/۸۱
۵	۹۶–۱۲۰	۱۰۸	۱۴۸	۱۸/۶۴	۶۰۹	۷۶/۷۰	۱۵۹۸۴	۱۸۹/۰۶	۲۷۹۸۱/۲۵
۶	۱۲۰–۱۴۴	۱۳۲	۸۵	۱۰/۷۱	۶۹۴	۸۷/۴۱	۱۱۲۲۰	۱۴۲۵/۰۶	۱۲۱۱۳۰/۳۱
۷	۱۴۴–۱۶۸	۱۵۶	۴۵	۵/۶۷	۷۳۹	۹۳/۰۷	۷۰۲۰	۳۸۱۳/۰۶	۱۷۱۵۸۷/۸۱
۸	۱۶۸–۱۹۲	۱۸۰	۳۰	۳/۷۸	۷۶۹	۹۶/۸۵	۵۴۰۰	۷۳۵۳/۰۶	۲۲۰۵۹۱/۸۸
۹	۱۹۲–۲۱۶	۲۰۴	۱۰	۱/۲۶	۷۷۹	۹۸/۱۱	۲۰۴۰	۱۲۰۴۵/۰۶	۱۲۰۴۵/۶۳
۱۰	۲۱۶–۲۴۰	۲۲۸	۵	۰/۶۳	۷۸۴	۹۸/۷۴	۱۱۴۰	۱۷۸۸۹/۰۶	۸۹۴۴۵/۳۱
۱۱	۲۴۰–۲۶۴	۲۵۲	۵	۰/۶۳	۷۸۹	۹۹/۳۷	۱۲۶۰	۲۴۸۸۵/۰۶	۱۲۴۴۲۵/۳۱
۱۲	۲۶۴–۲۸۸	۲۷۶	۱	۰/۱۳	۷۹۰	۹۹/۵	۲۷۶	۳۳۰۳۳/۰۶	۳۳۰۳۳/۰۶
۱۳	۲۸۸–۳۱۲	۳۰۰	۲	۰/۲۵	۷۹۲	۹۹/۷۵	۶۰۰	۴۲۳۳۳/۰۶	۸۴۶۶۶/۱۳
۱۴	۳۱۲–۳۳۶	۳۲۴	۱	۰/۱۳	۷۹۳	۹۹/۸۷	۳۲۴	۵۲۷۸۵/۰۶	۵۲۷۸۵/۰۶
۱۵	۳۳۶–۳۶۰	۳۴۸	۱	۰/۱۳	۷۹۴	۱۰۰	۳۴۸	۶۴۳۸۹/۰۶	۶۴۳۸۹/۰۶
	—	—	۷۹۴	۱۰۰/۰۲	—	—	۷۴۸۳۲	—	۱۷۳۹۶۷۱/۶۲

الف) نکته: مجموع ستون درصد فراوانی باید برابر ۱۰۰ باشد که در این جا تقریباً ۱۰۰/۰۲ به دست آمده که به خاطر گرد کردن اعداد می‌باشد.

ب)

$n = \sum_{i=1}^{۱۵} f_i = ۷۹۴$ میانگین $\bar{x} = \frac{1}{n} \sum_{i=1}^{۱۵} f_i x_i = \frac{۱}{۷۹۴}(۷۴۸۳۲) = ۹۴/۲۵$

واریانس $S^2 = \frac{1}{n} \sum_{i=1}^{۱۵} f_i (x_i - \bar{x})^2 = \frac{۱}{۷۹۴}(۱۷۳۹۶۷۱/۶۲) = ۲۱۹۱/۰۲$

انحراف معیار

ضریب تغییر $CV = \frac{S}{\bar{x}}$, $CV = \frac{۴۶/۸۱}{۹۴/۲۵} = ۰/۵۰$

معمولاً CV را به صورت درصد بیان می‌کنند پس: $\%CV = ۵۰$

ج) میانه $m = a_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \times l$

> یادآوری: ابتدا باید کلاس یا طبقه‌ای که میانه در آن قرار دارد به دست آوریم:
> ۱ـ $\frac{n}{2}$ را محاسبه می‌کنیم.
> ۲ـ در ستون فراوانی تجمعی از بالا به پایین اعداد را با $\frac{n}{2}$ مقایسه می‌کنیم اولین کلاسی که بزرگ‌تر یا مساوی $\frac{n}{2}$ باشد کلاس متناظر، کلاسی است که میانه در آن قرار دارد.
> که در فرمول بالا: a_i حد پایین کلاس میانه‌دار، f_i فراوانی کلاسی که میانه در آن قرار دارد،
> F_{i-1} فراوانی تجمعی کلاس ماقبل از کلاس میانه‌دار و l طول کلاس طبقه‌ی میانه‌دار می‌باشد.

$\frac{n}{2} = \frac{794}{2} = 397 \qquad 397 < 461 \Rightarrow i = 4$

پس کلاس چهارم، کلاسی است که میانه در آن قرار دارد.

$f_i = f_4 = 189 \qquad F_{i-1} = F_3 = 272$

$a_i = a_4 = 72 \qquad l = 24$

$m = 72 + \frac{397 - 272}{189} \times 24 = 72 + 15/87 = 87/87$

نما: $M = a_i + \frac{d_1}{d_1 + d_2} \times l$

> یادآوری: ابتدا باید کلاسی که نما در آن قرار دارد مشخص شود، در یک جدول توزیع فراوانی، نما در کلاسی است که بیش‌ترین فراوانی را دارد.
> در رابطه‌ی بالا: a_i حد پایین کلاس نمادار، l طول کلاس طبقه‌ی نمادار
> $d_1 = f_i - f_{i-1} \quad , \quad d_2 = f_i - f_{i+1}$

چون کلاس چهارم بیش‌ترین فراوانی را دارد پس نما در این کلاس قرار دارد و $i=4$ است.

$a_i = 72 \qquad l = 96 - 72 = 24$

$d_1 = 189 - 161 = 28 \qquad d_2 = 189 - 148 = 41$

$M = 72 + \frac{28}{28 + 41} \times 24 = 72 + 9/74 = 81/74$

چارک زام برای $j = 1, 2, 3$ و

$Q_j = a_i + \frac{\frac{j \times n}{4} - F_{i-1}}{f_i} \times l$

برای یافتن کلاسی که چارک اول، دوم (میانه)، سوم در آن قرار دارد به ترتیب $\frac{n}{4}$، $\frac{n}{2}$، $\frac{3n}{4}$ را با ستون فراوانی تجمعی مقایسه می‌کنیم و کلاس مورد نظر را پیدا می‌کنیم.

چارک اول:

$\frac{n}{4} = \frac{794}{4} = 198/5$

چارک اول در کلاس سوم قرار دارد. $198/5 \le 272 \to i = 3$

$a_i = 48 \ , \ f_i = 161 \ , \ F_{i-1} = 111 \ , \ l = 72 - 48 = 24$

$Q_1 = 48 + \frac{198/5 - 111}{161} \times 24 = 48 + 13/04 = 61/04$

چارک سوم

$\frac{3n}{4} = \frac{2382}{4} = 595/5$

چارک سوم در کلاس پنجم قرار دارد. $595/5 \le 609 \to i = 5$

$a_i = 96 \ , \ f_i = 148 \ , \ F_{i-1} = 461 \ , \ l = 120 - 96 = 24$

$Q_3 = 96 + \frac{595/5 - 461}{148} \times 24 = 96 + 21/81 = 117/81$

نکته: چارک دوم همان میانه است که قبلاً محاسبه شد.

۲۲ـ اندازه‌گیری‌های ۱۵ نقطه‌ی جوش از ترکیبات سیکیان بر حسب درجه‌ی سلسیوس به صورت زیر داده شده است. نمودار نقطه‌ای را رسم کنید.

۱۴۱ ۱۳۶ ۱۵۳ ۱۷۰ ۱۶۲ ۱۵۵ ۱۴۶ ۱۸۳ ۱۵۷ ۱۴۸ ۱۳۲ ۱۶۰
۱۵۷ ۱۵۰ ۱۶۶

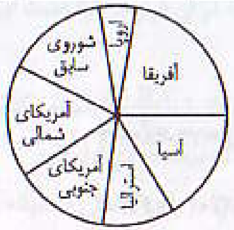
پاسخ:

۲۳ـ جدول زیر مساحت‌های قاره‌های سیاسی جهان را بر حسب میلیون کیلومتر مربع نشان می‌دهد. نمودار دایره‌ای آن را رسم کنید.

قاره	آفریقا	آسیا	آمریکای شمالی	اروپا	آمریکای جنوبی	شوروی سابق	استرالیا	کل
مساحت	۳۰/۳	۲۶/۹	۲۴/۳	۴/۹	۱۷/۹	۲۰/۵	۸/۵	۱۳۳/۳۰

پاسخ:

S	۸۱/۸۳	۷۶/۲۵	۶۵/۶۳	۱۳/۲۳	۴۸/۳۴	۵۵/۳۶	۲۲/۹۶	۳۶۰
درصد فراوانی	۲۲/۷۳	۲۰/۱۸	۱۸/۲۳	۳/۶۸	۱۳/۴۳	۱۵/۳۸	۶/۳۸	۱۰۰



یادآوری: برای به دست آوردن مقدار زاویه‌ی روی دایره برای هر طبقه:

$$S_i = \frac{۳۶۰}{\sum_{i=1}^{k} f_i} \times f_i \qquad \sum_{i=1}^{k} S_i = ۳۶۰$$

اگر در فرمول بالا به جای ۳۶۰، عدد ۱۰۰ قرار دهیم درصد فراوانی به دست می‌آید:

$$\text{درصد فراوانی } S_i = \frac{۱۰۰}{\sum f_i} \times f_i$$

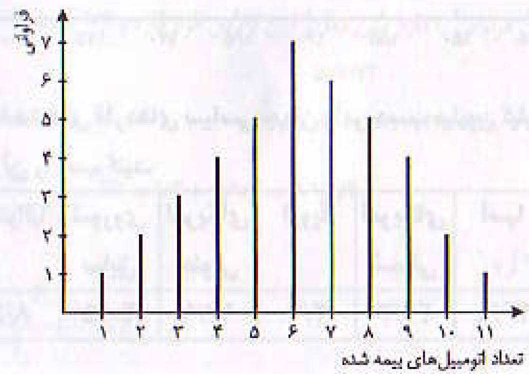
۲۴ـ تعداد اتومبیل‌های بیمه شده توسط یک شرکت بیمه در چهل روز عبارت است از:

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

نمودار میله‌ای را رسم کنید.

پاسخ:

تعداد اتومبیل بیمه‌شده	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	
فراوانی	۱	۲	۳	۴	۵	۷	۶	۵	۴	۲	۱	۴۰



۲۵ـ نمودار مستطیلی و نمودار چندضلعی فراوانی، جدول توزیع فراوانی را رسم کنید.

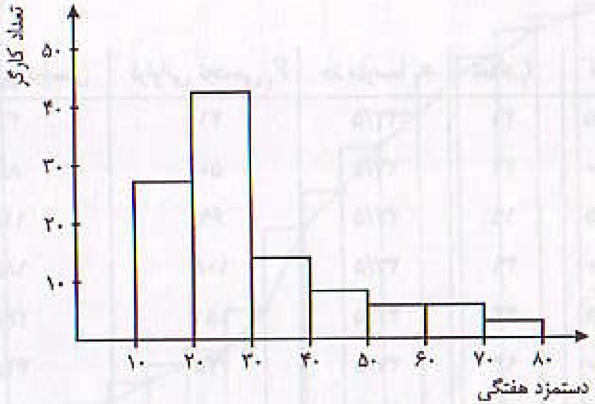
دستمزد هفتگی	۱۰-۱۵	۱۵-۲۰	۲۰-۲۵	۲۵-۳۰	۳۰-۴۰	۴۰-۶۰	۶۰-۸۰
تعداد کارگر	۷	۱۹	۲۷	۱۵	۱۲	۱۲	۸

راهنمایی: **فاصله‌ی کلاس مساوی نیست لذا قبل از رسم نمودار باید فاصله‌ها را یکسان کرد مثلاً کلاس ۴۰-۳۰ با فراوانی ۱۲ تبدیل می‌شود به ۳۵-۳۰ با فراوانی ۷ و ۴۰-۳۵ با فراوانی ۵.**

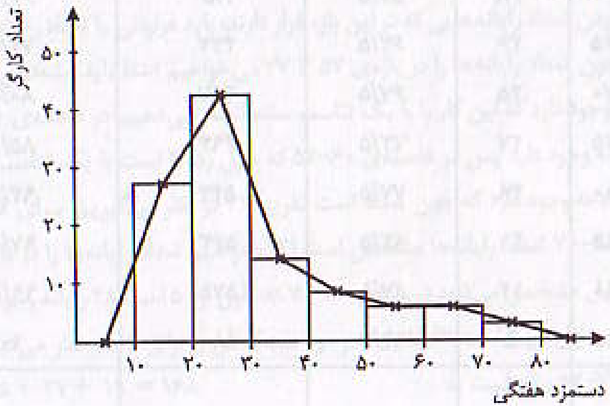
پاسخ: فاصله‌ی کلاس‌ها مساوی نیست لذا قبل از رسم نمودار باید فاصله یکسان شوند برای این که فراوانی طبقات خیلی کوچک نشود طول کلاس‌ها را ۱۰ در نظر می‌گیریم پس جدول به صورت زیر به دست می‌آید:

دستمزد هفتگی	۱۰-۲۰	۲۰-۳۰	۳۰-۴۰	۴۰-۵۰	۵۰-۶۰	۶۰-۷۰	۷۰-۸۰
تعداد کارگر	۲۶	۴۲	۱۲	۷	۵	۵	۳

نمودار مستطیلی



نمودار چندضلعی فراوانی



۲۶ـ جدول زیر قیمت ۵۸۰ رایانه‌ی فروخته شده در یک فروشگاه را نشان می‌دهد.

قیمت	۲۰-۲۵	۲۵-۳۰	۳۰-۳۵	۳۵-۴۰	۴۰-۴۵	۴۵-۵۰	۵۰-۵۵
تعداد	۲۱	۲۹	۱۹	۳۹	۴۳	۹۴	۷۳

قیمت	۵۵-۶۰	۶۰-۶۵	۶۵-۷۰	۷۰-۷۵	۷۵-۸۰	۸۰-۸۵	۸۵-۹۰	۹۰-۹۵
تعداد	۶۸	۳۶	۴۵	۲۷	۴۸	۲۱	۱۲	۵

الف) فراوانی تجمعی و درصد فراوانی تجمعی را به دست آورید.

ب) نمودار چندضلعی فراوانی تجمعی را رسم کنید.

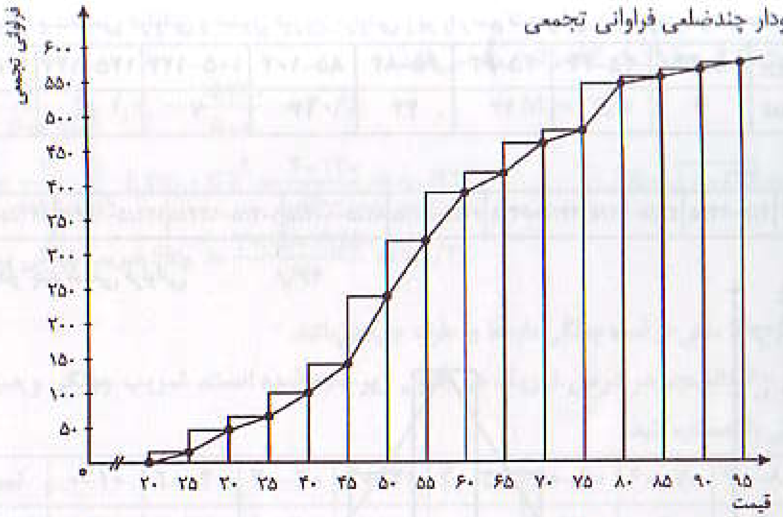
ج) نمودار چندضلعی درصد فراوانی تجمعی (اجایو) را رسم کنید.

د) تعداد رایانه‌هایی را که قیمت آن‌ها بین ۵۷ تا ۷۷ هستند به دست آورید.

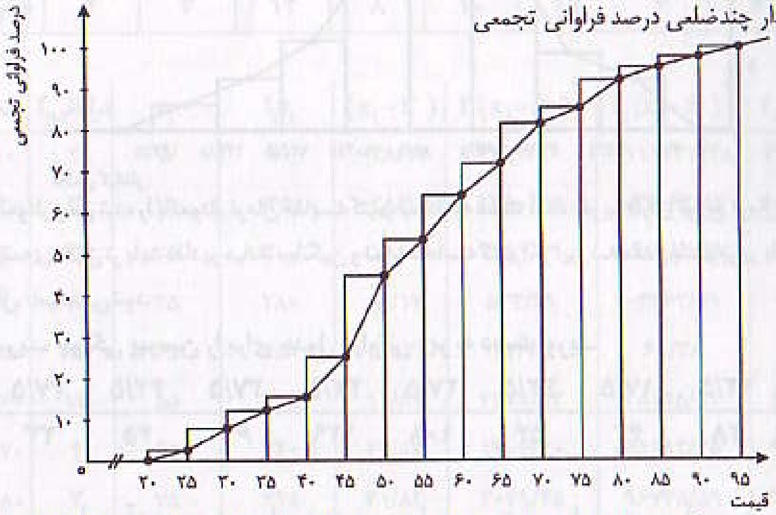
پاسخ: الف)

شماره کلاس	قیمت	تعداد f_i	حد متوسط x_i	فراوانی تجمعی F_i	درصد فراوانی تجمعی
۱	۲۰-۲۵	۲۱	۲۲/۵	۲۱	۳/۶۲
۲	۲۵-۳۰	۲۹	۲۷/۵	۵۰	۸/۶۲
۳	۳۰-۳۵	۱۹	۳۲/۵	۶۹	۱۱/۹۰
۴	۳۵-۴۰	۳۹	۳۷/۵	۱۰۸	۱۸/۶۲
۵	۴۰-۴۵	۴۳	۴۲/۵	۱۵۱	۲۶/۰۳
۶	۴۵-۵۰	۹۴	۴۷/۵	۲۴۵	۴۲/۲۴
۷	۵۰-۵۵	۷۳	۵۲/۵	۳۱۸	۵۴/۸۳
۸	۵۵-۶۰	۶۸	۵۷/۵	۳۸۶	۶۶/۵۵
۹	۶۰-۶۵	۳۶	۶۲/۵	۴۲۲	۷۲/۷۶
۱۰	۶۵-۷۰	۴۵	۶۷/۵	۴۶۷	۸۰/۵۲
۱۱	۷۰-۷۵	۲۷	۷۲/۵	۴۹۴	۸۵/۱۷
۱۲	۷۵-۸۰	۴۸	۷۷/۵	۵۴۲	۹۳/۴۵
۱۳	۸۰-۸۵	۲۱	۸۲/۵	۵۶۳	۹۷/۰۷
۱۴	۸۵-۹۰	۱۲	۸۷/۵	۵۷۵	۹۹/۱۴
۱۵	۹۰-۹۵	۵	۹۲/۵	۵۸۰	۱۰۰
		۵۸۰			

ب) نمودار چندضلعی فراوانی تجمعی



ج) نمودار چندضلعی درصد فراوانی تجمعی



د) برای به دست آوردن تعداد رایانه‌هایی که در این بازه قرار دارند، باید فراوانی یا تعداد رایانه‌ها را در این بازه با هم جمع کنیم. چون تعداد رایانه‌ها را در بازه‌ی ۵۷ تا ۷۷ می‌خواهیم ابتدا باید مشخص کنیم در بازه‌ی ۶۰-۵۷ چند رایانه وجود دارد که این کار را با یک تناسب ساده انجام می‌دهیم، در فاصله‌ی ۶۰-۵۵، که طول رده ۵ است ۶۸ رایانه وجود دارد پس در فاصله‌ی ۶۰-۵۷ که طول رده ۳ است با یک تناسب ساده مشخص می‌شود که ۴۰/۸ رایانه وجود دارد که چون تعداد است تقریباً ۴۱ در نظر می‌گیریم. برای فاصله‌های بعد ۶۵-۶۰ ۷۰-۶۵ ۷۵-۷۰ تعداد رایانه‌ها مشخص است اگر بخواهیم تعداد رایانه‌ها را در بازه‌ی ۷۷-۷۵ به دست آوریم مانند قبل مشخص می‌شود در بازه‌ی ۸۰-۷۵ که طول آن ۵ است ۴۸ رایانه وجود دارد حال برای فاصله‌ی ۷۷-۷۵ که طول آن ۲ است ۱۹/۲ رایانه وجود دارد که آن را برابر ۱۹ در نظر می‌گیریم. پس تعداد رایانه‌ها در بازه‌ی ۷۷-۵۷ برابر است با:

$$41 + 36 + 45 + 27 + 19 = 168$$

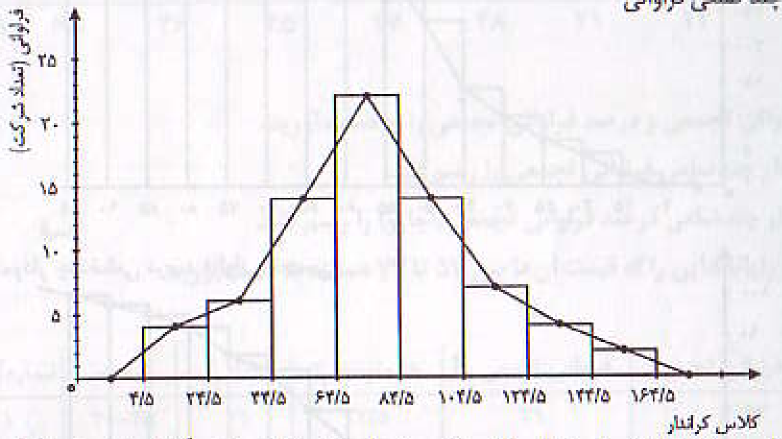
۲۷ـ نمودار چندضلعی فراوانی و جدول توزیع فراوانی زیر را رسم کنید و آن را با نمودار نرمال مقایسه کنید.

برق مصرفی	۵-۲۴	۲۵-۴۴	۴۵-۶۴	۶۵-۸۴	۸۵-۱۰۴	۱۰۵-۱۲۴	۱۲۵-۱۴۴	۱۴۵-۱۶۴
تعدادشرکت	۴	۶	۱۴	۲۲	۱۴	۷	۵	۳

پاسخ:

کلاس کرانداز	۴/۵-۲۴/۵	۲۴/۵-۴۴/۵	۴۴/۵-۶۴/۵	۶۴/۵-۸۴/۵	۸۴/۵-۱۰۴/۵	۱۰۴/۵-۱۲۴/۵	۱۲۴/۵-۱۴۴/۵	۱۴۴/۵-۱۶۴/۵

نمودار چند ضلعی فراوانی



اگر نمودار رسم شده را با نمودار نرمال مقایسه کنیم، نمودار به دست آمده تقریباً دارای توزیع نرمال است. برای مقایسه‌ی دقیق تر باید مقادیر میانه، میانگین و نما را محاسبه کنیم اگر این سه مقدار با هم برابر باشند منحنی نرمال نامیده می‌شود.

۲۸ـ ضریب چولگی پیرسن را برای جدول فراوانی زیر به دست آورید.

x	۱۲/۵	۱۷/۵	۲۲/۵	۲۷/۵	۳۲/۵	۳۷/۵	۴۲/۵	۴۷/۵
f	۲۸	۴۲	۵۴	۱۰۸	۱۲۹	۶۱	۴۵	۳۳

پاسخ:

i	x_i	f_i	$f_i x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
۱	۱۲/۵	۲۸	۳۵۰	-۱۸	۳۲۴	۹۰۷۲
۲	۱۷/۵	۴۲	۷۳۵	-۱۳	۱۶۹	۷۰۹۸
۳	۲۲/۵	۵۴	۱۲۱۵	-۸	۶۴	۳۴۵۶
۴	۲۷/۵	۱۰۸	۲۹۷۰	-۳	۹	۹۷۲
۵	۳۲/۵	۱۲۹	۴۱۹۲/۵	۲	۴	۵۱۶
۶	۳۷/۵	۶۱	۲۲۸۷/۵	۷	۴۹	۲۹۸۹
۷	۴۲/۵	۴۵	۱۹۱۲/۵	۱۲	۱۴۴	۶۴۸۰
۸	۴۷/۵	۳۳	۱۵۶۷/۵	۱۷	۲۸۹	۹۵۳۷
		۵۰۰	۱۵۲۳۰			۴۰۱۲۰

ضریب چولگی پیرسن $Sk_p = \dfrac{\bar{x} - M}{S}$

که در آن $\bar{x}$ میانگین، M نما و S انحراف معیار می‌باشد.

$$\bar{x} = \frac{۱}{۵۰۰} \sum_{i=۱}^{۸} f_i x_i = \frac{۱۵۲۳۰}{۵۰۰} = ۳۰/۵ \quad , \quad M = ۳۲/۵$$

$$S^۲ = \frac{۱}{۵۰۰} \sum_{i=۱}^{۸} f_i (x_i - \bar{x})^۲ = \frac{۴۰۱۲۰}{۵۰۰} = ۸۰/۲۴ \quad , \quad S = \sqrt{۸۰/۲۴} = ۸/۹۶$$

ضریب چولگی پیرسن $Sk_p = \dfrac{۳۰/۵ - ۳۲/۵}{۸/۹۶} = -۰/۲۲$

چون مقدار Sk_p منفی در آمده چولگی داده‌ها به طرف چپ می‌باشد.

۲۹ـ نمره‌ی ۶۰ دانشجو در درس فیزیک در جدول زیر داده شده است. ضریب چولگی و میزان برجستگی را حساب کنید.

نمره	۰-۱۰	۱۰-۲۰	۲۰-۳۰	۳۰-۴۰	۴۰-۵۰	۵۰-۶۰	۶۰-۷۰	۷۰-۸۰	۸۰-۹۰
فراوانی	۲	۳	۱۲	۸	۱۰	۱۷	۴	۳	۱

پاسخ:

نمره	فراوانی f_i	x_i	$f_i x_i$	$(x_i - \bar{x})$	$f_i (x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^3$	$f_i (x_i - \bar{x})^4$
۰-۱۰	۲	۵	۱۰	-۳۸/۱۷	۲۹۱۳/۹۰	-۱۱۱۲۳۳/۴۸	۴۲۴۵۴۰۰/۱۹
۱۰-۲۰	۳	۱۵	۴۵	-۲۸/۱۷	۲۳۸۰/۶۵	-۶۷۰۶۲/۸۲	۱۸۸۹۱۵۹/۵۷
۲۰-۳۰	۱۲	۲۵	۳۰۰	-۱۸/۱۷	۳۹۶۱/۷۹	-۷۱۹۸۵/۶۷	۱۳۰۷۹۷۹/۵۵
۳۰-۴۰	۸	۳۵	۲۸۰	-۸/۱۷	۵۳۳/۹۹	-۴۳۶۲/۷۱	۳۵۶۴۳/۳۳
۴۰-۵۰	۱۰	۴۵	۴۵۰	۱/۸۳	۳۳/۴۹	۶۱/۲۸	۱۱۲/۱۵
۵۰-۶۰	۱۷	۵۵	۹۳۵	۱۱/۸۳	۲۳۷۹/۱۳	۲۸۱۴۵/۱۲	۳۳۲۹۵۶/۸۱
۶۰-۷۰	۴	۶۵	۲۶۰	۲۱/۸۳	۱۹۰۶/۲۰	۴۱۶۱۲/۲۵	۹۰۸۳۹۵/۴۲
۷۰-۸۰	۳	۷۵	۲۲۵	۳۱/۸۳	۳۰۳۹/۴۵	۹۶۷۴۵/۵۹	۳۰۷۹۴۱۲/۰۸
۸۰-۹۰	۱	۸۵	۸۵	۴۱/۸۳	۱۷۴۹/۷۵	۷۳۱۹۲/۰۰	۳۰۶۱۶۲۱/۲۱
	۶۰		۲۵۹۰		۱۸۸۹۸/۳۵	-۱۴۸۷۸/۴۴	۱۴۸۶۰۶۸۰/۳۱

$$\bar{x} = \frac{۱}{۶۰} \sum_{i=۱}^{۹} f_i x_i = \frac{۲۵۹۰}{۶۰} = ۴۳/۱۷$$

$$S^۲ = \frac{۱}{۶۰} \sum_{i=۱}^{۹} f_i (x_i - \bar{x})^۲ = \frac{۱۸۸۹۸/۳۵}{۶۰} = ۳۱۴/۹۷ \quad , \quad S = ۱۷/۷۵$$

$$m_۳ = \frac{۱}{۶۰} \sum_{i=۱}^{۹} f_i (x_i - \bar{x})^۳ = \frac{-۱۴۸۷۸/۴۴}{۶۰} = -۲۴۷/۹۷$$

b: ضریب چولگی بر اساس گشتاور مرکزی مرتبه‌ی سوم $b=\frac{m_3}{S^3}$

$$b=\frac{-247/97}{5592/36}=-0/04$$

چون b (ضریب چولگی) مقداری منفی در آمده است پس چولگی داده‌ها به طرف چپ است.

برجستگی $k=\frac{m_4}{S^4}-3$

$$m_4=\frac{1}{60}\sum_{i=1}^{9} f_i(x_i-\bar{x})^4=\frac{14860680/31}{60}=247678/01 \quad , \quad S^4=99264/38$$

$$k=\frac{247678/01}{99264/38}-3=2/5-3=-0/5$$

چون مقدار $k=-0/5<0$ است منحنی چند ضلعی فراوانی در مقایسه با منحنی نرمال دارای پخی است.

۳۰ـ در یک سری مشاهدات هر گاه میزان چولگی خفیف باشد بین میانگین، میانه، نما رابطه‌ی تقریبی زیر برقرار است.

$$\bar{x}-M\approx 3(\bar{x}-m)$$

برای یک سری مشاهدات با $\bar{x}=17/2$، $m=15$ و $S=5$ ضریب چولگی پیرسن را به دست آورید.

پاسخ: ضریب چولگی پیرسن $Sk_p=\frac{\bar{x}-M}{S}$

$$\bar{x}-M\approx 3(\bar{x}-m)$$

$$17/2-M\approx 3(17/2-15)\rightarrow 17/2-M\approx 6/6\rightarrow M=10/60$$

مقدار نما را از رابطه‌ی بالا به دست آورده و در فرمول ضریب چولگی پیرسن قرار می‌دهیم:

$$Sk_p=\frac{17/2-10/60}{5}=\frac{6/60}{5}=1/32$$

چون مقدار Sk_p مثبت است چولگی داده‌ها به طرف راست است.

۳۱ـ هزینه‌ی زندگی ده خانوار ۶ نفره در جدول زیر آمده است:

خانوار	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
هزینه	۱۰۹۰	۱۲۷۰	۱۲۶۰	۱۲۰۰	۱۱۷۰	۱۰۸۰	۱۰۰۰	۱۳۱۰	۱۲۱۰	۱۱۳۰

میانگین، واریانس و ضریب تغییر را با استفاده از کدگذاری به دست آورید.

پاسخ: تذکر: چون داده دارای فراوانی نمی‌باشد a را دسته‌ی وسط در نظر می‌گیریم $a=1170$ و چون داده‌ها در جدول پیوسته نیست $b=1$ در نظر می‌گیریم.

$$u_i=\frac{x_i-1170}{1}\rightarrow u_i=x_i-1170$$

خانوار	هزینه x_i	$u_i=x_i-1170$	$(u_i-\bar{u})$	$(u_i-\bar{u})^2$
۱	۱۰۹۰	-۸۰	-۸۲	۶۷۲۴
۲	۱۲۷۰	۱۰۰	۹۸	۹۶۰۴
۳	۱۲۶۰	۹۰	۸۸	۷۷۴۴
۴	۱۲۰۰	۳۰	۲۸	۷۸۴
۵	۱۱۷۰	۰	-۲	۴
۶	۱۰۸۰	-۹۰	-۹۲	۸۴۶۴
۷	۱۰۰۰	-۱۷۰	-۱۷۲	۲۹۵۸۴
۸	۱۳۱۰	۱۴۰	۱۳۸	۱۹۰۴۴
۹	۱۲۱۰	۴۰	۳۸	۱۴۴۴
۱۰	۱۱۳۰	-۴۰	-۴۲	۱۷۶۴
		۲۰		۸۵۱۶۰

$$\bar{u}=\frac{1}{10}\sum_{i=1}^{10}u_i=\frac{20}{10}=2$$

$$S_u^2=\frac{1}{10}\sum_{i=1}^{10}(u_i-\bar{u})^2=\frac{85160}{10}=8516 \quad , \quad S_u=92/28$$

با توجه به این که $a=1170$ و $b=1$ میانگین و واریانس داده‌های قدیم (x_iها) به صورت زیر است:

$$\bar{u}=\frac{1}{b}(\bar{x}-a)\rightarrow \bar{x}=b\bar{u}+a=2+1170=1172$$

$$S_u^2=\frac{1}{b^2}S_x^2\rightarrow S_x^2=b^2S_u^2=8516 \quad , \quad S_x=92/28$$

و ضریب تغییر برابر است با:

$$CV_x=\frac{S}{\bar{x}}=\frac{92/28}{1172}=0/05 \qquad \%CV=5$$

۳۲ـ جدول توزیع فراوانی دستمزد کارمندان یک کارخانه را نشان می‌دهد. میانگین، واریانس و ضریب تغییر را با استفاده از کدگذاری به دست آورید.

دستمزد	۳۰۰-۴۰۰	۴۰۰-۵۰۰	۵۰۰-۶۰۰	۶۰۰-۷۰۰	۷۰۰-۸۰۰	۸۰۰-۹۰۰	۹۰۰-۱۰۰۰	۱۰۰۰-۱۱۰۰
تعدادکارمندان	۱۵	۲۲	۱۸	۱۴	۹	۷	۵	۴

پاسخ: ابتدا باید داده‌های قدیم را با استفاده از رابطه‌ی زیر به داده‌های جدید u_i تبدیل کنیم.

$$u_i = \frac{x_i - a}{b} \quad , \quad i = ۱, ۲, \dots , n$$

a مقدار ثابتی است که معمولاً بهتر است آن را برابر مرکز دسته‌ای از جدول با فراوانی ماکزیمم اختیار کنیم یا a را برابر مرکز دسته‌ای که در وسط جدول قرار دارد انتخاب کنیم و b را طول دسته یا کلاس در نظر می‌گیریم. در این جا a=۶۵۰ یعنی حد متوسط طبقه‌ی چهارم، b=۱۰۰ طول کلاس طبقه‌ی چهارم در نظر می‌گیریم.

دستمزد	تعدادکارمندان f_i	x_i	$u_i = \frac{x_i - ۶۵۰}{۱۰۰}$	$f_i u_i$	$(u_i - \bar{u})^۲$	$f_i (u_i - \bar{u})^۲$
۳۰۰–۴۰۰	۱۵	۳۵۰	-۳	-۴۵	۵/۹۵	۸۹/۳۰
۴۰۰–۵۰۰	۲۲	۴۵۰	-۲	-۴۴	۲/۰۷	۴۵/۶۲
۵۰۰–۶۰۰	۱۸	۵۵۰	-۱	-۱۸	۰/۱۹	۳/۴۸
۶۰۰–۷۰۰	۱۴	۶۵۰	۰	۰	۰/۳۱	۴/۳۹
۷۰۰–۸۰۰	۹	۷۵۰	۱	۹	۲/۴۳	۲۱/۹۰
۸۰۰–۹۰۰	۷	۸۵۰	۲	۱۴	۶/۵۵	۴۵/۸۸
۹۰۰–۱۰۰۰	۵	۹۵۰	۳	۱۵	۱۲/۶۷	۶۳/۳۷
۱۰۰۰–۱۱۰۰	۴	۱۰۵۰	۴	۱۶	۲۰/۷۹	۸۳/۱۷
	۹۴			-۵۳		۳۵۷/۱۱

$$\bar{u} = \frac{۱}{۹۴} \sum_{i=۱}^{۸} f_i u_i = \frac{-۵۳}{۹۴} = -۰/۵۶$$

$$S_u^۲ = \frac{۱}{۹۴} \sum_{i=۱}^{۸} f_i (u_i - \bar{u})^۲ = \frac{۳۵۷/۱۱}{۹۴} = ۳/۸۰ \rightarrow S_u = ۱/۹۵$$

با توجه به این که a=۶۵۰ و b=۱۰۰ است میانگین و واریانس داده‌های قدیم (x_i ها) را به دست می‌آوریم:

$$\bar{u} = \frac{۱}{b}(\bar{x} - a) \rightarrow \bar{x} = b\bar{u} + a = ۱۰۰ \times (-۰/۵۶) + ۶۵۰ = -۵۶ + ۶۵۰ = ۵۹۴$$

$$S_u^۲ = \frac{۱}{b^۲} S_x^۲ \rightarrow S_x^۲ = b^۲ S_u^۲ = (۱۰۰)^۲ \times ۳/۸۰ = ۳۸۰۰۰$$

$$S_x = \sqrt{۳۸۰۰۰} = ۱۹۴/۹۴$$

با به دست آمدن میانگین و انحراف معیار داده‌های قدیم از روش کدگذاری مقدار CV_x را به دست می‌آوریم:

$$CV_x = \frac{S_x}{\bar{x}} = \frac{۱۹۴/۹۴}{۵۹۴} = ۰/۳۳$$

که معمولاً آن را به صورت درصد بیان می‌کنند:

$$\%CV_x = ۳۳$$

۳۳ـ مقدار کود مصرفی (x) بر حسب تن در هکتار و محصول به دست آمده از هر هکتار (y) در جدول زیر آمده است. جدول دو متغیره‌ی x و y را تشکیل دهید.

x	۱	۲	۵	۳	۲	۳	۳	۵	۴
y	۱۴	۱۴	۱۸	۱۵	۱۵	۱۷	۱۶	۱۸	۱۷

پاسخ:

X \ Y	۱۴	۱۵	۱۶	۱۷	۱۸	$f_{i.}$
۱	۱	۰	۰	۰	۰	۱
۲	۱	۱	۰	۰	۰	۲
۳	۰	۱	۱	۱	۰	۳
۴	۰	۰	۰	۱	۰	۱
۵	۰	۰	۰	۰	۲	۲
$f_{.j}$	۲	۲	۱	۲	۲	$f_{..} = ۹$

احتمال شرطی

اگر پیشامدی همانند A به پیشامد دیگری همانند B مربوط باشد و بدانیم پیشامد B به‌وقوع پیوسته است، در این صورت احتمال وقوع A، به احتمال وقوع A به شرط B تغییر می‌یابد که آن را «احتمال شرطی» می‌گوییم و به صورت $P(A|B)$ نشان داده می‌شود.

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \quad , \quad P(B)>۰$$

نکته: اگر $P(A)\neq ۰$ از $P(A|B)=\frac{P(A\cap B)}{P(B)}$ نتیجه می‌شود که:

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

نکته: اگر $P(B)=۰$ باشد نتیجه می‌شود که $P(A\cap B)=۰$ در این صورت $P(A|B)$ تعریف نشده است.

دو پیشامد مستقل

دو پیشامد را «مستقل» گوییم، در صورتی که وقوع یا عدم‌وقوع یکی در وقوع و یا عدم‌وقوع دیگری هیچ تأثیری نداشته باشد؛ در قانون ضرب احتمالات داریم:

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

اگر A و B دو پیشامد مستقل باشند، وقوع پیشامد B در احتمال وقوع پیشامد A هیچ تأثیری ندارد، یعنی $P(A|B)=P(A)$. قانون ضرب احتمالات برای دو پیشامد مستقل A و B به این صورت ساده می‌شود:

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

بنابراین می‌توان وجود این رابطه را «شرط استقلال» دو پیشامد نیز دانست.

قضیه: اگر دو پیشامد A و B مستقل باشند آن گاه A^C و B، A و B^C، A^C و B^C نیز مستقل‌اند.

قضیه: اگر احتمال وقوع پیشامد $A_۱$ برابر $P_۱$ و احتمال وقوع پیشامد $A_۲$ برابر $P_۲$ و دو پیشامد $A_۱$ و $A_۲$ مستقل باشند آن گاه احتمال این که فقط یکی از آن‌ها اتفاق بیفتد برابر است با:

$$P_۱(۱-P_۲)+(۱-P_۱)P_۲$$

قضیه: *(قانون جمع احتمالات)*: فرض کنید پیشامدهای $A_۱$، $A_۲$، ...، A_k پیشامدهای دو به دو مجزا از هم و اجتماع آن‌ها S باشد و A یک پیشامد دلخواه از S باشد آن گاه:

$$P(A)=\sum_{i=۱}^{k} P(A|A_i)P(A_i)$$

فرمول بیز

اگر پیشامدهای $A_۱$، $A_۲$، ...، A_k دو به دو مجزا و $\bigcup_{i=۱}^{k} A_i=S$ باشد احتمال شرطی هر یک از A_i به شرط اتفاق پیشامد A از S برابر با:

$$P[A_i|A]=\frac{P(A_i\cap A)}{P(A)} \quad , \quad P(A)=\sum_{i=۱}^{k} P(A|A_i)P(A_i)$$

بنابراین:

$$P[A_i|A]=\frac{P(A|A_i)P(A_i)}{\sum_{i=۱}^{k} P(A|A_i)P(A_i)}$$

خودآزمایی فصل دوم

۱ ـ اگر $A=\{x \mid ۱<x<۹\}$ و $B=\{y \mid y<۵\}$ باشد. مقادیر A^C، $A\cup B$، $A\cap B$، $(A\cup B)^C$، $A-B$ را پیدا کنید.

پاسخ:

$$A^C=\{x \mid x\leq ۱ , x\geq ۹\}$$

$$A\cup B=\{x \mid x<۹\}$$

$$A\cap B=\{x \mid ۱<x<۵\}$$

$$(A\cup B)^C=\{x \mid x\geq ۹\}$$

$$A-B=\{x \mid ۵\leq x<۹\}$$

۲ ـ یک سکه سه بار پرتاب می‌شود:

الف) فضای نمونه را بنویسید.

ب) پیشامد این که در سه پرتاب دو شیر ظاهر شود چیست؟

ج) پیشامد این که حداقل یک خط ظاهر شود چیست؟

پاسخ: الف) اگر علامت H برای شیر بودن و T برای خط بودن سکه استفاده شود فضای نمونه‌ی سکه در سه بار پرتاب به صورت است:

S={(HHH) , (HHT) , (HTH) , (THH) , (TTH) , (THT) , (HTT) , (TTT)}

ب) پیشامد دو شیر در سه پرتاب را با A نشان می‌دهیم، داریم:

A={(HHT) , (HTH) , (THH)}

ج) پیشامد حداقل یک خط در سه پرتاب معادل این است که یک خط ظاهر شود یا دو خط یا سه خط و با B نشان می‌دهیم:

B={(HHT) , (HTH) , (THH) , (THT) , (TTH) , (HHT) , (TTT)}

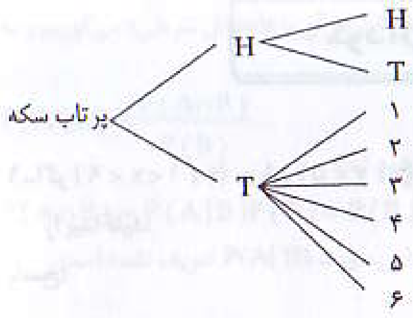
۳ ـ یک آزمایش شامل انداختن یک سکه و در صورت مشاهده‌ی شیر پرتاب مجدد آن می‌باشد و اگر در پرتاب اول خط ملاحظه شود، این بار یک تاس انداخته می‌شود.

الف) فضای نمونه را با رسم نمودار درختی بنویسید.

ب) فضای نمونه‌ی S مربوط به اتفاق A که در وجه فوقانی تاس عددی کم‌تر از ۴ ظاهر شده باشد را معین کنید.

پاسخ: الف)

$S=\{(H,H),(H,T),(T,۱),(T,۲),(T,۳),(T,۴),(T,۵),(T,۶)\}$



ب) $A=\{(T,۱),(T,۲),(T,۳)\}$

$$P(A)=\frac{n(A)}{n(S)}=\frac{۳}{۸}$$

۴ـ الف) به چند طریق ۵ نفر می‌توانند صف جهت سوار شدن به اتوبوس تشکیل دهند؟

ب) اگر ۲ نفر از آن‌ها قبول نکنند که پشت سر هم قرار گیرند چند طریق وجود دارد؟

پاسخ: الف) جایگشت ۵ شیء متمایز برابر است با:

$$۵!=۵\times۴\times۳\times۲\times۱=۱۲۰$$

ب) ابتدا باید تعداد حالاتی را که این دو نفر خاص پشت سر هم قرار می‌گیرند را به دست آورده و از تعداد حالات کلی که این ۵ نفر می‌توانند در صف قرار گیرند کم کنیم.

نکته: برای این که تعداد حالاتی را که این دو نفر خاص پشت سر هم قرار می‌گیرند محاسبه کنیم ابتدا این دو نفر را با هم و به عنوان یک عضو در نظر می‌گیریم، حال ۵ نفر تبدیل به ۴ نفر می‌شود که با !۴ می‌توانند با هم جابجا شوند همچنین دو نفری را که به عنوان یک عضو در نظر گرفتیم خود می‌توانند با !۲ با هم جابجا شوند.

تعداد حالاتی که دو نفر خاص پشت سر هم باشند $۴!\times۲!=۴۸$

تعداد حالاتی که دو نفر خاص پشت سر هم نباشند $۵!-(۴!\times۲!)=۱۲۰-۴۸=۷۲$

۵ـ ۶ رایانه را به چند طریق می‌توان دور یک میز دایره شکل چید؟

پاسخ: نکته: جایگشت n شیء متمایز روی محیط دایره برابر با $(n-1)!$ است.

$$(۶-۱)!=۵!=۵\times۴\times۳\times۲\times۱=۱۲۰$$

۶ـ از یک گروه ۵ نفر مهندس نرم‌افزار و ۳ نفر مهندس سخت‌افزار چند کمیته‌ی ۳ تایی می‌توان تشکیل داد اگر:

الف) محدودیتی نباشد.

ب) با دو مهندس نرم‌افزار و یک مهندس سخت‌افزار

ج) با یک مهندس نرم‌افزار و دو مهندس سخت‌افزار در صورتی که یک مهندس سخت‌افزار مشخصی باید در کمیته شرکت کند.

پاسخ: الف) چون محدودیت ندارد پس از بین مجموع مهندسین نرم‌افزار و سخت‌افزار (۸ نفر) می‌خواهیم ۳ نفر انتخاب کنیم:

$$۸C۳=\frac{۸!}{۳!\times۵!}=۵۶$$

ب)

تعداد ترکیبات ۲ از ۵ نفر مهندسین نرم‌افزار $۵C۲=\frac{۵!}{۲!\times۳!}=۱۰$

تعداد ترکیبات ۱ از ۳ نفر مهندسین سخت‌افزار $۳C۱=\frac{۳!}{۱!\times۲!}=۳$

با استفاده از اصل دوم شمارش تعداد کل کمیته‌ها برابر با: $۱۰\times۳=۳۰$

ج)

تعداد ترکیبات ۱ از ۵ نفر مهندسین نرم‌افزار $۵C۱=\frac{۵!}{۱!\times۴!}=۵$

انتخاب مهندس سخت‌افزار مشخص به یک طریق انجام می‌شود یا می‌توان نوشت: $۱C۱=\frac{۱!}{۱!\times۰!}=۱$

چون یک مهندس سخت‌افزار از قبل انتخاب شده تعداد ترکیبات یکی کم می‌شود پس:

تعداد ترکیبات ۱ از ۲ نفر مهندس سخت‌افزار باقی‌مانده $۲C۱=\frac{۲!}{۱!\times۱!}=۲$

با استفاده از اصل دوم شمارش تعداد کل کمیته‌ها برابر است با: $۵\times۱\times۲=۱۰$

۷ـ بسته‌ای شامل ۹ دستگاه رایانه است که ۳ تا از آن‌ها معیوب هستند. به چند طریق می‌توان ۴ دستگاه رایانه از این بسته انتخاب کرد که:

الف) حداقل ۳ دستگاه سالم باشد.

ب) حداکثر ۲ دستگاه معیوب باشد.

ج) فقط ۱ دستگاه معیوب باشد.

پاسخ:

الف) حداقل ۳ دستگاه سالم باشد معادل این است که ۳ دستگاه سالم باشد یا ۴ دستگاه سالم باشد.

۳ دستگاه سالم:

تعداد ترکیبات ۳ تا از ۶ رایانه‌ی سالم $۶C۳=\frac{۶!}{۳!\times۳!}=۲۰$

تعداد ترکیبات ۱ از ۳ رایانه‌ی معیوب $۳C۱=\frac{۳!}{۱!\times۲!}=۳$

با استفاده از اصل دوم شمارش تعداد ۳ دستگاه رایانه‌ی سالم و ۱ دستگاه معیوب: $۲۰\times۳=۶۰$

۴ دستگاه سالم:

تعداد ترکیبات ۴ از ۶ رایانه‌ی سالم $۶C۴=\frac{۶!}{۴!\times۲!}=۱۵$

تعداد ترکیبات صفر از ۳ رایانه‌ی معیوب $۳C۰=\frac{۳!}{۳!\times۰!}=۱$

با استفاده از اصل دوم شمارش داریم: $۱۵\times۱=۱۵$

حال برای به دست آوردن حداقل ۳ دستگاه سالم با استفاده از اصل اول شمارش داریم: $۶۰+۱۵=۷۵$

ب) حداکثر ۲ دستگاه معیوب معادل این است که ۲ دستگاه معیوب یا یک دستگاه معیوب یا صفر دستگاه معیوب باشد مانند قسمت قبل داریم:

۲ دستگاه معیوب $\quad {}_3C_2 = \frac{3!}{2! \times 1!} = 3$

پس ۲ دستگاه سالم $\quad {}_6C_2 = \frac{6!}{2! \times 4!} = 15$

با استفاده از اصل دوم شمارش: $\quad {}_3C_2 \times {}_6C_2 = 3 \times 15 = 45$

۱ دستگاه معیوب $\quad {}_3C_1 = \frac{3!}{1! \times 2!} = 3$

پس ۳ دستگاه سالم $\quad {}_6C_3 = \frac{6!}{3! \times 3!} = 20$

با استفاده از اصل دوم شمارش: $\quad {}_3C_1 \times {}_6C_3 = 3 \times 20 = 60$

صفر دستگاه معیوب $\quad {}_3C_0 = \frac{3!}{0! \times 3!} = 1$

پس ۴ دستگاه سالم $\quad {}_6C_4 = \frac{6!}{4! \times 2!} = 15$

با استفاده از اصل دوم شمارش: $\quad {}_3C_0 \times {}_6C_4 = 1 \times 15 = 15$

حال با استفاده از اصل اول شمارش حداکثر ۲ رایانه معیوب باشد، یعنی: $\quad 45 + 60 + 15 = 120$

ج) فقط یک دستگاه معیوب باشد پس ۳ دستگاه دیگر باید سالم باشد داریم:

تعداد ترکیبات ۱ از ۳ دستگاه معیوب $\quad {}_3C_1 = \frac{3!}{1! \times 2!} = 3$

تعداد ترکیبات ۳ از ۶ دستگاه سالم $\quad {}_6C_3 = \frac{6!}{3! \times 3!} = 20$

با استفاده از اصل دوم شمارش داریم: $\quad {}_3C_1 \times {}_6C_3 = 3 \times 20 = 60$

نکته: برای سهولت در محاسبه‌ی بعضی ترکیبات می‌توان از رابطه‌های زیر استفاده کرد:

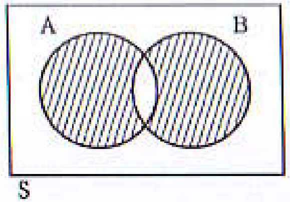
$nC_1 = n \qquad nC_{n-1} = n$

$nC_0 = 1 \qquad nC_n = 1$

تذکر: $0! = 1! = 1$

۸ـ پیشامد این که دقیقاً یکی از پیشامدهای A یا B رخ دهد را می‌توان به صورت $(A \cap B^C) \cup (A^C \cap B)$ بیان کرد نشان دهید که:

$$P[(A \cap B^C) \cup (A^C \cap B)] = P(A) + P(B) - 2P(A \cap B)$$



پاسخ: قسمت هاشورخورده در نمودار روبه‌رو پیشامد $[(A \cap B^C) \cup (A^C \cap B)]$ را نشان می‌دهد.

برهان: پیشامد A را می‌توان به دو پیشامد مجزای $A \cap B^C$ و $A \cap B$ تجزیه کرد.

$$A = (A \cap B^C) \cup (A \cap B)$$

رابطه‌ی (۱) $\quad P(A) = P(A \cap B^C) + P(A \cap B) \rightarrow P(A \cap B^C) = P(A) - P(A \cap B)$

همچنین پیشامد B را می‌توان به دو پیشامد مجزای $A^C \cap B$ و $A \cap B$ تجزیه کرد.

$$B = (A^C \cap B) \cup (A \cap B)$$

رابطه‌ی (۲) $\quad P(B) = P(A^C \cap B) + P(A \cap B) \rightarrow P(A^C \cap B) = P(B) - P(A \cap B)$

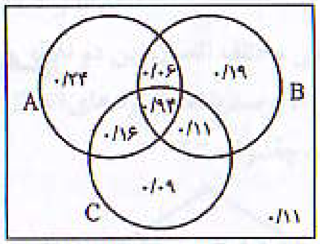
حال اگر بین دو رابطه‌ی ۱ و ۲ اجتماع بگیریم داریم:

$$P[(A \cap B^C) \cup (A^C \cap B)] = P(A) - P(A \cap B) + P(B) - P(A \cap B)$$
$$= P(A) + P(B) - 2P(A \cap B)$$

۹ـ الف) اگر A و B و C سه پیشامد از فضای نمونه‌ی S باشند. نشان دهید که:

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

ب) فرمول به دست آمده در الف را برای نمودار ون زیر امتحان کنید.



پاسخ: الف)

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

ابتدا با استفاده از قانون شرکت‌پذیری دو مجموعه‌ی A و B را داخل پرانتز قرار داده و C را جدا می‌گیریم و ۳ پیشامد را تبدیل به دو پیشامد می‌کنیم، با استفاده از قضیه‌ی ۲ـ۹ـ۶ برای دو پیشامد داریم:

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

رابطه‌ی (۱) $\quad = P(A) + P(B) - P(A \cap B) + P(C) - P[(A \cap C) \cup (B \cap C)]$

رابطه‌ی (۱) را با استفاده از رابطه‌ی اجتماع برای دو پیشامد باز می‌کنیم که پیشامد اول را $(A \cap C)$ و پیشامد دوم را $(B \cap C)$ می‌گیریم:

$$= P(A) + P(B) - P(A \cap B) + P(C) - [P(A \cap C) + P(B \cap C) - P[(A \cap C) \cap (B \cap C)]]$$

رابطه‌ی بالا را ساده می‌کنیم و منفی را داخل کروشه اثر می‌دهیم. پس داریم:

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

ب) با استفاده از اعدادی که داخل نمودار نوشته شده داریم:

$$P(A) = ۰/۲۴ + ۰/۱۶ + ۰/۰۴ + ۰/۰۶ = ۰/۵۰$$

$$P(B) = ۰/۱۹ + ۰/۱۱ + ۰/۰۴ + ۰/۰۶ = ۰/۴۰$$

$$P(C) = ۰/۰۹ + ۰/۱۱ + ۰/۱۶ + ۰/۰۴ = ۰/۴۰$$

$$P(A \cap B) = ۰/۰۶ + ۰/۰۴ = ۰/۱۰$$

$$P(A \cap C) = ۰/۱۶ + ۰/۰۴ = ۰/۲۰$$

$$P(B \cap C) = ۰/۰۴ + ۰/۱۱ = ۰/۱۵$$

$$P(A \cap B \cap C) = ۰/۰۴$$

حال با استفاده از فرمول الف داریم:

$$P(A \cup B \cup C) = ۰/۵۰ + ۰/۴۰ + ۰/۴۰ - ۰/۱۰ - ۰/۲۰ - ۰/۱۵ + ۰/۰۴ = ۰/۸۹$$

تذکر: همچنین $P(A \cup B \cup C)$ را می‌توان با استفاده از رابطه‌ی زیر به دست آورد:

$$P[(A \cup B \cup C) = ۱ - P(A \cup B \cup C)^C \rightarrow P(A \cup B \cup C) = ۱ - ۰/۱۱ = ۰/۸۹$$

۱۰ـ فرض کنید $P(A)=P(B)=\frac{۱}{۳}$ و $P(A \cap B)=\frac{۱}{۱۰}$ باشد احتمال‌های زیر را حساب کنید.

$P(A^C)$ $P(A \cup B^C)$ $P(B \cap A^C)$ $P(A^C \cup B^C)$

پاسخ:

* $P(A^C) = ۱ - P(A) = ۱ - \frac{۱}{۳} = \frac{۲}{۳}$

$P(B^C) = ۱ - P(B) = ۱ - \frac{۱}{۳} = \frac{۲}{۳}$

* $P(A \cup B^C) = P(A) + P(B^C) - P[A \cap B^C]$

با استفاده از قضیه‌ی ۲ـ۹ـ۵ داریم: $P(A \cap B^C) = P(A) - P(A \cap B)$

$$P(A \cup B^C) = P(A) + P(B^C) - [P(A) - P(A \cap B)]$$

$$= P(A) + P(B^C) - P(A) + P(A \cap B) = \frac{۲}{۳} + \frac{۱}{۱۰} = \frac{۲۰+۳}{۳۰} = \frac{۲۳}{۳۰}$$

* $P(B \cap A^C) = P(B) - P(A \cap B) = \frac{۱}{۳} - \frac{۱}{۱۰} = \frac{۱۰-۳}{۳۰} = \frac{۷}{۳۰}$

* $P(A^C \cup B^C) = P(A \cap B)^C = ۱ - P(A \cap B) = ۱ - \frac{۱}{۱۰} = \frac{۹}{۱۰}$

نکته: با استفاده از قانون دمورگان داریم:

$$(A \cap B)^C = A^C \cup B^C$$

۱۱ـ احتمال این که یک ایستگاه تلویزیون بعد از نمایش یک برنامه‌ی مناظره ۰، ۱، ۲، ۳، ۴، ...، ۸ و یا لااقل ۹ شکایت دریافت کند به ترتیب ۰/۰۱، ۰/۰۳، ۰/۰۷، ۰/۱۵، ۰/۱۹، ۰/۱۸، ۰/۱۴، ۰/۱۲، ۰/۰۹، ۰/۰۲ است احتمال‌های زیر را حساب کنید.

الف) بیش‌تر از ۴ شکایت دریافت کند.

ب) لااقل ۶ شکایت دریافت کند.

ج) از ۵ تا ۸ شکایت دریافت کند.

د) شکایتی دریافت نکند.

پاسخ: اگر تعریف کنیم $P(A_0)$ احتمال این که صفر شکایت دریافت کند، $P(A_1)$ احتمال این که یک شکایت دریافت کند و... و بالاخره $P(A_9)$ احتمال این که لااقل ۹ شکایت دریافت کند.

الف)

$$P(A_5) + P(A_6) + P(A_7) + P(A_8) + P(A_9) = ۰/۱۸ + ۰/۱۴ + ۰/۱۲ + ۰/۰۹ + ۰/۰۲ = ۰/۵۵$$

ب) لااقل ۶ شکایت دریافت کند یعنی ۶ شکایت یا بیش‌تر از آن دریافت کند.

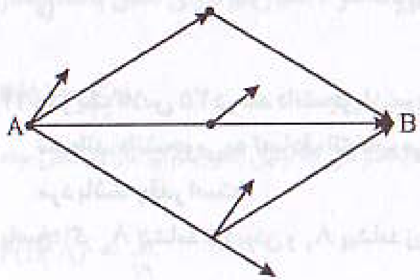
$$P(A_6) + P(A_7) + P(A_8) + P(A_9) = ۰/۱۴ + ۰/۱۲ + ۰/۰۹ + ۰/۰۲ = ۰/۳۷$$

ج)

$$P(A_5) + P(A_6) + P(A_7) + P(A_8) = ۰/۱۸ + ۰/۱۴ + ۰/۱۲ + ۰/۰۹ = ۰/۵۳$$

د) شکایتی دریافت نکند. $P(A_0) = ۰/۰۱$

۱۲ـ شخصی می‌خواهد از نقطه‌ی A به نقطه‌ی B برود. ولی راه‌های مختلف اتصال این دو نقطه را نمی‌داند و بسته به راهی که انتخاب می‌کند ممکن است به نقطه‌ی B برسد یا نه. اگر راه‌های اتصال به شکل زیر باشد احتمال این که این شخص به نقطه‌ی B برسد چقدر است؟



پاسخ: برای به دست آوردن احتمال، داریم:

$$P(A) = \frac{\text{تعداد حالات مساعد}}{\text{تعداد حالات کل}} = \frac{n(A)}{n(S)}$$

که در عبارت بالا $n(S)$ تعداد کل راه‌ها می‌باشد که ۷ راه وجود دارد و $n(A)$ راه‌هایی می‌باشد که از A شروع شده و به نقطه‌ی B ختم می‌شود که از این ۷ راه، ۳ راه به B ختم می‌شود.

$$P(A) = \frac{n(A)}{n(S)} = \frac{۳}{۷} = ۰/۴۳$$

۱۳ـ یک مؤسسه‌ی تحقیقاتی رایانه‌های موردنیاز خود را با مدل‌های ۱، ۲ و ۳ از چهار شرکت A، B، C و D خریداری می‌کند. آمار رایانه‌های خریداری شده در جدول زیر ثبت شده است:

مدل	A	B	C	D	جمع
۱	۲۰	۳۰	۱۵	۵	۷۰
۲	۱۰	۱۵	۱۰	۱۰	۴۵
۳	۱۵	۲۰	۲۵	۱۰	۷۰
جمع	۴۵	۶۵	۵۰	۲۵	۱۸۵

یک رایانه به تصادف از مؤسسه انتخاب می‌شود مطلوب است:

الف) احتمال این که رایانه مربوط به شرکت A باشد.

ب) احتمال این که رایانه مربوط به شرکت B یا C باشد.

ج) احتمال این که رایانه مربوط به شرکت D و دارای مدل ۳ باشد.

د) احتمال این که رایانه مربوط به شرکت D باشد به طوری که می‌دانیم دارای مدل ۲ است.

پاسخ: الف) $P(A)=\frac{۴۵}{۱۸۵}=۰/۲۴$

ب) چون دو شرکت B و C دو مجموعه‌ی ناسازگار یا جدا از هم می‌باشند. بنابراین $P(B\cap C)=۰$

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

$$P(B\cup C)=\frac{۶۵}{۱۸۵}+\frac{۵۰}{۱۸۵}=\frac{۱۱۵}{۱۸۵}=۰/۶۲$$

ج) $P(D\cap \text{مدل ۳})=\frac{۱۰}{۱۸۵}=۰/۰۵$

د) $$P(D|\text{مدل ۲})=\frac{P(D\cap \text{مدل ۲})}{P(\text{مدل ۲})}=\frac{\frac{۱۰}{۱۸۵}}{\frac{۴۵}{۱۸۵}}=\frac{۱۰}{۴۵}=۰/۲۲$$

۱۴ـ در یک کلاس ۳۵ درصد دانشجویان مرد هستند و ۲۰ درصد از مردان و ۲۵ درصد از زنان مردود شده‌اند. دانشجویی به تصادف انتخاب می‌شود. اگر این دانشجو مردود شده باشد، احتمال این که مرد باشد چقدر است؟

پاسخ: اگر A_1 پیشامد مرد بودن و A_2 پیشامد زن بودن و B پیشاد مردود شدن باشد، داریم:

$P(A_1)=۰/۳۵$ $\quad P(A_2)=۰/۶۵$

$P(B|A_1)=۰/۲۰$ $\quad P(B|A_2)=۰/۲۵$

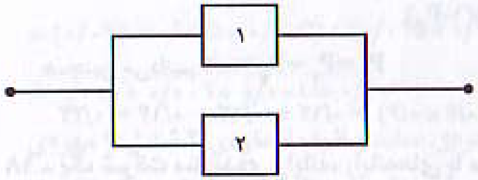
پس:
$$P(A_1|B)=\frac{P(B|A_1).P(A_1)}{P(B|A_1).P(A_1)+P(B|A_2).P(A_2)}$$

$$P(A_1|B)=\frac{۰/۲۰\times ۰/۳۵}{(۰/۲۰\times ۰/۳۵)+(۰/۲۵\times ۰/۶۵)}=۰/۳۰$$

۱۵ـ یک سیستم مهندسی دارای دو عامل است که به طور مستقل از هم عمل می‌کنند. اگر $P(\text{عامل اول خراب شود})=۰/۱$ و $P(\text{عامل دوم خراب شود})=۰/۲$ احتمال این که سیستم در موقعیت‌های زیر خراب نشود چقدر است؟

الف) دو عامل سری هستند.

ب) دو عامل موازی هستند.



پاسخ: الف) اگر A_1 را پیشامد این که عامل اول خراب شود و A_2 را پیشامد این که عامل دوم خراب شود در نظر بگیریم داریم:

$P(A_1^C\cap A_2^C)=P(A_1^C).P(A_2^C)$ $\qquad P(A_1^C)=۱-۰/۱=۰/۹$

$P(A_1^C\cap A_2^C)=۰/۹\times ۰/۸=۰/۷۲$ $\qquad P(A_2^C)=۱-۰/۲=۰/۸$

ب)
$$P(A_1\cap A_2^C)+P(A_1^C\cap A_2)+P(A_1^C\cap A_2^C)$$
$$=P(A_1).P(A_2^C)+P(A_1^C).P(A_2)+P(A_1^C).P(A_2^C)$$
$$=(۰/۱\times ۰/۸)+(۰/۹\times ۰/۲)+(۰/۹\times ۰/۸)=۰/۰۸+۰/۱۸+۰/۷۲=۰/۹۸$$

۱۶ـ جعبه‌ای شامل ۲۰ الگوریتم است که ۵ تا از آن‌ها ناقص است. اگر به تصادف ۳ الگوریتم متوالیاً و بدون جایگذاری از این جعبه انتخاب کنیم. احتمال این که هر ۳ الگوریتم دارای نقص باشد چقدر است؟

پاسخ: A : پیشامد این که اولین الگوریتم دارای نقص است. $\quad P(A)=\frac{۵}{۲۰}$

B| A : پیشامد این که دومین الگوریتم دارای نقص باشد به شرطی که اولین الگوریتم نیز دارای نقص بوده باشد.

$$P(B|A)=\frac{۴}{۱۹}$$

C| A∩B : پیشامد این که سومین الگوریتم دارای نقص باشد به شرطی که اولین و دومین الگوریتم نیز دارای نقص بوده باشد.

$$P(C|A\cap B)=\frac{۳}{۱۸}$$

A ∩ B ∩ C : پیشامد این که اولین و دومین و سومین الگوریتم دارای نقص است.

$$P(A\cap B\cap C)=P(A).P(B|A).P(C|A\cap B)=\frac{۵}{۲۰}\times\frac{۴}{۱۹}\times\frac{۳}{۱۸}=\frac{۱}{۱۱۴}=۰/۰۱$$

۱۷ـ احتمال این که در یک اندازه‌گیری کمیت فیزیکی خطا از حدی معین تجاوز کند ۰/۴ است. برای اندازه‌گیری این کمیت ۳ آزمایش مستقل انجام گرفته است. احتمال این که فقط در یکی از این آزمایش خطا از حد مجاز تجاوز کند چقدر است؟

پاسخ: اگر احتمال تجاوز خطا از حد مجاز در آزمایش اول را با $P_۱$ و در آزمایش دوم را با $P_۲$ و در آزمایش سوم با $P_۳$ نشان دهیم، داریم:

$$P_۱(۱-P_۲)(۱-P_۳)+P_۲(۱-P_۱)(۱-P_۳)+P_۳(۱-P_۱)(۱-P_۲)$$

همچنین می‌دانیم $P_۱=P_۲=P_۳=۰/۴$

$$(۰/۴\times ۰/۶\times ۰/۶)+(۰/۴\times ۰/۶\times ۰/۶)+(۰/۴\times ۰/۶\times ۰/۶)=۰/۱۴+۰/۱۴+۰/۱۴=۰/۴۲$$

۱۸ـ یک شرکت سازنده‌ی رایانه، رایانه‌های با مدل‌های ۱، ۲، ۳ با نسبت‌های ۲۰٪، ۳۰٪ و ۵۰٪ تولید می‌کند به طوری که به ترتیب ۵٪، ۳٪ و ۲٪ از مدل‌های تولید شده معیوب هستند. یک رایانه از تولیدات انتخاب می‌کنیم.

الف) احتمال این که معیوب باشد چقدر است؟

ب) می‌دانیم رایانه معیوب است. احتمال این که از مدل ۱ یا ۳ باشد چقدر است؟

پاسخ: A : پیشامد این که رایانه معیوب باشد.

$A_۱$: پیشامد این که رایانه با مدل ۱ باشد.

$A_۲$: پیشامد این که رایانه با مدل ۲ باشد.

$A_۳$: پیشامد این که رایانه با مدل ۳ باشد.

$$P(A_۱)=۰/۲۰ \qquad P(A_۲)=۰/۳۰ \qquad P(A_۳)=۰/۵۰$$

$$P(A|A_۱)=۰/۰۵ \qquad P(A|A_۲)=۰/۰۳ \qquad P(A|A_۳)=۰/۰۲$$

الف)

$$P(A)=\sum_{i=۱}^{۳}P(A|A_i)P(A_i)=(۰/۰۵\times ۰/۲۰)+(۰/۰۳\times ۰/۳۰)+(۰/۰۲\times ۰/۵۰)=۰/۰۲۹$$

ب) احتمال این که از مدل ۱ باشد:

$$P(A_۱|A)=\frac{P(A_۱\cap A)}{P(A)}=\frac{P(A|A_۱)P(A_۱)}{P(A)}=\frac{(۰/۰۵\times ۰/۲۰)}{۰/۰۲۹}=\frac{۰/۰۱}{۰/۰۲۹}=۰/۳۴۵$$

احتمال این که از مدل ۳ باشد:

$$P(A_۳|A)=\frac{P(A_۳\cap A)}{P(A)}=\frac{P(A|A_۳)P(A_۳)}{P(A)}=\frac{(۰/۰۲\times ۰/۵۰)}{۰/۰۲۹}=\frac{۰/۰۱}{۰/۰۲۹}=۰/۳۴۵$$

احتمال این که از مدل ۱ یا ۳ باشد به صورت مجموع دو احتمال بالا می‌باشد:

$$P(A_۱|A)+P(A_۳|A)=۰/۳۴۵+۰/۳۴۵=۰/۶۹$$

۱۹ـ یک سایت رایانه‌ای دارای سه رایانه‌ی فعال می‌باشد که مستقل از یک دیگر عمل می‌کنند. احتمال این که یکی از آن‌ها از کار بیفتد برابر ۰/۰۲۵ است. احتمال این که فقط یکی از آن‌ها فعال باشد، چقدر است؟

پاسخ: اگر از کار افتادن رایانه‌ی اول را با $P_۱$، رایانه‌ی دوم را با $P_۲$ و رایانه‌ی سوم را با $P_۳$ نشان دهیم، داریم:

$$P_۱.P_۲.(۱-P_۳)+P_۱.(۱-P_۲).P_۳+(۱-P_۱).P_۲.P_۳$$

$$=(۰/۰۲۵\times ۰/۰۲۵\times ۰/۹۷۵)+(۰/۰۲۵\times ۰/۹۷۵\times ۰/۰۲۵)+(۰/۹۷۵\times ۰/۰۲۵\times ۰/۰۲۵)$$

$$=۰/۰۰۱+۰/۰۰۱+۰/۰۰۱=۰/۰۰۳$$

۲۰ـ ظرف شماره‌ی ۱ شامل ۳ مهره‌ی قرمز و ۲ مهره‌ی سفید و ظرف شماره‌ی ۲ شامل ۲ مهره‌ی قرمز و ۵ مهره‌ی سفید است. یک تاس پرتاب می‌شود اگر تاس ۵ یا ۶ ظاهر شود یک مهره از ظرف ۱ انتخاب می‌شود و در غیر این صورت از ظرف شماره‌ی ۲. مطلوب است:

الف) احتمال این که یک مهره انتخاب شده سفید باشد.

ب) اگر مهره‌ی سفید انتخاب شود، احتمال این که از ظرف اول باشد چقدر است؟

پاسخ: در این جا ظرف‌ها با احتمال‌های متفاوت انتخاب می‌شوند. اگر $A_۱$ بیانگر پیشامد انتخاب ظرف ۱ و $A_۲$ بیانگر پیشامد انتخاب ظرف ۲ و A پیشامد سفید بودن مهره باشد داریم:

$$P(A_۱)=\frac{۲}{۶}=\frac{۱}{۳} \qquad P(A_۲)=\frac{۴}{۶}=\frac{۲}{۳}$$

$$P(A|A_۱)=\frac{۲}{۵} \qquad P(A|A_۲)=\frac{۵}{۷}$$

الف)

$$P(A)=\sum_{i=۱}^{۲}P(A|A_i)P(A_i)=(\frac{۲}{۵}\times\frac{۱}{۳})+(\frac{۵}{۷}\times\frac{۲}{۳})=\frac{۲}{۵}+\frac{۱۰}{۲۱}$$

$$=\frac{۴۲+۱۵۰}{۳۱۵}=\frac{۱۹۲}{۳۱۵}=۰/۶۱$$

ب)

$$P(A_۱|A)=\frac{P(A_۱\cap A)}{P(A)}=\frac{P(A|A_۱)P(A_۱)}{P(A)}=\frac{\frac{۲}{۵}\times\frac{۱}{۳}}{\frac{۱۹۲}{۳۱۵}}=\frac{۰/۱۳۳}{۰/۶۱}=۰/۲۱۸$$

خودآزمایی فصل سوم

۱ـ یک سکه‌ی ناریب سه بار پرتاب می‌شود. اگر متغیر X تعداد شیرها در سه پرتاب باشد تابع چگالی احتمال X را بنویسید و احتمالات زیر را حساب کنید.

$P(X=0)\ ,\ P(X=2)\ ,\ P(X\geq 3)\ ,\ P(X<3)$

پاسخ:

فضای نمونه $S=\{(HHH)\ (HHT)\ (HTH)\ (THH)\ (TTH)\ (THT)\ (HTT)\ (TTT)\}$

اگر متغیر تصادفی X را تعداد شیرها روی فضای نمونه‌ی S و A را فضای نمونه‌ی وابسته به X تعریف کنیم داریم:

$A=\{0\ ,\ 1\ ,\ 2\ ,\ 3\}$

تعداد شیرها X	0	1	2	3
P(X=x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

که تابع چگالی آن را به صورت زیر نیز می‌توان نوشت:

$$f(x)=P(X=x)=\frac{\binom{3}{x}}{8} \qquad x=0,1,2,3$$

$$P(X=0)=f(0)=\frac{\binom{3}{0}}{8}=\frac{1}{8}$$

$$P(X=2)=f(2)=\frac{\binom{3}{2}}{8}=\frac{3}{8}$$

$$P(X\geq 3)=P(X=3)=\frac{\binom{3}{3}}{8}=\frac{1}{8}$$

$$P(X<3)=P(X=0)+P(X=1)+P(X=2)$$

$$=\frac{\binom{3}{0}}{8}+\frac{\binom{3}{1}}{8}+\frac{\binom{3}{2}}{8}=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}=\frac{7}{8}$$

۲ـ سکه‌ای طوری تنظیم شده که شانس ظاهر شدن شیر در آن دو برابر خط است. اگر متغیر X تعداد شیرها در سه پرتاب باشد تابع چگالی احتمال X را بنویسید.

پاسخ: سکه به گونه‌ای درست شده که $P(H)=2P(T)$ از طرفی می‌دانیم، $P(T)+P(H)=1$.

پس با حل دو معادله‌ی فوق داریم: $P(T)=\frac{1}{3}\quad ,\quad P(H)=\frac{2}{3}$

فضای نمونه‌ی آزمایش عبارت است از:

$S=\{(HHH)\ (HHT)\ (HTH)\ (THH)\ (TTH)\ (THT)\ (HTT)\ (TTT)\}$

و اگر X را تعداد شیرها در سه پرتاب در نظر بگیریم: $A=\{0\ ,\ 1\ ,\ 2\ ,\ 3\}$

مقادیر مختلف X و احتمال X=x در جدولی به صورت زیر آمده است:

S	HHH	HTH	HHT	THH	TTH	THT	HTT	TTT
X	3	2	2	2	1	1	1	0
P(X=x)	$(\frac{2}{3})^3$	$(\frac{1}{3})(\frac{2}{3})^2$	$(\frac{1}{3})(\frac{2}{3})^2$	$(\frac{1}{3})(\frac{2}{3})^2$	$(\frac{1}{3})^2(\frac{2}{3})$	$(\frac{1}{3})^2(\frac{2}{3})$	$(\frac{1}{3})^2(\frac{2}{3})$	$(\frac{1}{3})^3$

مثلاً: $P(X=0)=P(\{TTT\})=P(T)\ .\ P(T)\ .\ P(T)=\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{27}$

که به صورت خلاصه می‌توان نوشت:

S	HHH	HHT,HTH,THH	THH,HTT,THT	TTT
X	3	2	1	0
P(X=x)	$\frac{8}{27}$	$\frac{4}{27}+\frac{4}{27}+\frac{4}{27}=\frac{12}{27}$	$\frac{2}{27}+\frac{2}{27}+\frac{2}{27}=\frac{6}{27}$	$\frac{1}{27}$

همچنین اگر احتمال شیر آمدن را برابر $\frac{2}{3}$ و خط آمدن را $\frac{1}{3}$ در نظر بگیریم تابع احتمال را به صورت زیر نیز می‌توان نوشت:

$$f(x)=P(X=x)=\binom{3}{x}\left[\frac{2}{3}\right]^x\left[\frac{1}{3}\right]^x \quad ; \quad x=0,1,2,3$$

۲ـ یک سکه‌ی ناریب را آن قدر پرتاب می‌کنیم که اولین شیر ظاهر شود. اگر متغیر تصادفی X تعداد پرتاب‌های موردنیاز برای ظاهر شدن اولین شیر باشد فضای نمونه‌ی X را بنویسید و یک قالب احتمال برای آن بنویسید.

پاسخ: فضای نمونه‌ی آزمایش عبارت است از: $S=\{H\ ,\ TH\ ,\ TTH\ ,\ TTTH\ ,\ ...\ \}$

چون سکه ناریب است پس شانس شیر آمدن برابر است با شانس خط آمدن یعنی:

$P(H)=P(T)=\frac{1}{2}$

S	H	TH	TTH	TTTH	TTTTH	...
X	1	2	3	4	5	...
P(X=x)	$\frac{1}{2}$	$(\frac{1}{2})^1\times\frac{1}{2}$	$(\frac{1}{2})^2\times\frac{1}{2}$	$(\frac{1}{2})^3\times\frac{1}{2}$	$(\frac{1}{2})^4\times\frac{1}{2}$	...

اگر $P(H)=P$ و $P(T)=q$ در نظر بگیریم می‌توان تابع چگالی را به صورت زیر تعریف کرد:

$f(x)=P(X=x)=p.q^{x-1} \qquad x=1\ ,\ 2\ ,\ 3\ ,\ ...$

۴ـ سایتی دارای ۶ رایانه است که دو تای آن معیوب هستند. اگر به تصادف دو تا از این رایانه‌ها را انتخاب کنیم و متغیر تصادفی X تعداد رایانه‌ی معیوب باشد تابع چگالی احتمال و تابع توزیع X را به دست آورید.

پاسخ: اگر متغیر تصادفی X را تعداد رایانه‌های معیوب در نظر بگیریم: $X=\{0,1,2\}$

$$P(X=x)=f(x)=\frac{\binom{2}{x}\binom{4}{2-x}}{\binom{6}{2}} \qquad x=0,1,2$$

$$P(X=0)=\frac{\binom{2}{0}\binom{4}{2}}{\binom{6}{2}}=\frac{1\times 6}{15}=\frac{6}{15}$$

$$P(X=1)=\frac{\binom{2}{1}\binom{4}{1}}{\binom{6}{2}}=\frac{2\times 4}{15}=\frac{8}{15}$$

$$P(X=2)=\frac{\binom{2}{2}\binom{4}{0}}{\binom{6}{2}}=\frac{1\times 1}{15}=\frac{1}{15}$$

که می‌توان مقادیر احتمال‌ها را به صورت زیر در جدول نوشت:

تعداد رایانه‌های معیوب X	۰	۱	۲	جمع
$P(X=x)$	$\frac{6}{15}$	$\frac{8}{15}$	$\frac{1}{15}$	۱

و تابع توزیع آن به صورت زیر می‌باشد:

$$F(x)=\sum_{x\le t} f(t)$$

$$F(-\infty)=P(X<0)=0$$

$$F(0)=P(X\le 0)=P(X=0)=\frac{6}{15}$$

$$F(1)=P(X\le 1)=P(X=0)+(X=1)=\frac{6}{15}+\frac{8}{15}=\frac{14}{15}$$

$$F(2)=P(X\le 2)=P(X=0)+(X=1)+P(X=2)=\frac{6}{15}+\frac{8}{15}+\frac{1}{15}=\frac{15}{15}=1$$

$$F(x)=\begin{cases} 0 & \text{if} \quad x<0 \\ \frac{6}{15} & \text{if} \quad 0\le x<1 \\ \frac{14}{15} & \text{if} \quad 1\le x<2 \\ 1 & \text{if} \quad x\ge 2 \end{cases}$$

۵ـ جعبه‌ای شامل چهار مهره به شماره‌های ۱ و ۲ و ۳ و ۴ است. دو مهره به تصادف از جعبه استخراج می‌کنیم. اگر X عدد بزرگ‌تر دو مهره انتخابی باشد تابع چگالی احتمال و تابع توزیع X را به دست آورید.

پاسخ: فضای نمونه‌ی آزمایش به صورت زیر است:

$$S=\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$$

تابع چگالی احتمال آن به صورت زیر است:

$X=x$	۱	۲	۳	۴	جمع
$f(x)$	۰	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	۱

و تابع توزیع آن به صورت زیر می‌باشد:

$$F(1)=P(X\le 1)=P(X=1)=0$$

$$F(2)=P(X\le 2)=P(X=1)+P(X=2)=0+\frac{1}{6}=\frac{1}{6}$$

$$F(3)=P(X\le 3)=P(X=1)+P(X=2)+P(X=3)=0+\frac{1}{6}+\frac{2}{6}=\frac{3}{6}$$

$$F(4)=P(X\le 4)=P(X=1)+P(X=2)+P(X=3)+P(X=4)=0+\frac{1}{6}+\frac{2}{6}+\frac{3}{6}=1$$

$$F(x)=\begin{cases} 0 & 1\le x<2 \\ \frac{1}{6} & 2\le x<3 \\ \frac{3}{6} & 3\le x<4 \\ 1 & x\ge 4 \end{cases}$$

۶ـ از فضای نمونه‌ی $S=\{s|\ 0<s<10\}$ یک نقطه اختیار می‌کنیم. فرض کنید $S'\subset S$. اگر تابع مجموعه‌ی احتمال $P(S')$ به صورت زیر باشد:

$$P(S')=\int_{S'}\frac{1}{10}\,dz$$

و $X=X(s)=2s-10$ تعریف کنیم تابع احتمال متغیر تصادفی X را پیدا کنید.

پاسخ: با توجه به فضای نمونه‌ی $S=\{s|\ 0<s<10\}$، فضای نمونه برای X را به دست می‌آوریم:

$$X=X(s)=2s-10$$

فضای نمونه X $\quad A=\{x|\ -10<x<10\}$

بنابراین داریم:

$$P_X(A)=P(A)=\int_A \frac{1}{20}\,dz=\int_{-10}^{10}\frac{1}{20}\,dz=\int_{0}^{1}\frac{1}{10}\,dz$$

۷ـ تابع f(x) به صورت زیر تعریف شده است:

$$f(x)=c\left[\left[\frac{1}{2}\right]^{x}-\frac{1}{2}\right] \qquad x=0,1,2,\dots$$

الف) به ازای چه مقدار c تابع f(x) یک تابع چگالی احتمال است.

ب) تابع توزیع آن را به دست آورید.

پاسخ: الف) ابتدا مقدار c را محاسبه می‌کنیم، با توجه به تعریف تابع چگالی احتمال $\sum_{x\in X} f(x)=1$ می‌باشد.

پس داریم:

$$\sum_{x=0}^{+\infty}\left[c\left[\left[\frac{1}{2}\right]^{x}-\frac{1}{2}\right]\right]=1$$

$$\to c\left[\sum_{x=0}^{+\infty}\left[\frac{1}{2}\right]^{x}-\sum_{x=0}^{+\infty}\frac{1}{2}\right]=1$$

از طرفی می‌دانیم $\sum_{x=0}^{+\infty}\left[\frac{1}{2}\right]^{x}$ یک سری هندسی است و زمانی همگراست که $r<1$ در نظر گرفته شود

همچنین $\sum_{x=0}^{+\infty}\frac{1}{2}=\frac{1}{2}\sum_{x=0}^{+\infty}1$ که این مقدار برابر بی‌نهایت می‌باشد و این تساوی هیچ‌گاه برقرار نیست پس f(x) نمی‌تواند تابع چگالی احتمال باشد.

نکته: مجموع و تفاضل یک سری همگرا و یک سری واگرا، واگرا می‌باشد.

ب) با توجه به این که تابع داده شده، تابع چگالی احتمال نمی‌باشد پس فاقد تابع توزیع نیز می‌باشد.

۸ـ متغیر تصادفی X دارای تابع چگالی احتمال به صورت زیر است:

$$f(x)=c(8-x) \qquad x=0,1,2,3,4,5$$

الف) مقدار c و F(x) را پیدا کنید.

ب) احتمال P(X> 2) را حساب کنید.

پاسخ: الف)

$$\sum_{x=0}^{5} f(x)=1\to f(0)+f(1)+f(2)+f(3)+f(4)+f(5)=1$$

$$\to 8c+7c+6c+5c+4c+3c=1$$

$$\to 33c=1\to c=\frac{1}{33}$$

بنابراین تابع چگالی احتمال به صورت زیر است:

$$f(x)=\frac{1}{33}(8-x) \qquad x=0,1,2,3,4,5$$

تابع به صورت زیر می‌باشد:

$$F(x)=P(X\le x)=\sum_{x\le t} f(t)$$

$$F(x)=\begin{cases}0 & \text{if } x<0\\ \frac{8}{33} & \text{if } 0\le x<1\\ \frac{15}{33} & \text{if } 1\le x<2\\ \frac{21}{33} & \text{if } 2\le x<3\\ \frac{26}{33} & \text{if } 3\le x<4\\ \frac{30}{33} & \text{if } 4\le x<5\\ 1 & \text{if } x\ge 5\end{cases}$$

ب)

$$P(X>2)=1-P(X\le 2)=1-[P(X=0)+P(X=1)+P(X=2)]$$

$$=1-\left[\frac{8}{33}+\frac{7}{33}+\frac{6}{33}\right]=1-\frac{21}{33}=\frac{12}{33}$$

یا می‌توان به صورت زیر نوشت:

$$P(X>2)=P(X=3)+P(X=4)+P(X=5)=\frac{5}{33}+\frac{4}{33}+\frac{3}{33}=\frac{12}{33}$$

۹ـ اگر X دارای تابع توزیع F(x) باشد:

$$F(x)=\begin{cases}0 & x<0\\ 1-\left(\frac{1}{2}\right)^{x+1} & x=0,1,2,\dots\end{cases}$$

الف) تابع چگالی احتمال X را پیدا کنید.

ب) احتمال $P(10< X<\ \le 20)$ را حساب کنید.

ج) مقدار P[X زوج باشد] را بیابید.

پاسخ: الف) X متغیر گسسته است با فضای $A=\{0,1,2,\dots\}$

$$F(-\infty)=0=P(X\le-\infty)$$

با توجه به خواص داریم:

$$f(x)=F(x)-F(x^-)$$

$$f(0)=F(0)-F(0^-)=\left(1-\left(\frac{1}{2}\right)^1\right)-0=\frac{1}{2}$$

$$f(1)=F(1)-F(0)=\left(1-\left(\tfrac{1}{2}\right)^2\right)-\left(1-\left(\tfrac{1}{2}\right)^1\right)=1-\frac{1}{2^2}-1+\frac{1}{2}=\frac{1}{2^2}$$

$$f(2)=F(2)-F(1)=\left(1-\left(\tfrac{1}{2}\right)^3\right)-\left(1-\left(\tfrac{1}{2}\right)^2\right)=1-\frac{1}{2^3}-1+\frac{1}{2^2}=\frac{1}{2^3}$$

$$f(3)=F(3)-F(2)=\left(1-\left(\tfrac{1}{2}\right)^4\right)-\left(1-\left(\tfrac{1}{2}\right)^3\right)=1-\frac{1}{2^4}-1+\frac{1}{2^3}=\frac{1}{2^4}$$

$$f(4)=F(4)-F(3)=\left(1-\left(\tfrac{1}{2}\right)^5\right)-\left(1-\left(\tfrac{1}{2}\right)^4\right)=1-\frac{1}{2^5}-1+\frac{1}{2^4}=\frac{1}{2^5}$$

چون متغیر تصادفی X گسسته و نامتناهی است پس می‌توان تابع چگالی را به صورت زیر نوشت:

$$f(x)=\begin{cases}\dfrac{1}{2^{x+1}} & x=0,1,2,\dots\\ 0 & \text{سایر جاها}\end{cases}$$

ب) با توجه به محاسبه‌ی احتمال از روی تابع توزیع داریم:

$$P(a<X\le b)=F(b)-F(a^-)$$

$$P(10<X\le 20)=F(20)-F(10^-)=\left[1-\left[\frac{1}{2}\right]^{21}\right]-\left[1-\left[\frac{1}{2}\right]^{10}\right]$$

$$=1-\frac{1}{2^{21}}-1+\frac{1}{2^{10}}=\frac{-1+2^{11}}{2^{21}}=\frac{2^{11}-1}{2^{21}}$$

ج) $P[\text{زوج باشد } X]=P(X\le 2k)=F(2k)$

۱۰ـ یک متغیر تصادفی پیوسته‌ی X که مقادیرش بین $x=2$ و $x=5$ فرض شده دارای تابع چگالی $f(x)=\dfrac{2(1+x)}{27}$ است. تابع توزیع X را به دست آورید و احتمالات زیر را حساب کنید.

$$P(X<4)\quad,\quad P(3\le X<4)\quad,\quad P(0<X<3)$$

پاسخ: تابع چگالی متغیر تصادفی پیوسته‌ی X به صورت زیر می‌باشد:

$$f(x)=\frac{2(1+x)}{27}\qquad 2\le x\le 5$$

تابع توزیع X به صورت زیر می‌باشد:

$$F(x)=P(X\le x)=\int_{-\infty}^{x}f(t)\,dt$$

$$F(x)=\begin{cases}0 & \text{if}\quad x<2\\ \displaystyle\int_2^x\frac{2(1+t)}{27}dt=\frac{2}{27}\left[t+\frac{t^2}{2}\right]_2^x=\frac{x^2+2x-8}{27} & \text{if}\quad 2\le x<5\\ 1 & \text{if}\quad x\ge 5\end{cases}$$

ابتدا احتمالات داده شده را با استفاده از تابع چگالی محاسبه می‌کنیم:

$$P(X<4)=\int_{-\infty}^{4}\frac{2(1+x)}{27}dx=\int_{-\infty}^{2}0\,dx+\int_2^4\frac{2(1+x)}{27}dx$$

$$=\frac{2}{27}\left[x+\frac{x^2}{2}\right]_2^4=\frac{2}{27}\left[4+8-2-2\right]=\frac{16}{27}$$

$$P(3\le x<4)=\int_3^4\frac{2(1+x)}{27}dx=\frac{2}{27}\int_3^4(1+x)dx=\frac{2}{27}\left[x+\frac{x^2}{2}\right]_3^4$$

$$=\frac{2}{27}\left[4+8-3-\frac{9}{2}\right]=\frac{9}{27}$$

$$P(0<x<3)=\int_0^3\frac{2(1+x)}{27}dx=\int_0^2 0\,dx+\int_2^3\frac{2(1+x)}{27}dx$$

$$=\frac{2}{27}\left[x+\frac{x^2}{2}\right]_2^3=\frac{2}{27}\left[3+\frac{9}{2}-2-2\right]=\frac{7}{27}$$

همچنین می‌توان احتمالات بالا را با استفاده از خواص تابع توزیع به دست آورد. داریم:

$$P(X<4)=P(-\infty<X<4)=F(4)-F(-\infty)=\frac{16}{27}-0=\frac{16}{27}$$

$$P(3\le X<4)=F(4)-F(3)=\frac{16}{27}-\frac{7}{27}=\frac{9}{27}$$

$$P(0<X<3)=F(3)-F(0)=\frac{7}{27}-0=\frac{7}{27}$$

یادآوری:

$$P(a\le X\le b)=F(b)-F(a)$$

$$\lim_{x\to-\infty}F(x)=F(-\infty)=0\quad,\quad\lim_{x\to+\infty}F(x)=F(+\infty)=1$$

نکته:

$$\int_a^a f(x)\,dx=0$$

پس:

$$P(a<x\le b)=P(a\le x<b)=P(a\le x\le b)=P(a<x<b)$$

۱۱ـ به ازای چه مقدار c تابع زیر یک تابع چگالی است؟

$$f(x)=cx^{-(c+1)}\qquad 0<x<\infty$$

پاسخ: برای این که تابع f(x) به ازای مقدار c یک تابع چگالی باشد باید:

$$\int_{-\infty}^{+\infty}f(x)\,dx=1\ \rightarrow\ \int_0^{+\infty}cx^{-(c+1)}dx=c\left[\frac{x^{-(c+1)+1}}{-(c+1)+1}\right]_0^{\infty}$$

$$=c\times\frac{1}{-c}x^{-c}=-x^{-c}\Big|_0^{+\infty}=-\infty^{-c}-0=-\infty^{-c}$$

که مقدار این انتگرال به ازای هیچ مقداری برابر یک نمی‌شود پس تابع بالا تابع چگالی احتمال نمی‌باشد.

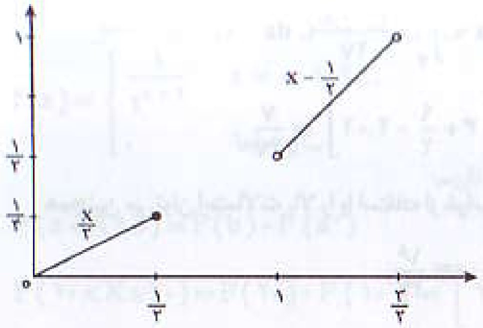
۱۲ـ متغیر تصادفی X دارای تابع توزیعی است به گونه‌ای که:

$$F(x)=\begin{cases}\dfrac{x}{2} & 0<x\le\dfrac{1}{2}\\ x-\dfrac{1}{2} & 1<x<\dfrac{3}{2}\end{cases}$$

الف) نمودار F(x) را رسم کنید.

ب) مطلوب است $P(X<\frac{1}{2})$

پاسخ: الف)



ب) $P\left(X < \frac{1}{2}\right) = P\left(-\infty < X < \frac{1}{2}\right) = F\left(\frac{1}{2}\right) - F(-\infty) = \frac{1}{4} - 0 = \frac{1}{4}$

۱۳ـ در یک فروشگاه رایانه، تقاضای سالانه برای یک بسته‌بندی نرم‌افزار یک متغیر تصادفی گسسته است. صاحب فروشگاه چهار نسخه از بسته‌بندی را با قیمت ۸۰۰۰ تومان برای هر نسخه سفارش می‌دهد و هر نسخه را به ۲۸۰۰۰ تومان به مشتری‌ها می‌فروشد. در پایان سال بسته‌بندی منسوخ می‌شود و فروشنده سرمایه‌اش را برای نسخه‌های فروش نرفته از دست می‌دهد. اگر تابع توزیع X به صورت زیر باشد تابع احتمال X را به دست آورید و احتمالات زیر را حساب کنید.

$P(X < 2)\ ,\quad P(1 < X < 3)\ ,\quad P(X = 4)$

$$F(X) = \begin{cases} 0 & x < 0 \\ 0/1 & 0 \le x < 1 \\ 0/4 & 1 \le x < 2 \\ 0/7 & 2 \le x < 3 \\ 0/9 & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$$

پاسخ: متغیر X گسسته است با فضای $A=\{0\ ,\ 1\ ,\ 2\ ,\ 3\ ,\ 4\}$

با توجه به خواص تابع توزیع داریم:

$F(-\infty) = 0 = P(X \le -\infty)$

$f(x) = F(x) - F(x^-)$

$f(0) = F(0) - F(0^-) = 0/1 - 0 = 0/1$

$f(1) = F(1) - F(0) = 0/4 - 0/1 = 0/3$

$f(2) = F(2) - F(1) = 0/7 - 0/4 = 0/3$

$f(3) = F(3) - F(2) = 0/9 - 0/7 = 0/2$

$f(4) = F(4) - F(3) = 1 - 0/9 = 0/1$

X	0	1	2	3	4	جمع
f(x)	0/1	0/3	0/3	0/2	0/1	1

$P(X < 2) = P(X = 0) + P(X = 1) = 0/1 + 0/3 = 0/4$

$P(1 < X < 3) = P(X = 2) = 0/3$

$P(X = 4) = 0/1$

۱۴ـ متغیر تصادفی X دارای تابع چگالی احتمال زیر می‌باشد. عدد m را میانه‌ی توزیع گوییم اگر داشته باشیم $P(X \le m) = P(X \ge m) = \frac{1}{2}$. میانه را برای متغیر X پیدا کنید.

$f(x) = \frac{2x}{9} \qquad 0 < x < 3$

پاسخ: باید مقدار انتگرال را تا نقطه‌ی m برابر $\frac{1}{2}$ (میانه) قرار دهیم.

$$P(X \le m) = \int_{-\infty}^{m} f(x)\,dx = \frac{1}{2}$$

$$\rightarrow \int_0^m \frac{2x}{9}\,dx = \frac{1}{2} \rightarrow \frac{2}{9}\int_0^m x\,dx = \frac{2}{9}\left[\frac{x^2}{2}\right]_0^m = \frac{1}{2}$$

$$\rightarrow \frac{m^2}{9} = \frac{1}{2} \quad \begin{cases} m = \frac{+3}{\sqrt{2}} & \text{قابل قبول} \\ m = \frac{-3}{\sqrt{2}} & \text{غیرقابل قبول} \end{cases}$$

بنابراین $m = \frac{3}{\sqrt{2}}$ میانه‌ی موردنظر ما می‌باشد چون $0<x<3$ می‌باشد.

۱۵ـ ضرایب معادله‌ی درجه‌ی دوم $x^2+bx+c=0$ را با ریختن دو تاس سالم تعیین می‌کنیم، به طوری که اولین برآمد b و دومین برآمد c در نظر گرفته شود. مطلوب است:

الف) فضای نمونه برای (b , c)

ب) احتمال این که معادله دارای جواب باشد چقدر است؟

پاسخ: الف)

$S=\{(b,c) \mid (b,c) \in \{(1,1)\ ,\ (1,2)\ ,\ (1,3)\ ,\ (1,4)\ ,\ \ldots\ ,\ (6,4)\ ,\ (6,5)\ ,\ (6,6)\}\}$

پس $n(S)=36$ است.

ب) می‌دانیم معادله‌ی $x^2 + bx + c = 0$ زمانی دارای جواب می‌باشد که $\Delta \geq 0$ * باشد از طرفی $\Delta = b^2 - 4ac$ ** می‌باشد. از مقایسه‌ی * و ** به این نتیجه می‌رسیم که باید $b^2 \geq 4ac$ باشد. همان طور که در معادله‌ی $x^2 + bx + c = 0$ ملاحظه می‌کنیم مقدار $a = 1$ است، بنابراین $b^2 \geq 4c$ می‌باشد. از این مطلب استفاده نموده و مجموعه‌ی A را به صورت زیر تعریف می‌کنیم:

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

$$P(A) = \frac{n(A)}{n(S)} = \frac{18}{36} = \frac{1}{2}$$

۱۶ـ از بسته‌ای که شامل ۳ رایانه با مدل I، ۲ رایانه با مدل II و ۳ رایانه با مدل III است یک نمونه‌ی تصادفی ۴ تایی انتخاب می‌کنیم. اگر متغیرهای تصادفی X و Y به ترتیب تعداد رایانه‌های مدل I و II باشند.

الف) تابع احتمال توأم (X , Y) را به دست آورید.

ب) احتمال $P[(X , Y) \in A]$ را به طوری که $A = \{(x , y) \mid x + y \leq 2\}$ باشد حساب کنید.

پاسخ: الف) تعداد کل نقاط فضای نمونه برابر است با:

$$n(S) = \binom{8}{4} = \frac{8!}{4! \times 4!} = 70$$

اگر x را تعداد رایانه‌های مدل I و y را تعداد رایانه‌های مدل II تعریف کنیم، $4-x-y$ تعداد رایانه‌های مدل III می‌باشد، و احتمال‌های مربوطه را می‌توان از فرمول زیر محاسبه کرد:

$$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{4-x-y}}{\binom{8}{4}} \qquad \begin{array}{l} x = 0, 1, 2, 3 \\ y = 0, 1, 2 \\ 1 \leq x + y \leq 4 \end{array}$$

جدول توزیع احتمال توأم X و Y به صورت زیر بیان می‌شود:

Y \ X	0	1	2	3	$f_Y(y)$
0	0	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	$\frac{15}{70}$
1	$\frac{2}{70}$	$\frac{18}{70}$	$\frac{18}{70}$	$\frac{2}{70}$	$\frac{40}{70}$
2	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	0	$\frac{15}{70}$
$f_X(x)$	$\frac{5}{70}$	$\frac{30}{70}$	$\frac{30}{70}$	$\frac{5}{70}$	1

ب) برای به دست آوردن احتمال داده شده از روی توزیع احتمال توأم باید حالت‌هایی را در نظر بگیریم که $x + y \leq 2$ است.

$$P[(x,y) \mid x + y \leq 2] = f(0,0) + f(0,1) + f(0,2) + f(1,0) + f(1,1) + f(2,0)$$

$$= 0 + \frac{2}{70} + \frac{3}{70} + \frac{3}{70} + \frac{18}{70} + \frac{9}{70} = \frac{35}{70} = 0/5$$

۱۷ـ اگر X و Y دارای تابع احتمال توأم f(x , y) باشد مطلوب است:

$$f(x , y) = c\frac{2^{x+y}}{x!\,y!} \qquad x = 0, 1, 2 \qquad y = 0, 1, 2$$

الف) مقدار c

ب) تابع چگالی حاشیه‌ای X و Y

ج) آیا X و Y از هم مستقل‌اند (راهنمایی $\sum_{x=0}^{\infty} \frac{2^x}{x!} = e^2$)

پاسخ: الف) برای به دست آوردن مقدار c با توجه به شرط دوم تابع چگالی داریم:

$$\sum_{x=0}^{2} \sum_{y=0}^{2} f(x,y) = 1$$

$$\sum_{x=0}^{2} \sum_{y=0}^{2} c\frac{2^{x+y}}{x!\,y!} = 1$$

$$\rightarrow f(0,0) + f(0,1) + f(0,2) + f(1,0) + f(1,1) + f(1,2) + f(2,0) + f(2,1) + f(2,2) = 1$$

$$\rightarrow c + 2c + 2c + 2c + 4c + 4c + 2c + 4c + 4c = 1 \rightarrow 25c = 1 \rightarrow c = \frac{1}{25}$$

بنابراین تابع احتمال توأم به صورت زیر است:

$$f(x , y) = \frac{2^{x+y}}{25x!\,y!} \qquad \begin{array}{l} x = 0, 1, 2 \\ y = 0, 1, 2 \end{array}$$

ب) برای به دست آوردن تابع چگالی حاشیه‌ای X داریم:

$$f(x) = \sum_{y} f(x,y) = \sum_{y=0}^{2} \frac{2^{x+y}}{25x!\,y!} = \frac{1}{25}\left[\frac{2^x}{x!} + \frac{2^{x+1}}{x!} + \frac{2^{x+2}}{2x!}\right]$$

با استفاده از خاصیت توان‌ها داریم: $a^x \cdot a^y = a^{x+y}$

$$= \frac{1}{25}\left[\frac{2^x}{x!} + \frac{2^x \cdot 2^1}{x!} + \frac{2^x \cdot 2^2}{2x!}\right] = \frac{1}{25}\left[\frac{2^x\,[1+2+2]}{x!}\right] = \frac{5}{25} \cdot \frac{2^x}{x!} = \frac{2^x}{5x!}$$

پس تابع چگالی حاشیه‌ای X به صورت زیر است:

$$f_X(x) = \frac{2^x}{5x!} \quad ; \quad x = 0, 1, 2$$

همچنین برای به دست آوردن تابع چگالی حاشیه‌ای Y داریم:

$$f(y)=\sum_x f(x,y)=\sum_{x=0}^{2}\frac{2^{x+y}}{25x!y!}=\frac{1}{25}\left[\frac{2^y}{y!}+\frac{2^{y+1}}{y!}+\frac{2^{y+2}}{2y!}\right]$$

$$=\frac{1}{25}\left[\frac{2^y[1+2+2]}{y!}\right]=\frac{5}{25}\cdot\frac{2^y}{y!}=\frac{2^y}{5y!}$$

پس تابع چگالی حاشیه‌ای Y به صورت زیر است:

$$f_Y(y)=\frac{2^y}{5y!}\quad ;\quad y=0,1,2$$

ج) دو متغیر تصادفی X و Y مستقلند اگر:

$$f(x,y)=f_X(x)\cdot f_Y(y)$$

$$f_X(x).f_Y(y)=\frac{2^x}{5x!}\times\frac{2^y}{5y!}=\frac{2^{x+y}}{25x!y!}=f(x,y)$$

پس دو متغیر تصادفی X و Y از هم مستقلند.

۱۸ـ فرض کنید X_1 و X_2 متغیرهای تصادفی گسسته با تابع توأم به صورت جدول زیر باشند:

$X_1 \backslash X_2$	۱	۲	۳	$f(x_1)$
۱	$\frac{1}{12}$	$\frac{1}{6}$	۰	$\frac{3}{12}$
۲	۰	$\frac{1}{9}$	$\frac{1}{5}$	$\frac{14}{45}$
۳	$\frac{1}{18}$	$\frac{1}{4}$	$\frac{2}{15}$	$\frac{79}{180}$
$f(x_2)$	$\frac{5}{36}$	$\frac{19}{36}$	$\frac{5}{15}$	۱

الف) تابع چگالی حاشیه‌ای X_1 و X_2 را به دست آورید.

ب) آیا X_1 و X_2 مستقل‌اند؟

ج) $P[X_1 \le X_2]$ را حساب کنید.

پاسخ: الف) تابع احتمال حاشیه‌ای X_1 و X_2 از روابط زیر به دست می‌آید که داخل جدول محاسبه شده است:

$$f(x_1)=\sum_{x_2} f(x_1,x_2)$$

X_1	۱	۲	۳	جمع
$f(x_1)$	$\frac{3}{12}$	$\frac{14}{45}$	$\frac{79}{180}$	۱

$$f(x_2)=\sum_{x_1} f(x_1,x_2)$$

X_2	۱	۲	۳	جمع
$f(x_2)$	$\frac{5}{36}$	$\frac{19}{36}$	$\frac{5}{15}$	۱

ب) برای بررسی مستقل بودن باید رابطه‌ی $f(x_1, x_2)=f(x_1).f(x_2)$ برای تمام $x_1=1,2,3$ و $x_2=1,2,3$ برقرار باشد. (اگر حتی برای یکی از اعضاء برقرار نباشد شرط مستقل بودن از بین می‌رود.)

$$f(1,1)=\frac{1}{12}\quad,\quad f(1)=\frac{3}{12}\quad,\quad f(1)=\frac{5}{36}$$

$$f(1,1)\neq f(1)\times f(1)\rightarrow \frac{1}{12}\neq\frac{5}{144}$$

چون برای اولین مقدار برقرار نشد پس مستقل نیستند و برای بقیه‌ی اعضاء ادامه نمی‌دهیم.

ج) $P[X_1\le X_2]=f(1,1)+f(1,2)+f(1,3)+f(2,2)+f(2,3)+f(3,3)$

$$=\frac{1}{12}+\frac{1}{6}+0+\frac{1}{9}+\frac{1}{5}+\frac{2}{15}=\frac{15+30+20+36+24}{180}=\frac{125}{180}=0/694$$

۱۹ـ تابع توزیع مشترک X و Y عبارت است از:

$$F(x,y)=P(X\le x, Y\le y)=\begin{cases}0 & x\le 0\ ,\ y\le 0\\ 1-2^{-x}-2^{-y}+2^{-xy} & x>0\ ,\ y>0\end{cases}$$

مطلوب است احتمال این که X یکی از مقادیر خود را در فاصله‌ی (۲ , ۱) و Y در فاصله‌ی (۵ , ۳) اختیار کند.

پاسخ: با توجه به تبصره‌ی ۳-۶-۳ داریم:

$$F(x,y)=P(X\le x, Y\le y)=\begin{cases}0 & x\le 0\ ,\ y\le 0\\ 1-2^{-x}-2^{-y}+2^{-xy} & x>0\ ,\ y>0\end{cases}$$

محاسبه‌ی احتمال (x , y) روی A به طوری که:

$$A=\{(x,y)|\ a\le x<b\ ,\ c\le y<d\}$$

تعریف شود از فرمول زیر قابل محاسبه است:

$$P[a\le X<b\ ,\ c\le Y<d]=F(b,d)-F(a,d)-F(b,c)+F(a,c)$$

حال برای این تمرین داریم:

$$P[1<X<2\ ,\ 3<Y<5]=F(2,5)-F(1,5)-F(2,3)+F(1,3)$$

$$=(1-2^{-2}-2^{-5}+2^{-10})-(1-2^{-1}-2^{-5}+2^{-5})-(1-2^{-2}-2^{-3}+2^{-6})+(1-2^{-1}-2^{-3}+2^{-3})$$

$$=1-2^{-2}-2^{-5}+2^{-10}-1+2^{-1}+2^{-5}-2^{-5}-1+2^{-2}+2^{-3}-2^{-6}+1-2^{-1}-2^{-3}-2^{-3}$$

$$=2^{-10}+2^{-3}-2^{-5}-2^{-6}=0/001+0/125-0/031-0/016=0/079$$

۲۰ـ اگر

$$f_X(x)=\frac{1}{10}\qquad x=0,1,2,\dots,9$$

$$f(y|x)=\frac{1}{10-x}\qquad y=x,x+1,\dots,9$$

مطلوب است تابع احتمال توأم $f(x,y)$ و تابع چگالی حاشیه‌ای $f_Y(y)$.

پاسخ: تابع احتمال توأم $f(x, y)$ را از رابطه‌ی زیر به دست می‌آوریم:

$$f(x,y)=f_X(x)\,f(y\mid x) \qquad x=0,1,\dots,9$$

$$f(x,y)=\frac{1}{10}\times\frac{1}{10-x}=\frac{1}{100-10x} \qquad y=x,x+1,\dots,9$$

برای به دست آوردن تابع چگالی حاشیه‌ای $f_Y(y)$ داریم:

$$f_Y(y)=\sum_{x=0}^{y}\frac{1}{100-10x}$$

$$f_Y(y)=P(Y=y)=\begin{cases}\frac{1}{100} & y=x=0\\ \frac{1}{100}+\frac{1}{90} & y=x+1=1\\ \vdots & \\ \frac{1}{100}+\frac{1}{90}+\dots+\frac{1}{10} & y=9\end{cases}$$

۲۱ـ اگر متغیر تصادفی X دارای تابع احتمال زیر باشد میانگین، واریانس و انحراف توزیع را به دست آورید.

x	۱۳۱	۱۴۰	۱۶۰	۱۸۰
f(x)	۰/۰۵	۰/۱۰	۰/۲۵	۰/۶۰

پاسخ:

میانگین $E(X)=\sum_x x.f(x)$

$$E(X)=(131\times 0/05)+(140\times 0/10)+(160\times 0/25)+(180\times 0/60)$$

$$=6/55+14+40+108=168/55$$

واریانس: $V(X)=E(X-\mu)^2=E(X^2)-\mu^2$

$$V(X)=28658/05-(168/55)^2=28658/05-28409/103=248/947$$

که در رابطه‌ی بالا $E(X^2)$ به صورت زیر به دست می‌آید:

$$E(X^2)=\sum_x x^2.f(x)$$

$$=(131^2\times 0/05)+(140^2\times 0/10)+(160^2+0/25)+(180^2\times 0/60)$$

$$=858/05+1960+6400+19440=28658/05$$

انحراف معیار: $\sqrt{V(X)}=\sqrt{248/947}=15/778$

۲۲ـ متغیر تصادفی گسسته‌ی X دو مقدار x_1 و x_2 را اختیار می‌کند ($x_2>x_1$). احتمال این‌که x_1 را اختیار کند برابر با ۰/۶ است. تابع چگالی X را مشخص کنید و میانگین و واریانس را به دست آورید.

پاسخ: تابع چگالی X به صورت زیر می‌باشد:

X	x_1	x_2	جمع
f(x)	۰/۶	۰/۴	۱

$$\sum_x f(x)=1\rightarrow f(x_1)+f(x_2)=1\rightarrow 0/6+f(x_2)=1\rightarrow f(x_2)=0/4$$

$$\mu=E(X)=\sum_x x.f(x)=(x_1\times 0/6)+(x_2\times 0/4)=0/6x_1+0/4x_2$$

$$E(X^2)=\sum_x x^2.f(x)=(x_1^2\times 0/6)+(x_2^2\times 0/4)=0/6x_1^2+0/4x_2^2$$

$$V(X)=E(X^2)-\mu^2$$

$$V(X)=(0/6x_1^2+0/4x_2^2)-[0/6x_1+0/4x_2]^2$$

با توجه به اتحاد مربع داریم: $(a+b)^2=a^2+b^2+2ab$

بنابراین:

$$V(X)=0/6x_1^2+0/4x_2^2-0/36x_1^2+0/16x_2^2-0/48x_1x_2$$

$$=0/24x_1^2+0/24x_2^2-0/48x_1x_2$$

$$=0/24[x_1^2+x_2^2-2x_1x_2]=0/24(x_1-x_2)^2$$

۲۳ـ اگر X دارای تابع چگالی احتمال زیر باشد مطلوب است $E(X)$ و $E(X^2+9)$

$$f(x)=\begin{cases}\frac{x}{2} & 0<x\leq 1\\ \frac{1}{2} & 1<x\leq 2\\ \frac{3-x}{2} & 2<x<3\end{cases}$$

پاسخ:

$$E(X)=\int_{-\infty}^{+\infty}x.f(x)dx=\int_0^1 x.\frac{x}{2}dx+\int_1^2 x.\frac{1}{2}dx+\int_2^3 x.\frac{3-x}{2}dx$$

$$=\frac{1}{2}\int_0^1 x^2dx+\frac{1}{2}\int_1^2 x\,dx+\frac{1}{2}\int_2^3(3x-x^2)dx$$

$$=\left[\frac{1}{6}x^3\right]_0^1+\left[\frac{1}{4}x^2\right]_1^2+\left[\frac{3x^2}{4}-\frac{x^3}{6}\right]_2^3$$

$$=\frac{1}{6}+\frac{3}{4}+\frac{7}{12}=\frac{2+9+7}{12}=\frac{18}{12}=1/5$$

اگر $h(X)=X^3+9$ در نظر بگیریم داریم:

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

$$E(X^3+9)=\int_0^1 (x^3+9).\frac{x}{2}\,dx+\int_1^2 (x^3+9).\frac{1}{2}\,dx+\int_2^3 (x^3+9)\left(\frac{3-x}{2}\right)dx$$

$$=\frac{1}{2}\int_0^1 (x^4+9x)dx+\frac{1}{2}\int_1^2 (x^3+9).\frac{1}{2}\,dx+\frac{1}{2}\int_2^3 (3x^3-x^4+27-9x)\,dx$$

$$=\frac{1}{2}\left[\frac{x^5}{5}+\frac{9x^2}{2}\right]_0^1+\frac{1}{2}\left[\frac{x^4}{4}+9x\right]_1^2+\frac{1}{2}\left[\frac{3x^4}{4}-\frac{x^5}{5}+27x-\frac{9x^2}{2}\right]_2^3$$

$$=\frac{47}{20}+\frac{51}{8}+\frac{221}{40}=\frac{94+255+221}{40}=\frac{570}{40}=14/25$$

۲۴ـ بازرگانی یک قلم جنس را به ارزش ۱۶۰۰ تومان می‌خرد و به قیمت ۲۰۰۰ تومان می‌فروشد. احتمال خرابی برای ۰ و ۱ و ۲ و ۳ و ۴ یا ۵ قلم بیشتر به ترتیب ۰/۰۵، ۰/۱۵، ۰/۳۰، ۰/۲۵، ۰/۱۵، ۰/۱۰ می‌باشد. امید سود حاصل از انبار کردن ۰، ۱، ۲، ۳، ۴ یا ۵ قلم یا بیش‌تر را محاسبه کنید.

پاسخ:

احتمال خرابی	۰	۱	۲	۳	۴	۵
f(x)	۰/۰۵	۰/۱۵	۰/۳۰	۰/۲۵	۰/۱۵	۰/۱۰

ابتدا باید سود حاصل را پیدا کنیم.

سود حاصل $=2000-1600=400$

$$E(x)=\sum_{x=0}^{5} x.f(x)=(400\times 0/05)+(400\times 0/15)+(400\times 0/30)+(400\times 0/25)$$
$$+(400\times 0/15)+(400\times 0/10)=400$$

۲۵ـ R شعاع اندازه‌گیری شده یک دایره دارای تابع چگالی احتمال زیر است. مطلوب است:

$f(r)=6r(1-r) \qquad 0<r<1$

الف) میانگین شعاع

ب) میانگین محیط

ج) میانگین سطح

پاسخ: الف) میانگین شعاع

$$E(R)=\int_{-\infty}^{+\infty} r.f(r)dr=\int_0^1 r.6r(1-r)dr$$

$$=6\int_0^1 (r^2-r^3)dr=6\left[\frac{r^3}{3}-\frac{r^4}{4}\right]_0^1=6\left[\frac{1}{3}-\frac{1}{4}\right]=6\times\frac{1}{12}=\frac{1}{2}=0/5$$

ب) میانگین محیط

محیط دایره = قطر × ۳/۱۴ $\qquad (\pi=3/14)$

محیط دایره $=2\pi r$

$$E(2\pi R)=\int_{-\infty}^{+\infty} 2\pi r.f(r)dr=\int_0^1 2\pi r.6r(1-r)dr$$

$$=12\pi\int_0^1 (r^2-r^3)dr=12\pi\left[\frac{r^3}{3}-\frac{r^4}{4}\right]_0^1=12\pi\left[\frac{1}{3}-\frac{1}{4}\right]=12\pi\times\frac{1}{12}=\pi=3/14$$

ج) میانگین سطح

مساحت دایره = (شعاع)^2 × ۳/۱۴

مساحت دایره $=\pi r^2$

$$E(\pi R^2)=\int_{-\infty}^{+\infty} \pi r^2.f(r)dr=\int_0^1 \pi r^2.6r(1-r)dr$$

$$=6\pi\int_0^1 (r^3-r^4)dr=6\pi\left[\frac{r^4}{4}-\frac{r^5}{5}\right]_0^1=6\pi\left[\frac{1}{4}-\frac{1}{5}\right]=\frac{6\pi}{20}=\frac{3}{10}\pi=0/942$$

۲۶ـ اگر X و Y دو متغیر تصادفی باشند:

الف) نشان دهید که $Cov(aX+c\,,\,bY+d)=ab\,Cov(X\,,Y)$

ب) اگر $E(X)=7$ و $E(Y)=1$ و $E(XY)=8$ باشد مقدار $Cov(2X-1\,,\,3Y+4)$ را محاسبه کنید.

پاسخ: الف) نکته: a و b و c و d مقادیر ثابت هستند.

می‌دانیم $Cov(X\,,Y)=E(XY)-E(X)\,E(Y)$

حال اگر رابطه‌ی خطی $aX+c$ را به عنوان X و $bY+d$ را به عنوان Y در نظر بگیریم بر اساس رابطه‌ی بالا داریم:

$$Cov(aX+c\,,\,bY+d)=E[(aX+c)(bY+d)]-E(aX+c)\,E(bY+d)$$
$$=E[abXY+adX+cbY+cd]-[aE(X)+c]\,[bE(Y)+d]$$

با توجه به خواص امید ریاضی داریم:

$$=abE(XY)+adE(X)+cbE(Y)+cd-abE(X)\,E(Y)-adE(X)-cbE(Y)-cd$$
$$=ab[E(XY)-E(X)\,E(Y)]=ab\,Cov(X\,,Y)$$

ب) با توجه به رابطه‌ی قسمت الف داریم:

$$Cov(2X-1\,,\,3Y+4)=2\times 3Cov(X\,,Y)=6\times 1=6$$

که در آن $Cov(X\,,Y)$ به صورت زیر به دست می‌آید:

$$Cov(X\,,Y)=E(XY)-E(X)E(Y)$$
$$Cov(X\,,Y)=8-(7\times 1)=8-7=1$$

۲۷ـ دو متغیر تصادفی X و Y دارای تابع احتمال توأم زیر است:

(x , y)	(۰ , ۰)	(۰ , ۱)	(۱ , ۰)	(۱ , ۱)	(۲ , ۰)	(۲ , ۱)
f(x , y)	$\frac{۱}{۱۸}$	$\frac{۳}{۱۸}$	$\frac{۴}{۱۸}$	$\frac{۳}{۱۸}$	$\frac{۶}{۱۸}$	$\frac{۱}{۱۸}$

مطلوب است $E(X+Y)$ و $E(X|Y=۰)$ و $Cov(X,Y)$ و ρ

پاسخ: برای سهولت در محاسبات می‌توان جدول را به صورت زیر رسم کرد:

X \ Y	۰	۱	f(x)
۰	$\frac{۱}{۱۸}$	$\frac{۳}{۱۸}$	$\frac{۴}{۱۸}$
۱	$\frac{۴}{۱۸}$	$\frac{۳}{۱۸}$	$\frac{۷}{۱۸}$
۲	$\frac{۶}{۱۸}$	$\frac{۱}{۱۸}$	$\frac{۷}{۱۸}$
f(y)	$\frac{۱۱}{۱۸}$	$\frac{۷}{۱۸}$	۱

$x = ۰ , ۱ , ۲$

$y = ۰ , ۱$

$$E(X+Y) = E(X) + E(Y)$$

$$E(X+Y)=\frac{۲۱}{۱۸}+\frac{۷}{۱۸}=\frac{۲۸}{۱۸}=۱/۵۵۶$$

که $E(X)$ و $E(Y)$ در زیر محاسبه شده است:

$$E(X)=\sum_{x=۰}^{۲} x.f(x)=(۰\times\frac{۴}{۱۸})+(۱\times\frac{۷}{۱۸})+(۲\times\frac{۷}{۱۸})=\frac{۷}{۱۸}+\frac{۱۴}{۱۸}=\frac{۲۱}{۱۸}=۱/۱۶۷$$

$$E(Y)=\sum_{y=۰}^{۱} y.f(y)=(۰\times\frac{۱۱}{۱۸})+(۱\times\frac{۷}{۱۸})=\frac{۷}{۱۸}=۰/۳۸۹$$

$$E(X|Y=۰)=\sum_{x=۰}^{۲} x.f(x|y=۰)$$

ابتدا باید تابع احتمال x به شرط $y=۰$ را به دست آوریم.

$$f(x|\ y=۰)=\frac{f(x,y)}{f(y=۰)} \quad x=۰,۱,۲$$

که تابع احتمال به صورت زیر است:

X	۰	۱	۲	جمع
f(x| y=۰)	$\frac{۱}{۱۱}$	$\frac{۴}{۱۱}$	$\frac{۶}{۱۱}$	۱

$$E(X|Y=۰)=(۰\times\frac{۱}{۱۱})+(۱\times\frac{۴}{۱۱})+(۲\times\frac{۶}{۱۱})=۰+\frac{۴}{۱۱}+\frac{۱۲}{۱۱}=\frac{۱۶}{۱۱}=۱/۴۵۵$$

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

$$Cov(X,Y)=\frac{۵}{۱۸}-(\frac{۲۱}{۱۸}\times\frac{۷}{۱۸})=\frac{۵}{۱۸}-\frac{۱۴۷}{۳۲۴}=\frac{-۵۷}{۳۲۴}=-۰/۱۷۶$$

مقدار $E(X)=\frac{۲۱}{۱۸}$ و $E(Y)=\frac{۷}{۱۸}$ را قبلاً محاسبه کردیم حال مقدار $E(XY)$ را به دست می‌آوریم:

$$E(XY)=\sum_{x=۰}^{۲}\sum_{y=۰}^{۱} xy.f(x,y)$$

$$=(۰\times۰\times\frac{۱}{۱۸})+(۰\times۱\times\frac{۳}{۱۸})+(۱\times۰\times\frac{۴}{۱۸})+(۱\times۱\times\frac{۳}{۱۸})$$

$$+(۲\times۰\times\frac{۶}{۱۸})+(۲\times۱\times\frac{۱}{۱۸})$$

$$=\frac{۳}{۱۸}+\frac{۲}{۱۸}=\frac{۵}{۱۸}=۰/۲۷۸$$

$$\rho=\frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}\Rightarrow\rho=\frac{-۰/۱۷۶}{\sqrt{۰/۵۸۲\times۰/۲۳۸}}=\frac{-۰/۱۷۶}{۰/۳۷۳}=-۰/۴۷۲$$

چون مقدار ρ منفی است پس همبستگی دو متغیر خلاف هم است یعنی افزایش یکی باعث کاهش دیگری می‌شود برای محاسبه‌ی ضریب همبستگی ابتدا باید $V(X)$ و $V(Y)$ را محاسبه کنیم.

$$V(X)=E(X^۲)-\mu_X^۲\rightarrow V(X)=۱/۹۴۴-(۱/۱۶۷)^۲=۰/۵۸۲$$

$$E(X^۲)=\sum_{x=۰}^{۲} x^۲ f(x)=(۰\times\frac{۴}{۱۸})+(۱\times\frac{۷}{۱۸})+(۴\times\frac{۷}{۱۸})$$

$$=۰+\frac{۷}{۱۸}+\frac{۲۸}{۱۸}=\frac{۳۵}{۱۸}=۱/۹۴۴$$

$$V(Y)=E(Y^۲)-\mu_Y^۲\rightarrow V(Y)=۰/۳۸۹-(۰/۳۸۹)^۲=۰/۲۳۸$$

$$E(Y^۲)=\sum_{y=۰}^{۱} y^۲.f(y)=(۰\times\frac{۱۱}{۱۸})+(۱\times\frac{۷}{۱۸})=\frac{۷}{۱۸}=۰/۳۸۹$$

۲۸ـ ثابت کنید گشتاورهای مرکزی دوم $\mu_۲$ از گشتاورهای مرتبه‌ی دوم نسبت به نقطه‌ی دلخواه c یعنی $\mu'_۲=E(X-c)^۲$ کوچک‌تر است.

پاسخ: می‌خواهیم نشان دهیم:

$$E(X-\mu)^۲<E(X-c)^۲$$

$$\sum_x (x-\mu)^۲ f(x)<\sum_x (x-c)^۲ f(x)$$

چون مقدار f(x) مقداری مثبت است بدون این که علامت نامساوی تغییر کند از دو طرف عبارت حذف می‌شود پس داریم:

$$\sum_x (x-\mu)^۲<\sum_x (x-c)^۲$$

با توجه به خواص میانگین رابطه‌ی به دست آمده در بالا را نامساوی میانگین گویند یعنی $\sum_x (x-c)^۲$ هنگامی کم‌ترین مقدار خود را اختیار می‌کند که c مساوی میانگین داده‌ها باشد.

بنابراین با ساده کردن نامساوی بالا داریم:

$$\sum_x (x^2-2x\mu+\mu^2) < \sum_x (x^2-2xc+c^2)$$

$$\rightarrow \sum_x x^2 - 2\mu\sum_x x + \mu^2\sum_x 1 < \sum_x x^2 + 2c\sum_x x + c^2\sum_x 1$$

با قرار دادن $n\mu=\sum_x x$ و $n=\sum_x 1$ در عبارت بالا داریم:

$$-2n\mu^2+n\mu^2 < -2cn\mu+nc^2 \rightarrow -n\mu^2 < n(-2c\mu+c^2)$$

تمام عبارت را بعد از ساده کردن به یک طرف منتقل می‌کنیم، داریم:

$$\mu^2-2c\mu+c^2>0 \rightarrow (\mu-c)^2>0$$

چون به یک عبارت بدیهی رسیدیم پس اثبات را یک اثبات بازگشتی گوییم یعنی:

$$E(X-\mu)^2<E(X-c)^2$$

۲۹ـ درستی تساوی زیر را نشان دهید:

$$\mu_3=E(X^3)-3E(X)E(X^2)+2(E(X))^3$$

پاسخ: می‌دانیم $\mu_r=E(X-\mu)^r$ می‌باشد، لذا:

$$\mu_3=E(X-\mu)^3=E\left[X^3-3X^2\mu+3X\mu^2-\mu^3\right]\rightarrow$$

با توجه به خواص میانگین (امید ریاضی) داریم:

$$\mu_3=E(X^3)-3\mu E(X^2)+3\mu^2E(X)-\mu^3\rightarrow$$

با توجه به این که $E(X)=\mu$ بنابراین می‌توان نوشت:

$$\mu_3=E(X^3)-3E(X)E(X^2)+3(E(X))^2E(X)-(E(X))^3\rightarrow$$

$$\mu_3=E(X^3)-3E(X)E(X^2)+3(E(X))^3-(E(X))^3\rightarrow$$

$$\mu_3=E(X^3)-3E(X)E(X^2)+2(E(X))^3$$

نکته: $(a-b)^3=a^3-3a^2b+3ab^2-b^3$

۳۰ـ اگر X و Y دو متغیر مستقل باشند و تعریف کنیم Z=X+Y نشان دهید که گشتاورهای مرکزی مرتبه‌ی سوم Z برابر با مجموع گشتاورهای مرکزی مرتبه‌ی سوم X و Y است. یعنی:

$$\mu_3^Z=\mu_3^X+\mu_3^Y$$

پاسخ: اگر X و Y دو متغیر مستقل باشند و $Z=X+Y$ با توجه به تبدیلی که روی داده‌ها انجام شده داریم: حال داریم: $\mu_Z=\mu_X+\mu_Y$

$$\mu_3^Z=E[Z-\mu_Z]^3=E[Z^3-3Z^2\mu_Z+3Z\mu_Z^2-\mu_Z^3]$$

$$=E[Z^3]-3\mu_Z E[Z^2]+3\mu_Z^2E[Z]-\mu_Z^3$$

حال با جایگزین کردن $Z=X+Y$ و $\mu_Z=\mu_X+\mu_Y$ داریم:

$$=E[(X+Y)^3]-3(\mu_X+\mu_Y)E[(X+Y)^2]+3(\mu_X+\mu_Y)^2E(X+Y)-(\mu_X+\mu_Y)^3$$

با ساده کردن رابطه‌های بالا داریم:

$$=E(X^3+3X^2Y+3XY^2+Y^3)-3(\mu_x+\mu_y)E[X^2+2XY+Y^2]$$

$$+3(\mu_x^2+2\mu_x\mu_y+\mu_y^2)E(X+Y)-[\mu_x^3+3\mu_x^2\mu_y+3\mu_x\mu_y^2+\mu_y^3]$$

حال با توجه به مستقل بودن X و Y (یعنی $E(XY)=E(X)(Y)$) و تأثیر امید ریاضی در جملات داریم:

$$=E(X^3)+3E(X^2)E(Y)+3E(X)E(Y^2)+E(Y^3)-3\mu_x E(X^2)$$

$$-6\mu_x E(X)E(Y)-3\mu_x E(Y^2)-3\mu_y E(X^2)-6\mu_y E(X)E(Y)-3\mu_y E(Y^2)$$

$$+3\mu_x^2 E(X)+3\mu_y^2E(X)+6\mu_x\mu_y E(X)+3\mu_x^2E(Y)+3\mu_y^2E(Y)$$

$$+6\mu_x\mu_y E(Y)-\mu_x^3-3\mu_x^2\mu_y-3\mu_x\mu_y^2-\mu_y^3$$

با توجه به این که $E(X)=\mu_x$ و $E(Y)=\mu_y$ رابطه‌های بالا را ساده می‌کنیم.

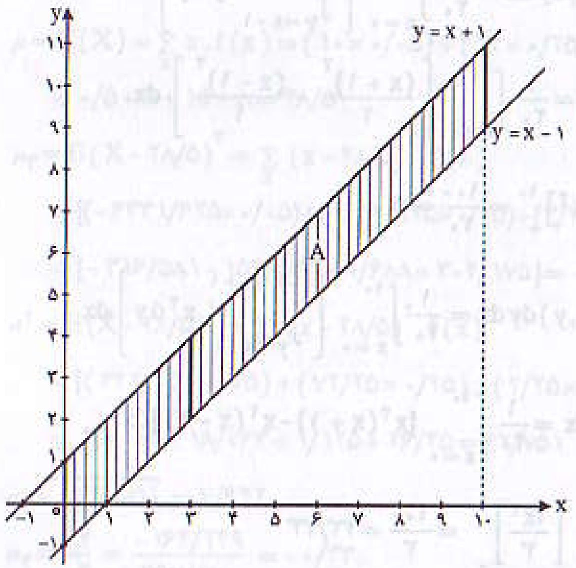
$$=E(X^3)-3\mu_x E(X^2)+3\mu_x^2E(X)-\mu_x^3+E(Y^3)-3\mu_y E(Y^2)+3\mu_y^2E(Y)-\mu_y^3$$

$$=E[X-\mu_x]^3+E[Y-\mu_y]^3=\mu_3^x+\mu_3^y$$

۳۱ـ فرض کنید X و Y دارای تابع چگالی احتمال توأم زیر باشند:

$$f(x,y)=\frac{1}{20}\qquad (x,y)\in A$$

که $A=\{(x,y)\mid 0<x<10,\ x-1<y<x+1\}$. $f(x,y)$ را رسم کنید و ضریب همبستگی X و Y را محاسبه کنید.



پاسخ: برای مشخص کردن $f(x,y)$ ابتدا ناحیه‌ی A را مشخص می‌کنیم. برای این منظور دو خط $y=x+1$ و $y=x-1$ را رسم می‌کنیم. $f(x,y)=\frac{1}{20}$، صفحه‌ای است که به ارتفاع $\frac{1}{20}$ روی ناحیه‌ی A را می‌پوشاند. می‌خواهیم ضریب همبستگی X و Y را محاسبه کنیم داریم:

$$\rho=\frac{Cov(X,Y)}{\sqrt{V(X).V(Y)}}$$

چون مقدار α_4 منفی است چولگی وجود دارد و چولگی داده‌ها به طرف چپ است.

$$\alpha_4 = \frac{\mu_4}{\sigma^4} - ۳ = \frac{E(X-\mu)^4}{\sigma^4} - ۳$$

$$\mu_4 = E(X - ۲۸/۵)^4 = \sum_x (x - ۲۸/۵)^4 . f(x)$$

$$= [(۱۱۷۱۳۵/۰۶۲ \times ۰/۰۵) + (۵۲۲۰/۰۶۲ \times ۰/۲۵) + (۵/۰۶۳ \times ۰/۵) + (۱۷۴۹۰/۰۶۲ \times ۰/۲۰)]$$

$$= ۵۸۵۶/۷۵۳ + ۱۳۰۵/۰۱۶ + ۲/۵۳۲ + ۳۴۹۸/۰۱۲ = ۱۰۶۶۲/۳۱۳$$

$$\alpha_4 = \frac{E(X-\mu)^4}{\sigma^4} - ۳ = \frac{۱۰۶۶۲/۳۱۳}{۳۹۳۸/۵۷۷} - ۳ = ۲/۷۰۷ - ۳ = -۰/۲۹۳$$

چون $\alpha_4 < ۰$ می‌باشد پس نمودار نسبت به توزیع نرمال دارای پخی است.

۳۳ـ تابع مولد گشتاورها را برای توابع چگالی احتمال زیر محاسبه کنید.

۱ـ $f(x) = p^x(۱-p)^{۱-x} \qquad x = ۰, ۱$

۲ـ $f(x) = p(۱-p)^x \qquad x = ۰, ۱, ۲, \ldots$

۳ـ $f(x) = \frac{۱}{\theta} e^{-\frac{x}{\theta}} \qquad x > ۰$

۴ـ $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-۱}(۱-x)^{\beta-۱} \qquad x > ۰$

پاسخ: ۱ـ

$$f(x) = p^x(۱-p)^{۱-x} \qquad x = ۰, ۱$$

x متغیر تصادفی برنولی با پارامتر p می‌باشد یعنی:

$$f(x) = P(X = x) = \begin{cases} ۱-p & , \quad x = ۰ \\ p & , \quad x = ۱ \\ ۰ & , \quad \text{سایر نقاط} \end{cases}$$

بنا به تعریف تابع مولد گشتاورها داریم:

$$M_X(t) = E(e^{tX}) = (۱-p)e^{t \times ۰} + pe^{t \times ۱} = (۱-p) + pe^t$$

۲ـ

$$f(x) = p(۱-p)^x \qquad x = ۰, ۱, ۲, \ldots$$

اگر $۱-p = q$ در نظر بگیریم داریم:

$$M_X(t) = E(e^{tX}) = \sum_{x=۰}^{\infty} e^{tx} pq^x = p\sum_{x=۰}^{\infty} (qe^t)^x$$

این سری زمانی همگراست اگر داشته باشیم $qe^t < ۱$ *

بنابراین تحت این شرایط داریم:

$$M_X(t) = p\frac{۱}{۱-qe^t} = p(۱-qe^t)^{-۱}$$

$$* \Leftrightarrow e^t < \frac{۱}{q} \Rightarrow t < -\ln q$$

نکته: سری هندسی با قدر نسبت r که $|r| < ۱$ همگراست و مقدارش برابر است با: $\frac{\text{اولین جمله}}{\text{۱ - قدر نسبت}}$

۳ـ

$$f(x) = \frac{۱}{\theta} e^{-\frac{x}{\theta}} \qquad x > ۰$$

بنا به تعریف تابع مولد گشتاورها داریم:

$$M_X(t) = E(e^{tX}) = \int_۰^{+\infty} e^{tx} f(x)dx = \int_۰^{+\infty} e^{tx} . \frac{۱}{\theta} e^{-\frac{x}{\theta}} dx$$

$$= \int_۰^{+\infty} \frac{۱}{\theta} e^{-x(\frac{۱}{\theta} - t)} dx = \frac{-۱}{\theta(\frac{۱}{\theta} - t)} e^{-x(\frac{۱}{\theta} - t)} \Big|_{x=۰}^{x=+\infty} = \frac{۱}{۱-\theta t}$$

۴ـ

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-۱}(۱-x)^{\beta-۱} \qquad x > ۰$$

بنا به تعریف تابع مولد گشتاورها داریم:

$$M_X(t) = E(e^{tX}) = \int_۰^{+\infty} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-۱}(۱-x)^{\beta-۱} . e^{tx} dx$$

اثبات به خواننده واگذار می‌شود.

۳۴ـ اگر متغیر تصادفی X دارای تابع مولد گشتاورهای زیر باشد میانگین و واریانس X را حساب کنید.

$$M_X(t) = e^{\mu t + \frac{۱}{۲}t^2\sigma^2}$$

پاسخ:

$$M'_X(t) = (\mu + \frac{۲t\sigma^2}{۲})e^{t\mu + \frac{t^2\sigma^2}{۲}}$$

$$M'_X(۰) = E(X) = \mu$$

$$M''_X(۰) = \sigma^2 e^{t\mu + \frac{t^2\sigma^2}{۲}} + (\mu + t\sigma^2)^2 e^{t\mu + \frac{t^2\sigma^2}{۲}}$$

$$M''_X(۰) = E(X^2) = \sigma^2 + \mu^2$$

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(X) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

بنابراین $E(X) = \mu$ و $V(X) = \sigma^2$ می‌باشد.

۳۵ـ اگر متغیر تصادفی X دارای تابع احتمال زیر باشد تابع مولد گشتاورهای X را به دست آورید.

x	$-\sqrt{3}$	0	$\sqrt{3}$
f(x)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

پاسخ:

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

$$=(e^{-\sqrt{3}t}\times\frac{1}{3})+(e^{0t}\times\frac{1}{3})+(e^{\sqrt{3}t}\times\frac{1}{3})=\frac{1}{3}(1+e^{-\sqrt{3}t}+e^{\sqrt{3}t})$$

۳۶ـ نشان دهید که تابع چگالی احتمال زیر دارای تابع مولد گشتاورهای $M_X(t)=\frac{1}{1-t^2}$ است.

$$f(x)=\frac{1}{2}e^{-|x|} \qquad -\infty<x<+\infty$$

پاسخ: بنا به تعریف داریم:

$$f(x)=\begin{cases}\frac{1}{2}e^{-x} & x\geq 0\\ \frac{1}{2}e^{x} & x<0\end{cases}$$

$$M_X(t)=E(e^{tX})=\int_{-\infty}^{+\infty}\frac{1}{2}e^{-|x|}.e^{tx}dx$$

$$=\int_{-\infty}^{0}\frac{1}{2}e^{x}.e^{tx}dx+\int_{0}^{+\infty}\frac{1}{2}e^{-x}.e^{tx}dx$$

$$=\frac{1}{2}\int_{-\infty}^{0}e^{(1-t)x}dx+\frac{1}{2}\int_{0}^{+\infty}e^{(t-1)x}dx$$

برای این که انتگرال‌ها وجود داشته باشند لازم است تا حوزه‌ی تعریف تابع مولد گشتاورهای متغیر تصادفی X را به $\{t\in R: -1<t<1\}$ محدود کنیم. برای مقادیر t در این حوزه‌ی تعریف،

$$M_X(t)=E(e^{tX})=\frac{1}{2(1+t)}e^{(1+t)x}\Big|_{-\infty}^{0}+\frac{1}{2(t-1)}e^{(t-1)x}\Big|_{0}^{+\infty}$$

$$=\frac{1}{2(1+t)}+\frac{1}{2(1-t)}=\frac{1}{1-t^2}$$

بنابراین $M_X(t)=\frac{1}{1-t^2}$ می‌باشد.

۳۷ـ اگر X دارای تابع چگالی احتمال $f(x)=1$ و $0<x<1$ احتمال زیر را به طریق مستقیم و با استفاده از نامساوی چبیشف حساب کنید و آن‌ها را مقایسه کنید.

$$P[|X-\mu|>4]$$

که $\mu=E(X)$ است.

پاسخ: با استفاده از نامساوی چبیشف داریم:

$$E(X)=\int_{-\infty}^{+\infty}x.f(x)dx=\int_0^1 x.1dx=\frac{1}{2}x^2\Big|_0^1=\frac{1}{2}=0/5$$

$$E(X^2)=\int_0^1 x^2dx=\frac{1}{3}x^3\Big|_0^1=\frac{1}{3}$$

$$V(X)=E(X^2)-[E(X)]^2=\frac{1}{3}-(\frac{1}{2})^2=\frac{1}{3}-\frac{1}{4}=\frac{1}{12}$$

بنابراین با استفاده از نامساوی چبیشف داریم:

$$P(|X-\mu|>4)\leq\frac{V(X)}{16}=\frac{\frac{1}{12}}{16}=\frac{1}{192}$$

همچنین با استفاده از روش مستقیم داریم:

$$P(|X-\mu|>4)=1-P(|X-0/5|\leq 4)=1-P(-4<X-0/5<4)$$

$$=1-P(-3/5<X<4/5)=1-1=0$$

۳۸ـ در نامساوی چبیشف، اگر بخواهیم وقتی متغیر تصادفی مقداری بین $\mu-k\sigma$ و $\mu+k\sigma$ اختیار می‌کند با احتمال حداقل ۰/۹۵ همراه باشد، کم‌ترین مقدار k چه خواهد بود؟ (μ و σ به ترتیب میانگین و انحراف معیار متغیر است.)

پاسخ: با توجه به رابطه‌ی نامساوی چبیشف داریم:

$$P[|X-\mu|\geq k]\leq\frac{\sigma^2}{k^2}$$

که این نامساوی را به صورت زیر نیز می‌توان بیان کرد:

$P[|X-\mu|<k\sigma]\geq 1-\frac{1}{k^2}$ به ازای مقدار مثبت k

بنابراین داریم:

$$P(\mu-k\sigma<X<\mu+k\sigma)\geq 0/95$$

$$\rightarrow P(-k\sigma<X-\mu<k\sigma)\geq 0/95$$

$$\rightarrow P(|X-\mu|<k\sigma)\geq 0/95$$

احتمال خواسته شده به صورت نامساوی چبیشف در آمده پس داریم:

$$1-\frac{1}{k^2}=0/95\rightarrow\frac{1}{k^2}=0/05\rightarrow 0/05k^2=1\rightarrow k^2=\frac{1}{0/05}=20$$

پس $k=\pm\sqrt{20}$ با توجه به این که k مقداری مثبت است پس داریم: $k=\sqrt{20}$

۳۹ـ متغیر تصادفی X دارای میانگین $\mu=8$ و واریانس $\sigma^2=9$ می‌باشد. مطلوب است:

$$P[-4<X<20]$$

پاسخ:

$$P[-4<X<20]=P[-12<X-8<12]=P[-12<X-\mu<12]$$

$$=P[|X-\mu|<12]=1-P[|X-\mu|\geq 12]$$

اما

$$=P[|X-\mu|\geq 12]\leq \frac{\sigma^2}{k^2}=\frac{9}{144}=0/063$$

بنابراین:

$$P[-4<X<20]\geq 1-\frac{9}{144}$$

$$P[-4<X<20]\geq \frac{135}{144}=0/937$$

۴۰ـ اگر X دارای تابع چگالی احتمالی زیر باشد مطلوب است $P[0<X<2]$

x	0	1	3
f(x)	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

پاسخ:

$$E(X)=\sum_x x.f(x)=(0\times\frac{1}{3})+(1\times\frac{1}{2})+(3\times\frac{1}{6})=0+\frac{1}{2}+\frac{3}{6}=\frac{6}{6}=1$$

$$E(X^2)=\sum_x x^2.f(x)=(0\times\frac{1}{3})+(1\times\frac{1}{2})+(9\times\frac{1}{6})=0+\frac{1}{2}+\frac{9}{6}=\frac{12}{6}=2$$

$$V(X)=E(X^2)-[E(X)]^2=2-1=1$$

با توجه به این که $\mu=E(X)=1$ و $\sigma^2=V(X)=1$ حال احتمال موردنظر را محاسبه می‌کنیم:

$$P[0<X<2]=P[0-1<X-1<2-1]$$

$$=P[-1<X-1<1]=P[|X-\mu|<1]=1-P[|X-\mu|\geq 1]$$

اما داریم:

$$P[|X-\mu|\geq 1]\leq \frac{\sigma^2}{k^2}=1$$

بنابراین:

$$P[0<X<2]\geq 1-1=0$$

پس:

$$P[0<X<2]\geq 0$$

خودآزمایی فصل چهارم

۱ـ توزیع یکنواخت را برای نمونه‌ی تصادفی سه تایی که از ۵ رایانه‌ی مختلف می‌تواند تشکیل شود به دست آورید.

پاسخ: تعداد حالت‌هایی که می‌توان از بین ۵ رایانه‌ی مختلف ۳ تا را انتخاب کرد برابر است با:

$$\binom{۵}{۳} = \frac{۵!}{۳!\,۲!} = ۱۰$$

حال با توجه به تابع احتمال یکنواخت، احتمال انتخاب هر کدام از حالت‌ها برابر است با:

$$f(x) = P(X = x) = \frac{۱}{k} \qquad x = ۱, ۲, \ldots, k$$

بنابراین:

$$f(x) = P(X = x) = \frac{۱}{\binom{۵}{۳}} = \frac{۱}{۱۰} \qquad x = ۱, ۲, \ldots, ۱۰$$

۲ـ احتمال این که یک تیرانداز هدفی را بزند برابر با ۰/۷۵ است. یک قالب احتمال برای پیروزی و عدم پیروزی تیرانداز ارائه دهید و شانس عدم موفقیت ایشان را به دست آورید.

پاسخ: متغیر تصادفی X دارای تابع احتمال برنولی با پارامتر $p = ۰/۷۵$ و $q = ۱ - p = ۱ - ۰/۷۵ = ۰/۲۵$ می‌باشد. با توجه به تابع احتمال برنولی داریم:

$$f(x) = P(X = x) = p^x(۱ - p)^{۱-x} \qquad x = ۰, ۱$$

پس:

$$f(x) = P(X = x) = (۰/۷۵)^x(۰/۲۵)^{۱-x} \qquad x = ۰, ۱$$

و احتمال (شانس) عدم موفقیت به صورت زیر است:

$$P(x = ۰) = (۰/۷۵)^۰(۰/۲۵)^۱ = ۰/۲۵$$

۳ـ احتمال این که در یک خط تولید کالایی معیوب باشد برابر با ۰/۰۰۱ است. در یک نمونه‌ی ۷ تایی مطلوب است احتمال این که:

الف) کالای معیوب پیدا نشود.

ب) بیش از ۵ کالا معیوب باشد.

ج) دقیقاً ۳ کالا معیوب باشد.

پاسخ: اگر احتمال معیوب بودن $p=0/001$ و $n=7$ باشد متغیر تصادفی X دارای توزیع دو جمله‌ای به صورت زیر است:

$$f(x)=P(X=x)=\binom{7}{x}(0/001)^x(0/999)^{7-x} \qquad x=0,1,\dots,7$$

الف) $P(X=0)=\binom{7}{0}(0/001)^0(0/999)^7=1\times 1\times 0/993=0/993$

ب) $P(X>5)=1-P(X\le 5)=1-F(5)$

$$=1-\sum_{t=0}^{5}\binom{7}{t}(0/001)^t(0/999)^{7-t}$$

ج) $P(X=3)=\binom{7}{3}(0/001)^3(0/999)^4=3/5\times 10^{-8}$

۴ـ یک محموله از ۵۰ قطعه‌ی مکانیکی شامل ۴۲ قطعه‌ی سالم و ۸ قطعه‌ی معیوب است. یک بازرس ۵ قطعه را بدون جایگذاری انتخاب می‌کند.

الف) یک قالب احتمال ارائه دهید.

ب) احتمال این که درست سه قطعه سالم باشد چقدر است؟

ج) احتمال این که حداکثر سه قطعه سالم باشد چقدر است؟

پاسخ: الف) متغیر تصادفی X دارای تابع احتمال فوق هندسی با پارامترهای $N=50$ و $k=42$ و $n=5$ است و تابع احتمال آن به صورت زیر می‌باشد:

$$f(x)=P(X=x)=\frac{\binom{7}{x}\binom{N-k}{n-x}}{\binom{N}{n}}=\frac{\binom{42}{x}\binom{8}{5-x}}{\binom{50}{5}} \qquad x=0,1,\dots,5$$

ب) $P(X=3)=\dfrac{\binom{42}{3}\binom{8}{2}}{\binom{50}{5}}=\dfrac{11480\times 28}{2118760}=0/152$

ج) $P(X\le 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)$

$$=\frac{\binom{42}{0}\binom{8}{5}}{\binom{50}{5}}+\frac{\binom{42}{1}\binom{8}{4}}{\binom{50}{5}}+\frac{\binom{42}{2}\binom{8}{3}}{\binom{50}{5}}+\frac{\binom{42}{3}\binom{8}{2}}{\binom{50}{5}}$$

$$=0/00003+0/00139+0/02276+0/15171=0/17589$$

۵ـ احتمال این که یک وسیله‌ی اندازه‌گیری یک جابجایی مفرط را نشان دهد برابر با ۰/۰۵ است احتمال این که ششمین وسیله‌ای که آزمون شده است اولین وسیله‌ای باشد که جابجایی را نشان می‌دهد چقدر است؟

پاسخ: متغیر تصادفی X دارای تابع احتمال هندسی با پارامتر $p=0/05$ است و تابع احتمال آن به صورت زیر می‌باشد:

$$f(x)=P(X=x)=(0/05)(0/95)^x \qquad x=0,1,2,\dots$$

حال داریم:

$$P(X=6)=(0/05)(0/95)^6=0/037$$

۶ـ برای تابع چگالی احتمال هندسی امید ریاضی و واریانس را به دست آورید.

$$f(x)=P(X=x)=P(1-P)^x \qquad x=0,1,2,\dots$$

پاسخ: برای به دست آوردن میانگین و واریانس می‌توانیم از تابع مولد گشتاور مشتق گرفته و سپس t را صفر قرار دهیم همچنین می‌توان به طور مستقیم میانگین و واریانس متغیر تصادفی هندسی را به دست آورد که در این جا به طور مستقیم به دست می‌آوریم:

$$E(X)=\sum_{x=0}^{\infty}x\,pq^x=pq\sum_{x=0}^{\infty}x\,q^{x-1}=pq\sum_{x=0}^{\infty}\frac{d}{dq}(q^x)=pq\frac{d}{dq}\sum_{x=0}^{\infty}q^x$$

همچنین می‌دانیم سری هندسی با قدر نسبت r که $|r|<1$ همگراست و مقدارش برابر با $\dfrac{\text{اولین جمله}}{\text{قدر نسبت}-1}$

بنابراین برای سری بالا داریم:

$$\sum_{x=0}^{\infty}q^x=\frac{1}{1-q}$$

$$E(X)=pq\frac{d}{dq}\sum_{x=0}^{\infty}q^x=pq\frac{d}{dq}\left(\frac{1}{1-q}\right)=pq\frac{1}{(1-q)^2}=\frac{q}{p}$$

پس $E(X)=\dfrac{q}{p}$ که در آن $q=1-p$ می‌باشد.

برای محاسبه‌ی واریانس ابتدا $E(X(X-1))$ را محاسبه می‌کنیم تا مقدار $E(X^2)$ را به دست آوریم:

$$E(X(X-1))=\sum_{x=0}^{\infty}x(x-1)pq^x=pq^2\sum_{x=0}^{\infty}x(x-1)q^{x-2}$$

$$=pq^2\sum_{x=0}^{\infty}\frac{d^2}{dq^2}(q^x)=pq^2\frac{d^2}{dq^2}\left(\sum_{x=0}^{\infty}q^x\right)=pq^2\frac{d^2}{dq^2}\left(\frac{1}{1-q}\right)$$

$$=pq^2\frac{d}{dq}\left(\frac{1}{(1-q)^2}\right)=pq^2\left(\frac{2(1-q)}{(1-q)^4}\right)=\frac{2q^2}{(1-q)^2}=\frac{2q^2}{p^2}$$

بنابراین:

$$E(X(X-1)) = E(X^2) - E(X) = \frac{2q^2}{p^2}$$

از رابطه‌ی بالا داریم:

$$E(X^2) = E(X(X-1)) + E(X)$$

$$E(X^2) = \frac{2q^2}{p^2} + \frac{q}{p} = \frac{2q^2+pq}{p^2}$$

حال برای محاسبه‌ی واریانس داریم:

$$Var(X) = E(X^2) - (E(X))^2$$

$$Var(X) = \frac{2q^2+pq}{p^2} - \frac{q^2}{p^2} = \frac{q^2+pq}{p^2} = \frac{q(q+p)}{p^2} = \frac{q(1)}{p^2} = \frac{q}{p^2}$$

۷ـ در مسئله‌ی ۶ نشان دهید که:

$$P[X \ge s+t \mid X \ge t] = P(X \ge s)$$

پاسخ: متغیر تصادفی X با تابع احتمال هندسی با پارامتر θ عبارت است از:

$$f(X) = \theta(1-\theta)^{x-1} \quad ; \quad 0<\theta<1 \quad , \quad x=1,2,\dots$$

از طرفی توزیع هندسی دارای خاصیت بی‌حافظگی می‌باشد بنابراین:

$$P[X \ge s+t \mid X \ge t] = P(X \ge s)$$

ابتدا $P(X \ge t)$ را به دست می‌آوریم، داریم:

$$P(X \ge t) = P(X=t+1) + P(X=t+2) + \dots$$

$$P(X \ge t) = \theta(1-\theta)^t + \theta(1-\theta)^{t+1} + \dots$$

$$P(X \ge t) = \theta(1-\theta)^t + \underbrace{[1+(1-\theta)+(1-\theta)^2+\dots]}_{*}$$

* با استفاده از تصاعد هندسی برابر با $\frac{1}{1-(1-\theta)}$ می‌باشد.

$$P(X \ge t) = \theta(1-\theta)^t \times \frac{1}{1-(1-\theta)} = \theta(1-\theta)^t \times \frac{1}{\theta} = (1-\theta)^t$$

حال با استفاده از احتمال شرطی داریم:

$$P(X \ge s+t \mid X \ge t] = \frac{P(X \ge s+t \,,\, X \ge t)}{P(X \ge t)}$$

$$\Rightarrow P[X \ge s+t \mid X \ge t] = \frac{(1-\theta)^{s+t-1}}{(1-\theta)^t}$$

$$\Rightarrow P[X \ge s+t \mid X \ge t] = (1-\theta)^{s-1}$$

بنابراین:

$$\Rightarrow P[X \ge s+t \mid X \ge t] = P(X \ge s)$$

رابطه ثابت شد.

۸ـ در آبگیری ۲۰۰ ماهی قزل‌آلا وجود دارد. از این آبگیر ۵۰ ماهی را به تصادف صید کرده و پس از علامت‌گذاری آن‌ها را به آبگیر بر می‌گردانیم. اگر مجدداً ۵۰ ماهی را به تصادف صید کنیم احتمال این که ۵ ماهی از آن‌ها علامت‌گذاری شده باشند چقدر است؟ و احتمال این که حداکثر ۳ ماهی از ماهی‌ها علامت‌گذاری شده باشد چقدر است؟

تذکر: این تمرین را از توزیع فوق هندسی با $N=200$، $k=50$ و $n=50$ نیز می‌توان حل کرد.

پاسخ: متغیر تصادفی X دارای تابع احتمال دو جمله‌ای با پارامتر $p = \frac{50}{200} = \frac{1}{4}$ و $n=200$ می‌باشد چون n و p به ترتیب از حد تصور بزرگ‌تر و کوچک‌تر است به طوری که $np=\lambda$ می‌باشد می‌توان توزیع دو جمله‌ای را با استفاده از توزیع پواسن تقریب زد. پس داریم:

$$\lambda = np = 200 \times \frac{1}{4} = 50$$

پس:

$$f(x) = P(X=x) = \frac{e^{-50} 50^x}{x!} \qquad x = 0, 1, 2, \dots$$

$$P(X=5) = \frac{e^{-50} 50^5}{5!}$$

$$P(X \le 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= \frac{e^{-50} 50^0}{0!} + \frac{e^{-50} 50^1}{1!} + \frac{e^{-50} 50^2}{2!} + \frac{e^{-50} 50^3}{3!}$$

$$= e^{-50} + 50 e^{-50} + 1250 e^{-50} + 1041/67\, e^{-50}$$

۹ـ یک چاپگر رایانه طوری طراحی شده که در هر ۱۵ ثانیه، ۲ صفحه از اطلاعات ذخیره شده‌ی خود را چاپ می‌کند. اگر این چاپگر به مدت ۳ دقیقه کار کند مطلوب است احتمال این که:

الف) صفحه‌ای چاپ نکند.

ب) حداقل چهار صفحه چاپ کند.

پاسخ: متغیر تصادفی X دارای تابع احتمال پواسن می‌باشد ابتدا λ را برای ۳ دقیقه (۱۸۰ ثانیه) به دست می‌آوریم:

با یک تناسب ساده داریم:

صفحه	ثانیه
۲	۱۵
$\lambda=24$	۱۸۰

پس تابع احتمال با توجه به $\lambda=24$ به صورت زیر است:

$$f(x) = P(X=x) = \frac{e^{-24} 24^x}{x!} \qquad x = 0, 1, 2, \dots$$

الف) $P(X=0)=\frac{e^{-24}24^0}{0!}=e^{-24}$

ب) $P(X\ge 4)=1-P(X<4)=1-\left[P(X=0)+P(X=1)+P(X=2)+P(X=3)\right]$

$$=1-\left[\frac{e^{-24}24^0}{0!}+\frac{e^{-24}24^1}{1!}+\frac{e^{-24}24^2}{2!}+\frac{e^{-24}24^3}{3!}\right]$$

$$=1-\left[e^{-24}+24e^{-24}+288e^{-24}+2304e^{-24}\right]=1-2617e^{-24}$$

۱۰ـ تعداد تصادفاتی که در فواصل مساوی جاده اتفاق می‌افتد دارای توزیع پواسن با پارامتر $\lambda=0.3$ است. مطلوب است احتمال این که:

الف) تصادفی صورت نگیرد.

ب) حداقل یک تصادف اتفاق بیفتد.

ج) بین ۱ تا ۲ تصادف اتفاق بیفتد.

پاسخ:
$$f(x)=P(X=x)=\frac{e^{-0.3}0.3^x}{x!}\qquad x=0,1,2,\dots$$

الف) $P(X=0)=\frac{e^{-0.3}0.3^0}{0!}=e^{-0.3}=0.74$

ب) $P(X\ge 1)=1-P(X<1)=1-P(X=0)=1-0.74=0.26$

ج) $P(1\le X\le 2)=P(X=1)+P(X=2)$

$$=\frac{e^{-0.3}0.3^1}{1!}+\frac{e^{-0.3}0.3^2}{2!}=0.3e^{-0.3}+0.045e^{-0.3}=0.345e^{-0.3}=0.255$$

۱۱ـ در توزیع پواسن، تابع مولد گشتاورها را به دست آورید و از روی آن میانگین و واریانس توزیع را حساب کنید. برای محاسبه‌ی $E(X(X-1))$ چه راهی پیشنهاد می‌کنید.

پاسخ: ابتدا تابع مولد گشتاورهای توزیع پواسن را به دست می‌آوریم. داریم:

$$M_X(t)=E(e^{tX})=\sum_{x=0}^{\infty}e^{tx}f(x)=\sum_{x=0}^{\infty}e^{tx}\frac{e^{-\lambda}\lambda^x}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!}$$

همچنین می‌دانیم: $e^x=\sum_{x=0}^{\infty}\frac{x^k}{k!}$

بنابراین:

$$=e^{-\lambda}.e^{\lambda e^t}=e^{\lambda(e^t-1)}$$

برای محاسبه‌ی میانگین باید از تابع مولد گشتاور مشتق اول گرفته و $t=0$ قرار دهیم:

$$E(X)=\frac{d}{dt}M_X(t)\Big|_{t=0}=\frac{d}{dt}e^{\lambda(e^t-1)}\Big|_{t=0}=\lambda e^t e^{\lambda(e^t-1)}\Big|_{t=0}=\lambda$$

برای محاسبه‌ی واریانس ابتدا $E(X^2)$ را محاسبه کرده پس داریم:

$$E(X^2)=\frac{d^2}{dt^2}M_X(t)\Big|_{t=0}=\lambda e^t e^{\lambda(e^t-1)}+\lambda^2e^{2t}e^{\lambda(e^t-1)}\Big|_{t=0}=\lambda+\lambda^2$$

بنابراین:

$$Var(X)=E(X^2)-(E(X))^2=\lambda+\lambda^2-\lambda^2=\lambda$$

برای محاسبه‌ی $E(X(X-1))$ داریم:

$$E(X(X-1))=E(X^2-X)=E(X^2)-E(X)$$

با توجه به این که $E(X)=\lambda$ و $E(X^2)=\lambda+\lambda^2$ پس داریم:

$$E(X(X-1))=\lambda+\lambda^2-\lambda=\lambda^2$$

۱۲ـ اگر X دارای تابع احتمال سری لگاریتمی با پارامتر $\alpha=0.6379$ باشد جدول زیر را کامل کنید.

x	۱	۲	۳	۴	۵	۶
f(x)	۰/۶۳۷۹	۰/۲۰۰۳	۰/۰۸۵۲	۰/۰۴۰۷	۰/۰۲۰۸	۰/۰۱۱۱

پاسخ: اگر متغیر تصادفی X دارای تابع احتمال سری لگاریتمی با پارامتر $\alpha=0.6379$ باشد تابع احتمال به صورت زیر است:

$$f(x)=P(X=x)=\frac{-1}{\ln(1-0.6379)}\cdot\frac{(0.6379)^x}{x}\qquad x=1,2,\dots$$

$$f(x)=P(X=x)=0.9844\,\frac{(0.6379)^x}{x}\qquad x=1,2,\dots$$

مثلاً برای $x=1$ داریم:

$$f(1)=P(X=1)=0.9844\times(0.6379)=0.6279$$

۱۳ـ تجربه نشان داده که چاپگرهای شرکتی در هر ۱۰۰۰ صفحه یک اشتباه چاپ می‌کند. اگر ۱۰۰۰۰ صفحه به تصادف انتخاب شود احتمال این که ۶، ۷ یا ۸ صفحه اشتباه چاپ شود چقدر است؟

پاسخ: متغیر تصادفی X دارای توزیع دو جمله‌ای با پارامتر $n=10000$ و $p=\frac{1}{1000}$ است چون n بزرگ و p کوچک است با استفاده از تقریب پواسن داریم:

$$\lambda=np=10000\times\frac{1}{1000}=10$$

حال با $\lambda=10$ داریم:

$$f(x)=P(X=x)=\frac{e^{-10}10^x}{x!}\qquad x=0,1,2,\dots$$

$$P(X=6)+P(X=7)+P(X=8)=\frac{e^{-10}10^6}{6!}+\frac{e^{-10}10^7}{7!}+\frac{e^{-10}10^8}{8!}$$

همچنین با توجه به محاسبه‌ی احتمالات از روی جدول توزیع پواسن داریم:

$$P(X=6)+P(X=7)+P(X=8)=P(X\le 8)-P(X\le 5)=0.333-0.067=0.266$$

۱۴ـ سوابق در یک سایت رایانه نشان می‌دهد که احتمال این که رایانه‌ای از رده خارج شود برابر ۰/۰۰۱۲ است. احتمال این که از بین ۱۰۰۰ رایانه حداقل ۲ رایانه از رده خارج شود چقدر است؟

پاسخ: متغیر تصادفی X دارای توزیع دوجمله‌ای با پارامتر $n=1000$ و $p=0.0012$ می‌باشد چون n از حد تصور بزرگ‌تر و p از حد تصور کوچک‌تر است با استفاده از تقریب پواسن داریم:

$$\lambda = np = 1000 \times 0.0012 = 1.2$$

$$f(x) = P(X=x) = \frac{e^{-1.2}1.2^x}{x!} \qquad x=0,1,2,\ldots$$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - [P(X=0) + P(X=1)]$$

$$= 1 - \left[\frac{e^{-1.2}1.2^0}{0!} + \frac{e^{-1.2}1.2^1}{1!}\right] = 1 - [e^{-1.2} + 1.2e^{-1.2}]$$

$$= 1 - 2.2e^{-1.2} = 1 - 0.6626 = 0.3374$$

۱۵ـ سختی یک آلیاژ معین (اندازه‌گیری شده با مقیاس شکنندگی) یک متغیر تصادفی یکنواخت با $a=50$ و $b=75$ است. تابع توزیع آن را به دست آورید و احتمال $P[60<X<70]$ را پیدا کنید.

پاسخ: متغیر تصادفی X دارای تابع احتمال یکنواخت با پارامترهای $a=50$ و $b=75$ است و تابع چگالی آن به صورت زیر می‌باشد:

$$f(x) = \frac{1}{75-50} = \frac{1}{25} \qquad 50 < x < 75$$

و تابع توزیع برای تابع چگالی احتمال یکنواخت به صورت زیر است:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

که برای این تمرین داریم:

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < 50 \\ \frac{x-50}{25} & 50 \leq x < 75 \\ 1 & x \geq 75 \end{cases}$$

$$P(60 < X < 70) = \int_{60}^{70} f(x)\,dx = \int_{60}^{70} \frac{1}{25}dx = \frac{x}{25}\Big|_{60}^{70} = \frac{70}{25} - \frac{60}{25} = \frac{10}{25} = 0.4$$

۱۶ـ در آزمایش‌های معینی، خطاهای حاصل از تعیین چگالی یک ماده، متغیر تصادفی است با تابع چگالی یکنواخت که در آن $a=-0.025$ و $b=0.025$ می‌باشد مطلوب است احتمال خطا:

الف) مابین ۰/۰۱ و ۰/۱۵

ب) مابین ۰/۰۱۲- و ۰/۰۱۲

پاسخ: متغیر تصادفی X دارای تابع چگالی یکنواخت می‌باشد.

$$f(x) = P(X=x) = \frac{1}{0.025-(-0.025)} = \frac{1}{0.05} \qquad -0.025 < x < 0.025$$

پس داریم:

$$f(x) = P(X=x) = 20 \qquad -0.025 < x < 0.025$$

الف)

$$P(0.01 < X < 0.15) = \int_{0.01}^{0.15} f(x)\,dx = \int_{0.01}^{0.025} 20\,dx + \int_{0.025}^{0.15} 0\,dx$$

$$= 20x\Big|_{0.01}^{0.025} = 0.5 - 0.2 = 0.3$$

ب)

$$P(-0.012 < X < 0.012) = \int_{-0.012}^{0.012} f(x)\,dx = \int_{-0.012}^{0.012} 20\,dx$$

$$= 20x\Big|_{-0.012}^{0.012} = 0.24 - (-0.24) = 0.48$$

۱۷ـ اگر متغیر تصادفی Z دارای توزیع نرمال استاندارد باشد با استفاده از جدول ضمیمه‌ی (۳) احتمالات زیر را حساب کنید.

$P[Z<1.65]$, $P[|Z|<1.96]$, $P[Z<0.2]$

$P[|Z|<1.64]$, $P[-0.5<Z<3.2]$, $P[Z^3 \leq 8]$

پاسخ:

* $P[Z<1.65] = 0.9505$

* $P[|Z|<1.96] = P[-1.96<Z<1.96] = P(Z<1.96) - P(Z<-1.96)$

$= P(Z<1.96) - P(Z>1.96) = P(Z<1.96) - [1 - P(Z<1.96)]$

$= 2P(Z<1.96) - 1 = 2(0.9750) - 1 = 0.95$

* $P[Z<0.2] = 0.5793$

* $P[|Z|<1.64] = P[-1.64<Z<1.64] = P(Z<1.64) - P(Z<-1.64)$

$= P(Z<1.64) - P(Z>1.64) = P(Z<1.64) - [1 - P(Z<1.64)]$

$= 2P(Z<1.64) - 1 = 2(0.9495) - 1 = 0.899$

* $P[-0.5<Z<3.2] = P(Z<3.2) - P(Z<-0.5) = P(Z<3.2) - P(Z>0.5)$

$= P(Z<3.2) - [1 - P(Z<0.5)] = P(Z<3.2) + P(Z<0.5) - 1$

$= 1 + 0.6915 - 1 = 0.6915$

* $P[Z^3 \leq 8] = P[Z \leq 2] = 0.9772$

۱۸ ـ اگر متغیر تصادفی X دارای توزیع نرمال با میانگین $\mu=۳$ و واریانس $\sigma^{۲}=۰/۱۶$ باشد احتمال‌های زیر را حساب کنید.

$P[X>۳]$ $\quad$ $P[X>۳/۳]$ $\quad$ $P[|X|<۳/۳]$

$P[|X-۳|<۱/۵]$ $\quad$ $P[X<۴/۹۶]$ $\quad$ $P[۲/۸<X<۳/۱]$

و مقدار c را به گونه‌ای بیابید که $P[۳-c<X<۳+c]=۰/۹۰۱$

پاسخ: ابتدا باید متغیر تصادفی X را به متغیر نرمال استاندارد Z با استفاده از رابطه‌ی $Z=\frac{X-\mu}{\sigma}$ تبدیل کرد سپس احتمالات مختلف را با استفاده از جدول نرمال استاندارد محاسبه نمود.

$$P[X>۳]=۱-P[X<۳]=۱-P[\frac{X-۳}{۰/۴}<\frac{۳-۳}{۰/۴}]=۱-[Z<۰]=۱-۰/۵=۰/۵$$

$$P[X>۳/۳]=۱-P[X<۳/۳]=۱-P[\frac{X-۳}{۰/۴}<\frac{۳/۳-۳}{۰/۴}]=۱-[Z<۰/۷۵]=۱-۰/۷۷۳۴=۰/۲۲۶۶$$

$$P[|X|<۳/۳]=P[-۳/۳<X<۳/۳]=P[\frac{-۳/۳-۳}{۰/۴}<\frac{X-۳}{۰/۴}<\frac{۳/۳-۳}{۰/۴}]$$

$$=P[-۱۵/۷۵<Z<۰/۷۵]=P[Z<۰/۷۵]-P[Z<-۱۵/۷۵]$$

$$=P(Z<۰/۷۵)-[۱-P(Z>۱۵/۷۵)]=P(Z<۰/۷۵)=۰/۷۷۳۴$$

$$P[|X-۳|<۱/۵]=P[-۱/۵<X-۳<۱/۵]=P[\frac{-۱/۵}{۰/۴}<\frac{X-۳}{۰/۴}<\frac{۱/۵}{۰/۴}]=P[-۳/۷۵<Z<۳/۷۵]$$

$$=P(Z<۳/۷۵)-P(Z<-۳/۷۵)=P(Z<۳/۷۵)-[۱-P(Z<۳/۷۵]$$

$$=۲P(Z<۳/۷۵)-۱=(۲\times۱)-۱=۱$$

نکته: در توزیع نرمال استاندارد $P(-۳<Z<۳)$ باشد تقریباً برابر یک می‌باشد.

$$P[X<۴/۹۶]=P[\frac{X-۳}{۰/۴}<\frac{۴/۹۶-۳}{۰/۴}]=P[Z<۴/۹]=۱$$

$$P[۲/۸<X<۳/۱]=P[\frac{۲/۸-۳}{۰/۴}<\frac{X-۳}{۰/۴}<\frac{۳/۱-۳}{۰/۴}]=P[-۰/۵<Z<۰/۲۵]$$

$$=P(Z<۰/۲۵)-P(Z<-۰/۵)=P(Z<۰/۲۵)-[۱-P(Z<۰/۵)]$$

$$=P(Z<۰/۲۵)+P(Z<۰/۵)-۱=۰/۵۹۸۷+۰/۶۹۱۵-۱=۰/۲۹۰۲$$

برای به دست آوردن c ابتدا احتمال را به نرمال استاندارد تبدیل می‌کنیم.

$$P[۳-c<X<۳+c]=۰/۹۰۱$$

$$P[\frac{۳-c-۳}{۰/۴}<\frac{X-۳}{۰/۴}<\frac{۳+c-۳}{۰/۴}]=P[\frac{-c}{۰/۴}<Z<\frac{c}{۰/۴}]$$

$$=P[Z<\frac{c}{۰/۴}]-P[Z<\frac{-c}{۰/۴}]=P[Z<\frac{c}{۰/۴}]-[۱-P[Z<\frac{c}{۰/۴}]]=۲P(Z<\frac{c}{۰/۴})-۱=۰/۹۰۱$$

$$۲P(Z<\frac{c}{۰/۴})=۱/۹۰۱\rightarrow P(Z<\frac{c}{۰/۴})=۰/۹۵۰۵$$

مقدار احتمال ۰/۹۵۰۵ را در داخل جدول نرمال استاندارد پیدا کرده و عدد متناظر با آن را به دست می‌آوریم پس داریم:

$$P(Z<۱/۶۵)=۰/۹۵۰۵$$

حال برای به دست آوردن C داریم:

$$\frac{c}{۰/۴}=۱/۶۵\rightarrow c=۱/۶۵\times۰/۴=۰/۶۶$$

پس:

$$P(۲/۳۴<X<۳/۶۶)=۰/۹۰۱$$

۱۹ ـ تجربه نشان داده که توزیع نمرات دانشجویان در یک درس دارای توزیع نرمال با میانگین ۷۰ و انحراف معیار ۳ است. یک دانشجو به تصادف انتخاب می‌شود مطلوب است:

الف) احتمال این که نمره‌ی ایشان بین ۷۴ و ۹۶ باشد.

ب) احتمال نمره‌ی ایشان کم‌تر از ۷۴ باشد.

ج) چند درصد دانشجویان نمراتشان بیش از ۹۶ است؟

پاسخ: الف) با توجه به این که $\mu=۷۰$ و $\sigma=۳$ است داریم:

$$P(۷۴<X<۹۶)=P(\frac{۷۴-۷۰}{۳}<\frac{X-۷۰}{۳}<\frac{۹۶-۷۰}{۳})=P(۱/۳۳<Z<۸/۶۷)$$

$$=P(Z<۸/۶۷)-P(Z<۱/۳۳)=۱-۰/۹۰۸۲=۰/۰۹۱۸$$

ب)

$$P(X<۷۴)=P(\frac{X-۷۰}{۳}<\frac{۷۴-۷۰}{۳})=P(Z<۱/۳۳)=۰/۹۰۸۲$$

ج)

$$P(X<۹۶)=P(\frac{X-۷۰}{۳}>\frac{۹۶-۷۰}{۳})=P(Z>۸/۶۷)=۱-P(Z<۸/۶۷)=۱-۱=۰$$

صفر درصد از دانشجویان یا هیچ دانشجویی نمره‌اش بیش از ۹۶ نمی‌باشد.

۲۰ ـ استادی از ۵ دانشجو امتحان گرفته که نتایج آن عبارت است از:

۰/۷۵ ۴/۲۵ ۹/۲۵ ۱۲ ۱۸

میانگین و واریانس نمونه برابر با ۸/۷۵ و ۳۶/۱۲ . او علاقه‌مند است به دانشجویان متناسب با نمره‌ای که گرفته‌اند، نمره‌ای اضافه کند به طوری که میانگین نمرات جدید ۱۲ و واریانس آن ۱۶ بشود. چگونه می‌توان این کار را انجام داد؟ (راهنمایی: نمرات را به داده‌های استاندارد تبدیل کنید و سپس نمرات جدید را از رابطه‌ی $Y=۴Z+۱۲$ به دست آورید.)

پاسخ: ابتدا نمرات ۵ دانشجو را از رابطه‌ی $Z=\frac{X-\mu}{\sigma}$ استاندارد می‌کنیم:

$$Z_{۱}=\frac{۰/۷۵-۸/۷۵}{۶/۰۱}=-۱/۳۵ \qquad Z_{۲}=\frac{۴/۲۵-۸/۷۵}{۶/۰۱}=-۰/۷۷$$

$$Z_{۳}=\frac{۹/۲۵-۸/۷۵}{۶/۰۱}=۰/۰۷ \qquad Z_{۴}=\frac{۱۲-۸/۷۵}{۶/۰۱}=۰/۵۲$$

$$Z_{۵}=\frac{۱۸-۸/۷۵}{۶/۰۱}=۱/۵۲$$

حال با استفاده از رابطه $Y = 4Z + 12$ نمرات جدید را به دست می‌آوریم:

$$Y_1 = 4\times(-1/35)+12 = 6/60 \qquad Y_2 = 4\times(-0/77)+12 = 8/92$$
$$Y_3 = 4\times(0/07)+12 = 12/28 \qquad Y_4 = 4\times(0/52)+12 = 14/08$$
$$Y_5 = 4\times(1/52)+12 = 18/08$$

حال برای نمرات زیر که از رابطه‌ی جدید به دست آمده مقدار $\mu=12$ و $\sigma^2=16$ می‌باشد.

۶/۶۰ ۸/۹۲ ۱۲/۲۸ ۱۴/۰۸ ۱۸/۰۸

۲۱ـ فرض کنید متغیر تصادفی X دارای توزیع با میانگین μ و واریانس σ^2 باشد به طوری که $P[X\le 60]=0/1$ و $P[X\le 90]=0/95$ مطلوب است μ و σ^2.

پاسخ: ابتدا احتمالات داده شده را به نرمال استاندارد تبدیل می‌کنیم:

$$P[X\le 60]=P\left[\frac{X-\mu}{\sigma} \le \frac{60-\mu}{\sigma}\right]=0/1$$

احتمال ۰/۱ را داخل جدول نرمال استاندارد به دست می‌آوریم به طور تقریب داریم:

رابطه‌ی (۱) $P[Z\le -1/28]=0/1$

همچنین برای $P[X\le 90]=0/95$ داریم:

$$P[X\le 90]=P\left[\frac{X-\mu}{\sigma} \le \frac{90-\mu}{\sigma}\right]=0/95$$

با استفاده از جدول نرمال استاندارد داریم:

رابطه‌ی (۲) $P[Z\le 1/64]=0/95$

با توجه به رابطه‌ی (۱) و (۲) داریم:

$$\frac{60-\mu}{\sigma}=-1/28 \rightarrow \mu-1/28\sigma=60$$
$$\frac{90-\mu}{\sigma}=1/64 \rightarrow \mu+1/64\sigma=90$$

با توجه به دستگاه معادلات ایجاد شده مقدار μ و σ را به دست می‌آوریم:

$$-\begin{cases}\mu-1/28\sigma=60\\ \mu+1/64\sigma=90\end{cases} \rightarrow \begin{cases}-\mu+1/28\sigma=-60\\ \mu+1/64\sigma=90\end{cases}$$

$$\Rightarrow 2/92\sigma=30 \rightarrow \sigma=\frac{30}{2/92}=10/27$$

حال مقدار $\sigma=10/27$ را داخل یکی از معادلات قرار داده و مقدار μ را به دست می‌آوریم:

$$\mu-1/28(10/27)=60 \rightarrow \mu=60+13/15=73/15$$

پس مقدار $\mu=73/15$ و $\sigma^2=105/47$ می‌باشد.

۲۲ـ در فواصل زمانی مساوی مثلاً یک ساعت، تعداد تلفن به یک مرکز مخابره می‌شود اگر فواصل زمانی از هم مستقل و دارای توزیع نمایی با $\theta=2$ باشد احتمالات زیر را حساب کنید.

الف) $P(X > 3)$

ب) $P(|X| < 3)$

ج) $P(X > 7)$

پاسخ: متغیر تصادفی X دارای تابع چگالی احتمال نمایی با پارامتر $\theta=2$ است.

$$f(x)=\frac{1}{2}e^{-\frac{x}{2}} \qquad x>0 \;,\; \theta>0$$

برای سهولت در محاسبه‌ی احتمالات مختلف در توزیع نمایی می توانیم از تابع توزیع آن استفاده کرده و احتمالات مختلف را حساب کنیم.

$$F(x)=P(X<x)=\int_{-\infty}^{x} f(t)\,dt = 1-e^{-\frac{x}{\theta}}$$

که برای این تمرین داریم:

$$F(x)=P(X<x)=1-e^{-\frac{x}{2}}$$

الف) $P(X>3)=1-P(X\le 3)=1-F(3)=1-(1-e^{-\frac{3}{2}})=e^{-\frac{3}{2}}=0/22$

ب) $P(|X|<3)=P(-3<X<3)=P(0<X<3)=F(3)-F(0)$

$$=1-e^{-\frac{3}{2}}-(1-e^{0})=1-e^{-\frac{3}{2}}=1-0/22=0/78$$

تذکر: تابع چگالی احتمال نمایی در $x>0$ تعریف شده بنابراین $P(X<0)$ برابر صفر می‌باشد.

ج) $P(X>7)=1-P(X\le 7)=1-F(7)=1-(1-e^{-\frac{7}{2}})=e^{-\frac{7}{2}}=0/03$

> نکته: در توابع چگالی احتمال‌های پیوسته، مقدار احتمال در یک نقطه‌ی خاص برابر صفر است چون:
>
> $$P(X=a)=\int_a^a f(x)\,dx=0$$
>
> بنابراین $P(a\le X\le b)$ برابر است با:
>
> $P(a\le X<b)\;,\;P(a<X\le b)\;,\;P(a<X<b)$

۲۳ـ اگر X دارای توزیع نمایی با پارامتر θ باشد $M_X(t)$ و $M'_X(0)$ و $M''_X(0)$ را به دست آورید.

پاسخ: ابتدا تابع مولد گشتاور متغیر تصادفی نمایی را به دست می‌آوریم، می‌دانیم تابع چگالی توزیع نمایی به صورت زیر است:

$$f(x)=P(X=x)=\frac{1}{\theta}e^{-\frac{x}{\theta}} \qquad x>0 \;,\; \theta>0$$

حال داریم:

$$M_X(t)=E(e^{tX})=\int_0^{+\infty} e^{tx}.f(x)\,dx$$

$$=\int_0^{+\infty} e^{tx}\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=\int_0^{+\infty}\frac{1}{\theta}e^{-x(\frac{1}{\theta}-t)}dx$$

$$=\frac{-1}{\theta(\frac{1}{\theta}-t)}e^{-x(\frac{1}{\theta}-t)}\Big|_{x=0}^{x=+\infty}=\frac{1}{1-\theta t}$$

پس:

$$M_X(t)=\frac{1}{1-\theta t}$$

$$E(X)=M'_X(0)=\frac{d}{dt}M_X(t)\Big|_{t=0}=\frac{\theta}{(1-\theta t)^2}\Big|_{t=0}=\theta$$

$$E(X^2)=M''_X(0)=\frac{d^2}{dt^2}M_X(t)\Big|_{t=0}=\frac{2\theta^2}{(1-\theta t)^3}\Big|_{t=0}=2\theta^2$$

همچنین می‌توان با به دست آمدن $E(X)$ و $E(X^2)$ مقدار $Var(X)$ را محاسبه نمود.

$$Var(X)=E(X^2)-[E(X)]^2=2\theta^2-\theta^2=\theta^2$$

۲۴ـ مقدار روزانه (بر حسب اینچ) اندازه‌ی رسوب در مسیر رودخانه متغیر تصادفی گاما با پارامترهای $\alpha=6$ و $\beta=0/2$ است. احتمال این که مقدار رسوب از یک حدی مثلاً دواینچ تجاور کند چقدر است؟

پاسخ: متغیر تصادفی X دارای تابع چگالی احتمال گاما با پارامترهای $\alpha=6$ و $\beta=0/2$ است.

$$f(x)=\frac{1}{(0/2)^2\Gamma(6)}x^{6-1}e^{-\frac{x}{0/2}} \qquad x>0\ ,\ \alpha>0\ ,\ \beta>0$$

که برای محاسبه‌ی احتمالات مختلف توزیع گاما برای سهولت از $F(x)$ استفاده می‌کنیم:

$$F(x)=P(X\le x)=\int_{-\infty}^{x}\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}dx=1-\sum_{i=0}^{\alpha}\frac{1}{i!}\left(\frac{x}{\beta}\right)^i e^{-\frac{x}{\beta}}$$

بنابراین:

$$P(X>2)=1-P(X\le 2)=1-\left[1-\sum_{i=0}^{6}\frac{1}{i!}\left[\frac{2}{0/2}\right]^i e^{-\frac{2}{0/2}}\right]$$

$$=\sum_{i=0}^{6}\frac{1}{i!}(10)^i e^{-10}$$

$$=e^{-10}[1+10+50+166/67+416/67+833/33+1388/89]$$

$$=2866/57e^{-10}=0/13$$

۲۵ـ تابع مولد گشتاورها را برای توزیع گاما به دست آورید و نشان دهید که تحت چه شرایط توزیع گاما به توزیع نمایی تبدیل می‌شود.

پاسخ: اگر متغیر تصادفی X دارای تابع احتمال گاما باشد داریم:

$$f(x)=\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} \qquad x>0\ ,\ \alpha>0\ ,\ \beta>0$$

تابع مولد گشتاور این توزیع به صورت زیر می‌باشد:

$$M_X(t)=E(e^{tX})=\int_0^{\infty}\frac{e^{tx}}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}dx$$

$$=\frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^{\infty}x^{\alpha-1}e^{-x(\frac{1}{\beta}-t)}dx=\frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^{\infty}x^{\alpha-1}e^{-x(\frac{1-\beta t}{\beta})}dx$$

با در نظر گرفتن تغییر متغیر

$$y=x\left(\frac{1-\beta t}{\beta}\right)\ \rightarrow\ x=\frac{\beta y}{1-\beta t}$$

$$dy=\frac{1-\beta t}{\beta}dx$$

$$M_X(t)=\frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^{\infty}\left(\frac{\beta y}{1-\beta t}\right)^{\alpha-1}e^{-y}\frac{\beta}{1-\beta t}dy$$

$$=\frac{1}{\beta^\alpha\Gamma(\alpha)}\int_0^{\infty}\frac{\beta^\alpha}{(1-\beta t)^\alpha}y^{\alpha-1}e^{-y}dy$$

$$=\frac{1}{(1-\beta t)^\alpha\Gamma(\alpha)}\int_0^{\infty}y^{\alpha-1}e^{-y}dy=\frac{1}{(1-\beta t)^\alpha}=(1-\beta t)^{-\alpha}$$

نکته: $\int_0^{\infty}y^{\alpha-1}e^{-y}dy=\Gamma(\alpha)$

بنابراین داریم:

$$M_X(t)=\frac{1}{(1-\beta t)^\alpha}=(1-\beta t)^{-\alpha}$$

توزیع نمایی حالت خاصی از توزیع گاما می‌باشد اگر در توزیع گاما قرار دهیم $\alpha=1$ و $\beta=\theta$.

پس خواهیم داشت:

$$f(x)=\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} \qquad x>0\ ,\ \beta>0\ ,\ \alpha>0$$

حال اگر $\alpha=1$ و $\beta=\theta$ داریم:

$$f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}} \qquad x>0\ ,\ \theta>0$$

۲۶ـ اگر X دارای توزیع کی دو با ۵ درجه‌ی آزادی باشد مطلوب است:

$$E(X+7) \quad , \quad E(X^2+1) \quad , \quad V(X+4) \quad , \quad V(2X+3)$$

پاسخ: می‌دانیم توزیع کی دو حالت خاص توزیع گاماست که در آن $\beta=2$ و $\alpha=\frac{r}{2}$ می‌باشد و این توزیع دارای میانگین r و واریانس 2r است حال با توجه به خواص میانگین و واریانس عبارت‌های بالا را محاسبه می‌کنیم (r را درجه‌ی آزادی توزیع گوییم)

$$E(X)=r=5$$

$$V(X)=2r=10$$

$$1)\ E(X+7)=E(X)+7=5+7=12$$

$$2)\ E(X^2+1)=E(X^2)+1$$

برای محاسبه‌ی عبارت باید مقدار $E(X^2)$ را به دست آوریم با توجه به رابطه‌ی واریانس داریم:

$$V(X)=E(X^2)-[E(X)]^2 \rightarrow E(X^2)=V(X)+[E(X)]^2$$

$$E(X^2)=2r+r^2 \rightarrow E(X^2)=10+25=35$$

حال داریم:

$$E(X^2)+1=35+1=36$$

$$3)\ V(X+4)=V(X)+0=10$$

$$4)\ V(2X+3)=2^2V(X)+0=4\times 10=40$$

۲۷ـ اگر X دارای تابع چگالی بتا با پارامترهای α و β باشد نشان دهید که:

$$E(X)=\frac{\alpha}{\alpha+\beta} \quad , \quad V(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

پاسخ: متغیر تصادفی X دارای توزیع بتا است اگر تابع احتمال آن به صورت زیر باشد:

$$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} \qquad 0<x<1 \quad , \quad \alpha>0 \quad , \quad \beta>0$$

برای محاسبه‌ی $E(X)$ بنا به تعریف داریم:

$$E(X)=\int_{-\infty}^{+\infty}x.f(x)dx=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x.x^{\alpha-1}(1-x)^{\beta-1}dx$$

همچنین می‌دانیم:

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}=B(\alpha,\beta)$$

پس داریم:

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)}=\frac{\alpha}{\alpha+\beta}$$

که در آن دیده می‌شود که انتگرال همان $B(\alpha+1,\beta)$ و از واقعیت $\Gamma(\alpha+1)=\alpha.\Gamma(\alpha)$ و $\Gamma(\alpha+\beta+1)=(\alpha+\beta)\Gamma(\alpha+\beta)$ نیز استفاده شده است.

حال برای محاسبه‌ی واریانس باید ابتدا $E(X^2)$ را محاسبه کرد که طی مراحل مشابهی مانند بالا داریم:

$$E(X^2)=\int_{-\infty}^{+\infty}x^2.f(x)dx=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^2.x^{\alpha-1}(1-x)^{\beta-1}dx$$

$$=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)}=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{(\alpha+1)\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)}$$

$$=\frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}$$

حال برای محاسبه‌ی واریانس داریم:

$$V(X)=E(X^2)-[E(X)]^2$$

$$V(X)=\frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}-\left[\frac{\alpha}{\alpha+\beta}\right]^2=\frac{(\alpha+1)\alpha(\alpha+\beta)-\alpha^2(\alpha+\beta+1)}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

$$=\frac{\alpha^3+\alpha^2\beta+\alpha^2+\alpha\beta-\alpha^3-\alpha^2\beta-\alpha^2}{(\alpha+\beta+1)(\alpha+\beta)^2}=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

۲۸ـ در یک استان، بخش‌هایی از بزرگراهی که در یک سال احتیاج به تعمیر دارد، دارای توزیع بتا با پارامترهای $\alpha=3$ و $\beta=2$ مطلوب است:

الف) به طور متوسط چند درصد از بخش‌های بزرگراه در یک سال به تعمیر احتیاج دارد؟

ب) احتمال این که حداکثر نیمی از بخش‌های بزرگراه در یک سال به تعمیر احتیاج داشته باشد چقدر است؟

پاسخ: الف) متغیر تصادفی X دارای توزیع بتا با پارامترهای $\alpha=3$ و $\beta=2$ می‌باشد اگر تابع چگالی آن به صورت زیر باشد:

$$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} \qquad 0<x<1 \quad , \quad \alpha>0 \quad , \quad \beta>0$$

که میانگین توزیع به صورت $\frac{\alpha}{\alpha+\beta}$ می‌باشد پس:

$$E(x)=\frac{\alpha}{\alpha+\beta}=\frac{3}{3+2}=\frac{3}{5}=0/60$$

پس ۶۰ درصد بزرگراه در یک سال به تعمیر احتیاج دارد.

ب)

$$P(X<\tfrac{1}{2})=\int_0^{\frac{1}{2}}\frac{\Gamma(3+2)}{\Gamma(3)\Gamma(2)}x^2(1-x)dx=12\int_0^{\frac{1}{2}}x^2(1-x)dx$$

$$=12\int_0^{\frac{1}{2}}x^2-x^3dx=12\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^{\frac{1}{2}}=12\left[\frac{4x^3-3x^4}{12}\right]_0^{\frac{1}{2}}$$

$$=(4\times\tfrac{1}{8})-(3\times\tfrac{1}{16})=\frac{8-3}{16}=\frac{5}{16}=0/31$$

۲۹ـ اگر X دارای استودنت با r درجه‌ی آزادی باشد نشان دهید که $E(X)=0$.

نکته: اگر توزیع داده‌ها متقارن باشد، در این صورت گشتاور مرکزی مراتب فرد (یعنی ... , ۵ , ۳ , ۱=r) همگی مساوی صفر است.

پاسخ: با توجه به این که توزیع استودنت یک توزیع متقارن می‌باشد پس گشتاور مرکزی مراتب فرد آن همگی مساوی صفر است پس $E(X)=m_r=0$

$$E(X)=\int_{-\infty}^{+\infty} x.f(x)\,dx=\int_{-\infty}^{+\infty} x.\frac{\Gamma\left[\frac{(r+1)}{2}\right]}{\Gamma(\frac{r}{2})}.\frac{1}{\sqrt{r\pi}}.\frac{1}{(1+\frac{x^2}{r})^{(\frac{r+1}{2})}}\,dx$$

$$=\frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})}\times\frac{1}{\sqrt{r\pi}}\int_{-\infty}^{+\infty}\frac{x}{(1+\frac{x^2}{r})^{(\frac{r+1}{2})}}\,dx=0$$

۳۰ـ فرض کنید X دارای توزیع فیشر با درجات آزادی $r_1=7$ و $r_2=25$ باشد احتمالات زیر را با استفاده از جدول ضمیمه‌ی (۶) حساب کنید.

$P[X>3/46]$, $P[X<3/46]$, $P[|X|<3/46]$

پاسخ: چون $F_{0/01}(7, 25)=3/46$ بنابراین:

$P(X>3/46)=0/01$

$P(X<3/46)=1-P(X>3/46)=1-0/01=0/99$

$P(|X|<3/46)=P(-3/46<X<3/46)$

چون برای، توزیع فیشر $X>0$ تعریف شده است پس داریم:

$P(X<3/46)=1-P(X>3/46)=1-0/01=0/99$

خودآزمایی فصل پنجم

۱ـ اگر مشاهدات نمونه‌ی تصادفی $X_1 , X_2 , X_3 , X_4 , X_5$ برابر با ۰/۲، ۰/۷، ۰/۱-، ۰، ۵ باشند مقدار آماره‌های زیر را به دست آورید.

$$\bar{X}=\frac{1}{5}\sum_{i=1}^{5} X_i \quad , \quad S^2=\frac{1}{4}\sum_{i=1}^{5}(X_i-\bar{X})^2 \quad , \quad M=\sum_{i=1}^{5}\frac{|X_i-\bar{X}|}{4}$$

پاسخ:

$$\bar{X}=\frac{1}{5}\sum_{i=1}^{5} X_i=\frac{1}{5}(0.2+0.7-0.1+0+5)=\frac{5.80}{5}=1.16$$

$$S^2=\frac{1}{4}\sum_{i=1}^{5}(X_i-\bar{X})^2$$

$$=\frac{1}{4}[(0.2-1.16)^2+(0.7-1.16)^2+(-0.1-1.16)^2+(0-1.16)^2+(5-1.16)^2]$$

$$=\frac{1}{4}[0.92+0.21+1.59+1.35+14.75]=\frac{18.82}{4}=4.71$$

$$M=\sum_{i=1}^{5}\frac{|X_i-\bar{X}|}{4}$$

$$=\frac{1}{4}[|0.2-1.161|+|0.7-1.16|+|-0.1-1.161|+|0-1.161|+|5-1.161|]$$

$$=\frac{1}{4}[0.96+0.26+1.26+1.16+3.84]=\frac{7.68}{4}=1.92$$

۲ـ اگر X_1 , X_2 , X_3 یک نمونه‌ی سه تایی از جامعه‌ای با میانگین μ و واریانس σ^2 باشد کدام یک از آماره‌های زیر دارای کم‌ترین واریانس است؟

$$T_1=\frac{2X_1+3X_2+2X_3}{6} \qquad T_2=\frac{X_1+X_2+X_3}{3}$$

$$T_3=\frac{X_3+X_1}{2} \qquad T_4=\frac{2X_1+X_2}{3}+\frac{1}{5}$$

پاسخ: ابتدا برای چهار آماره‌ی بالا مقدار میانگین و واریانس را محاسبه می‌کنیم بهترین آماره، آماره‌ای است که ناریب و دارای کم‌ترین واریانس باشد.

$$E(T_1)=E(\frac{2X_1+3X_2+2X_3}{6})=\frac{1}{6}[2E(X_1)+3E(X_2)+2E(X_3)]$$

$$=\frac{1}{6}[2\mu+3\mu+2\mu]=\frac{7}{6}\mu=1.17\mu$$ آماره دارای اریبی است.

$$V(T_1)=V(\frac{2X_1+3X_2+2X_3}{6})=\frac{4}{36}V(X_1)+\frac{9}{36}V(X_2)+\frac{4}{36}V(X_3)$$

$$=\frac{4}{36}\sigma^2+\frac{9}{36}\sigma^2+\frac{4}{36}\sigma^2=\frac{17}{36}\sigma^2=0.47\sigma^2$$

$$E(T_2)=E(\frac{X_1+X_2+X_3}{3})=\frac{1}{3}[E(X_1)+E(X_2)+E(X_3)]$$

$$=\frac{1}{3}[\mu+\mu+\mu]=\frac{3\mu}{3}=\mu$$ آماره ناریب می‌باشد.

$$V(T_2)=V(\frac{X_1+X_2+X_3}{3})=\frac{1}{9}[V(X_1)+V(X_2)+V(X_3)]$$

$$=\frac{1}{9}[\sigma^2+\sigma^2+\sigma^2]=\frac{3}{9}\sigma^2=\frac{\sigma^2}{3}=0.33\sigma^2$$

$$E(T_3)=E(\frac{X_3+X_1}{2})$$

$$=\frac{1}{2}[E(X_3)+E(X_1)]=\frac{1}{2}[\mu+\mu]=\frac{2\mu}{2}=\mu$$ آماره ناریب می‌باشد.

$$V(T_3)=V(\frac{X_3+X_1}{2})=\frac{1}{4}[V(X_3)+V(X_1)]$$

$$=\frac{1}{4}[\sigma^2+\sigma^2]=\frac{2}{4}\sigma^2=0.5\sigma^2$$

$$E(T_4)=E(\frac{2X_1+X_2}{3}+\frac{1}{5})=\frac{2}{3}E(X_1)+\frac{1}{3}E(X_2)+\frac{1}{5}$$

$$=\frac{2}{3}\mu+\frac{1}{3}\mu+\frac{1}{5}=\mu+\frac{1}{5}$$ آماره دارای اریبی است.

$$V(T_4)=V(\frac{2X_1+X_2}{3}+\frac{1}{5})=\frac{4}{9}V(X_1)+\frac{1}{9}V(X_2)+0$$

$$=\frac{4}{9}\sigma^2+\frac{1}{9}\sigma^2=\frac{5}{9}\sigma^2=0.56\sigma^2$$

آماره‌ی T_2 کاراترین آماره می‌باشد چون علاوه بر این که دارای کم‌ترین واریانس است، ناریب نیز می‌باشد.

نکته: چون مقدار $\sigma^2>0$ است پس آماره‌ای کم‌ترین واریانس را دارد که ضریب σ^2 کوچک‌تر باشد.

۳ـ اگر $\bar{X}$ میانگین یک نمونه m تایی از جامعه‌ای با میانگین μ_1 و واریانس σ_1^2 و $\bar{Y}$ میانگین یک نمونه‌ی n تایی از جامعه‌ای دیگر با میانگین μ_2 و واریانس σ_2^2 باشند مطلوب است امید ریاضی و واریانس آماره‌های زیر.

$$T_1=\frac{m\bar{X}+n\bar{Y}}{m+n} \quad , \quad T_2=\frac{m\bar{X}}{m^2+1}+\frac{n\bar{Y}}{n^2+1}$$

پاسخ: $\bar{X}$ دارای میانگین μ_1 و واریانس $\frac{\sigma_1^2}{m}$ و $\bar{Y}$ دارای میانگین μ_2 و واریانس $\frac{\sigma_2^2}{n}$ می‌باشد پس داریم:

$$E(T_1)=E\left[\frac{m\bar{X}+n\bar{Y}}{m+n}\right]=\frac{1}{m+n}\left[mE(\bar{X})+nE(\bar{Y})\right]=\frac{m\mu_1+n\mu_2}{m+n}$$

$$V(T_1)=V\left[\frac{m\bar{X}+n\bar{Y}}{m+n}\right]=\frac{1}{(m+n)^2}\left[m^2V(\bar{X})+n^2V(\bar{Y})\right]$$

$$=\frac{1}{(m+n)^2}\left[m^2\cdot\frac{\sigma_1^2}{m}+n^2\cdot\frac{\sigma_2^2}{n}\right]=\frac{m\sigma_1^2+n\sigma_2^2}{(m+n)^2}$$

$$E(T_2)=E\left[\frac{m\bar{X}}{m^2+1}+\frac{n\bar{Y}}{n^2+1}\right]=\frac{m}{m^2+1}E(\bar{X})+\frac{n}{n^2+1}E(\bar{Y})=\frac{m\mu_1}{m^2+1}+\frac{n\mu_2}{n^2+1}$$

$$V(T_2)=V\left[\frac{m\bar{X}}{m^2+1}+\frac{n\bar{Y}}{n^2+1}\right]=\frac{m^2}{(m^2+1)^2}V(\bar{X})+\frac{n^2}{(n^2+1)^2}V(\bar{Y})$$

$$=\frac{m^2}{(m^2+1)^2}\cdot\frac{\sigma_1^2}{m}+\frac{n^2}{(n^2+1)^2}\cdot\frac{\sigma_2^2}{n}=\frac{m\sigma_1^2}{(m^2+1)^2}+\frac{n\sigma_2^2}{(n^2+1)^2}$$

۴ـ اگر $\bar{X}$ و S^2 به ترتیب میانگین و واریانس یک نمونه‌ی n تایی از توزیع نرمال با میانگین μ و واریانس σ^2 باشند امید ریاضی و واریانس آماره‌های زیر را به دست آورید (k و c ثابت هستند).

$$T_1=k\bar{X}+c \quad , \quad T_2=kS^2+c$$

$$T_3=\frac{\bar{X}}{k+c}-10 \quad , \quad T_4=\frac{S^2}{k+c}+10$$

تذکر: می‌دانیم اگر $X_1, X_2, \ldots, X_n$ یک نمونه‌ی تصادفی n تایی از جامعه‌ی نرمال با میانگین μ و واریانس σ^2 باشند آن‌گاه:

$$E(\bar{X})=\mu \quad , \quad V(\bar{X})=\frac{\sigma^2}{n}$$

همچنین برای S^2 داریم:

$$E(S^2)=\sigma^2 \quad , \quad V(S^2)=\frac{2\sigma^4}{n-1}$$

پاسخ:

$$E(T_1)=E(k\bar{X}+c)=kE(\bar{X})+c=k\mu+c$$

$$V(T_1)=V(k\bar{X}+c)=k^2V(\bar{X})+0=k^2+\frac{\sigma^2}{n}=\frac{(k\sigma)^2}{n}$$

$$E(T_2)=E(kS^2+c)=kE(S^2)+c=k\sigma^2+c$$

$$V(T_2)=V(kS^2+c)=k^2V(S^2)+0=k^2\cdot\frac{2\sigma^4}{n-1}=\frac{2k^2\sigma^4}{n-1}$$

$$E(T_3)=E\left(\frac{\bar{X}}{k+c}-10\right)=\frac{1}{k+c}E(\bar{X})-10=\frac{\mu}{k+c}-10$$

$$V(T_3)=V\left(\frac{\bar{X}}{k+c}-10\right)=\frac{1}{(k+c)^2}V(\bar{X})-0=\frac{1}{(k+c)^2}\cdot\frac{\sigma^2}{n}=\frac{\sigma^2}{n(k+c)^2}$$

$$E(T_4)=E\left(\frac{S^2}{k+c}+10\right)=\frac{1}{k+c}E(S^2)+10=\frac{\sigma^2}{k+c}+10$$

$$V(T_4)=V\left(\frac{S^2}{k+c}+10\right)=\frac{1}{(k+c)^2}V(S^2)+0=\frac{1}{(k+c)^2}\cdot\frac{2\sigma^4}{(n-1)}=\frac{2\sigma^4}{(k+c)^2(n-1)}$$

۵ـ اگر X دارای توزیع پواسن با پارامتر λ و Y دارای توزیع دو جمله‌ای با پارامترهای n و p باشند توزیع مشترک X و Y را به دست آورید.

پاسخ: می‌دانیم $X\sim P(\lambda)$ است پس تابع احتمال آن به صورت زیر می‌باشد:

$$f(x)=\frac{e^{-\lambda}\lambda^x}{x!} \qquad x=0,1,\ldots$$

همچنین $Y\sim B(n,p)$ می‌باشد و تابع احتمال آن به صورت زیر می‌باشد:

$$f(y)=\begin{bmatrix}n\\y\end{bmatrix}p^y q^{n-y} \qquad y=0,1,\ldots,n$$

حال توزیع مشترک آن‌ها به صورت زیر تعریف می‌شود اگر X و Y مستقل از هم باشند:

$$f(x,y)=f(x).f(y)$$

$$f(x,y)=\frac{e^{-\lambda}\lambda^x}{x!}\cdot\begin{bmatrix}n\\y\end{bmatrix}p^y q^{n-y}$$

۶ـ اگر $X_1, X_2, \ldots, X_n$ یک نمونه‌ی n تایی از تابع چگالی زیر باشد توزیع مشترک $X_1, X_2, \ldots, X_n$ را به دست آورید.

$$f(x)=\frac{x+1}{\theta(\theta+1)}e^{-\frac{x}{\theta}} \qquad x>0$$

پاسخ: تابع احتمال برای X_1:

$$f(x_1)=\frac{x_1+1}{\theta(\theta+1)}e^{-\frac{x_1}{\theta}} \qquad x_1>0$$

تابع احتمال برای X_2:

$$f(x_2)=\frac{x_2+1}{\theta(\theta+1)}e^{-\frac{x_2}{\theta}} \qquad x_2>0$$

$\vdots$

تابع احتمال برای X_n:

$$f(x_n)=\frac{x_n+1}{\theta(\theta+1)}e^{-\frac{x_n}{\theta}} \qquad x_n>0$$

حال توزیع مشترک $X_۱ , X_۲ , \dots , X_n$ به صورت زیر می‌باشد:

$$f(x_۱,x_۲,\dots,x_n)=f(x_۱).f(x_۲)\dots f(x_n)=\prod_{i=۱}^{n} f(x_i)$$

$$=\frac{(x_۱+۱)(x_۲+۱)\dots(x_n+۱)}{\theta^n(\theta+۱)^n}.e^{-\frac{x_۱}{\theta}}.e^{-\frac{x_۲}{\theta}}.\dots.x^{-\frac{x_n}{\theta}}$$

$$=\frac{\prod_{i=۱}^{n}(x_i+۱)}{\theta^n(\theta+۱)^n}.e^{-\frac{۱}{\theta}\sum_{i=۱}^{n}x_i} \qquad x_i>۰$$

پس توزیع مشترک $X_۱ , X_۲ , \dots , X_n$ به صورت زیر می‌باشد:

$$f(x_۱,x_۲,\dots,x_n)=\frac{\prod_{i=۱}^{n}(x_i+۱)}{\theta^n(\theta+۱)^n}.e^{-\frac{۱}{\theta}\sum_{i=۱}^{n}x_i} \qquad x_i>۰$$

۷ـ جامعه‌شناسی ادعا می‌کند که قد افراد شهری دارای توزیع نرمال با میانگین ۱۷۰Cm و انحراف معیار ۴ است. در یک نمونه ۲۵تایی احتمال این که متوسط قد افراد نمونه بین ۱۷۵ تا ۱۶۵ باشد چقدر است؟

پاسخ: $\bar{X}$ دارای توزیع نرمال با میانگین ۱۷۰ و واریانس $\frac{۱۶}{۲۵}$ است و استاندارد آن برابر است با:

$$Z=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{\bar{X}-۱۷۰}{\frac{۴}{۵}}=\frac{۵(\bar{X}-۱۷۰)}{۴}$$

$$P(۱۶۵<\bar{X}<۱۷۵)=P\left[\frac{۱۶۵-\mu}{\frac{\sigma}{\sqrt{n}}}<\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}<\frac{۱۷۵-\mu}{\frac{\sigma}{\sqrt{n}}}\right]$$

$$=P\left[\frac{۵(۱۶۵-۱۷۰)}{۴}<Z<\frac{۵(۱۷۵-۱۷۰)}{۴}\right]=P\,[-۶/۲۵<Z<۶/۲۵]$$

$$=P\,[Z<۶/۲۵]-P\,[Z<-۶/۲۵]=P\,[Z<۶/۲۵]-[۱-P\,(Z<۶/۲۵)]$$

$$=۲P\,[Z<۶/۲۵]-۱=(۲\times ۱)-۱=۱$$

نکته: در توزیع نرمال استاندارد مقدار احتمال تقریباً بین (۳/۵ و ۳/۵-) می‌باشد و مقدار احتمال خارج از این دو عدد خیلی ناچیز است که تقریباً برابر صفر در نظر می‌گیریم.

$$P(Z<-۳/۵)\simeq ۰ \quad , \quad P(Z<۳/۵)\simeq ۱ \quad , \quad P(Z>۳/۵)\simeq ۰$$

۸ـ مدیر کارخانه‌ی سازنده‌ی رایانه‌ای ادعا می‌کند که ۰/۲ درصد تولیدات رایانه کارخانه معیوب هستند. در یک نمونه‌ی ۱۰۰ تایی مطلوب است $P(\bar{X}>۰/۳)$.

پاسخ: متغیر تصادفی X دارای توزیع دوجمله‌ای با پارامترهای $n=۱۰۰$ و $p=\frac{۰/۲}{۱۰۰}=۰/۰۰۲$ است و چون مقدار n بزرگ است و p به سمت صفر میل می‌کند با استفاده از تقریب پواسن داریم:

$$\lambda=np=۱۰۰\times ۰/۰۰۲=۰/۲$$

که در این توزیع $E(X)=\lambda=۰/۲$ و $V(X)=\lambda=۰/۲$ می‌باشد. پس:

$$X_i \sim P(۰/۲)$$

حال می‌خواهیم $V(\bar{X})$ و $E(\bar{X})$ را به دست آوریم پس داریم:

$$E(\bar{X})=\frac{۱}{n}\sum_{i=۱}^{n}E(X_i)=\frac{n\lambda}{n}=\lambda$$

$$V(\bar{X})=V(\frac{۱}{n}\sum_{i=۱}^{n}X_i)=\frac{۱}{n^۲}\sum_{i=۱}^{n}V(X_i)=\frac{n\lambda}{n^۲}=\frac{\lambda}{n}$$

حال داریم:

$$P(\bar{X}>۰/۳)=P\left[\frac{\bar{X}-\lambda}{\sqrt{\frac{\lambda}{n}}}>\frac{۰/۳-\lambda}{\sqrt{\frac{\lambda}{n}}}\right]$$

$$=P(Z>\frac{۰/۳-۰/۲}{\sqrt{\frac{۰/۲}{۱۰۰}}})=P(Z>۲/۲۲)=۱-P\,(Z<۲/۲۲)=۱-۰/۹۸۶۸=۰/۰۱۳۲$$

همچنین می‌توان احتمال داده شده را با استفاده از تقریب دوجمله‌ای به نرمال محاسبه نمود بنابراین داریم:

$$X\sim B(۱۰۰ \,,\, ۰/۰۰۲)$$

همچنین می‌دانیم که در توزیع دوجمله‌ای $E(X)=np$ و $V(X)=npq$ می‌باشد حال می‌خواهیم میانگین و واریانس $\bar{X}$ را به دست آوریم، بنابراین داریم:

$$E(\bar{X})=E(\frac{۱}{n}\sum_{i=۱}^{n}X_i)=\frac{۱}{n}\sum_{i=۱}^{n}E(X_i)=\frac{n.np}{n}=np$$

$$V(\bar{X})=V(\frac{۱}{n}\sum_{i=۱}^{n}X_i)=\frac{۱}{n^۲}\sum_{i=۱}^{n}V(X_i)=\frac{n.npq}{n^۲}=pq$$

پس $\bar{X}\sim N(np \,,\, pq)$ بنابراین داریم:

$$n=۱۰۰ \quad , \quad p=\frac{۰/۲}{۱۰۰}=۰/۰۰۲ \quad , \quad q=۱-۰/۰۰۲=۰/۹۹۸$$

$$P(\bar{X}>۰/۳)=P(\frac{\bar{X}-np}{\sqrt{pq}}>\frac{۰/۳-np}{\sqrt{pq}})=P(Z>\frac{۰/۳-۰/۲}{\sqrt{۰/۰۰۲\times ۰/۹۹۸}})$$

$$=P(Z>۲/۲۲)=۱-P(Z<۲/۲۲)=۱-۰/۹۸۶۸=۰/۰۱۳۲$$

مشاهده می‌شود از هر دو روش تقریب پواسن و تقریب نرمال، چون مقدار n بزرگ است جواب‌ها یکسان در می‌آید.

۹ـ اگر متغیر تصادفی X دارای تابع احتمال زیر باشد برای یک نمونه‌ی ۲۵ تایی احتمالات زیر را محاسبه کنید.

x	۱	۲	۳	۴
f(x)	۰/۲۵	۰/۲۵	۰/۲۵	۰/۲۵

$$P(\overline{X}>3) \quad , \quad P(1<\overline{X}<2.5) \quad , \quad P(\overline{X}<0.7)$$

> نکته: چون متغیر گسسته می‌باشد و می‌خواهیم از تقریب نرمال استفاده کنیم لازم است از تصحیح پیوستگی استفاده شود ولی چون در کتاب برای توزیع پواسن نیز تصحیح پیوستگی را به کار نبرده ما نیز در این تمرین استفاده نمی‌کنیم.

پاسخ: ابتدا مقدار μ و σ^2 را برای متغیر تصادفی X به دست می‌آوریم:

$$\mu=E(X)=\sum_{x=1}^{4} x.f(x)=0.25+0.5+0.75+1=2.5$$

$$E(X^2)=\sum_{x=1}^{4} x^2.f(x)=0.25+1+2.25+4=7.5$$

$$\sigma^2=V(X)=E(X^2)-[E(X)]^2=7.5-(2.5)^2=7.5-6.25=1.25$$

$$\sigma=\sqrt{1.25}=1.12$$

پس متغیر تصادفی X دارای میانگین ۲/۵ و واریانس ۱/۲۵ می‌باشد و $\overline{X}$ دارای میانگین $\mu=2.5$ و واریانس $\frac{\sigma^2}{n}=\frac{1.25}{25}$ می‌باشد بنابراین:

$$P[\overline{X}>3]\approx P\left[\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}>\frac{3-\mu}{\frac{\sigma}{\sqrt{n}}}\right]=P\left[Z>\frac{3-2.5}{\frac{1.12}{5}}\right]$$

$$=P[Z>2.27]=1-P[Z<2.27]=1-0.9884=0.0116$$

$$P[1<\overline{X}<2.5]\approx P\left[\frac{1-2.5}{\frac{1.12}{5}}<\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}<\frac{2.5-2.5}{\frac{1.12}{5}}\right]=P[-6.82<Z<0]$$

$$=P[Z<0]-P[Z<-6.82]=0.5-0=0.5$$

$$P[\overline{X}>0.7]\approx P\left[\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}>\frac{0.7-2.5}{\frac{1.12}{5}}\right]=P[Z>-8.18]=1$$

۱۰ـ اگر یک نمونه‌ی ۷ تایی از جامعه‌ای نرمال با میانگین ۵ دارای انحراف معیار S=۱/۰۸ باشد مطلوب است $P(\overline{X}>6)$.

پاسخ:

$$P[\overline{X}>6]=P\left[\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}>\frac{6-\mu}{\frac{S}{\sqrt{n}}}\right]=p\left[T>\frac{6-5}{\frac{1.08}{\sqrt{7}}}\right]=P[T>2.44]$$

با توجه به جدول، متغیر t دارای درجه‌ی آزادی n-1=۶ است پس داریم:

$$P[T>2.44]=0.025$$

۱۱ـ تعداد دقایقی که پزشکی صرف معاینه‌ی هر بیمار می‌کند دارای توزیع نمایی با $\theta=9$ است. در یک نمونه‌ی ۲۵ تایی احتمال این که به طور متوسط بیش از ۱۵ دقیقه صرف معاینه‌ی بیماری کند چقدر است؟

پاسخ: در این تمرین حجم نمونه n=۲۵ تا حدود بزرگ و $\sigma^2=V(X)=81$ متناهی است. شرایط قضیه‌ی حد مرکزی برقرار است. بنابراین داریم:

$$E(X)=\theta=\mu=9$$

$$V(X)=\theta^2=\sigma^2=81$$

$$P[\overline{X}>15]\approx P\left[\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}>\frac{15-\mu}{\frac{\sigma}{\sqrt{n}}}\right]=p\left[Z>\frac{15-9}{\frac{9}{5}}\right]=P[Z>3.33]$$

$$=1-P[Z\le 3.33]=1-0.9996=0.004$$

۱۲ـ تعداد ماهی‌هایی که یک ماهیگیر در هر ساعت از دریاچه‌ای صید می‌کند دارای توزیع پواسن با $\lambda=1.5$ است احتمال این که ماهیگیر در ۴۸ ساعت کاری به طور متوسط بین ۱ تا ۳ ماهی صید کند چقدر است؟

پاسخ: اگر n=۴۸ و $\lambda=1.5$ در نظر بگیریم، با استفاده از قضیه‌ی حد مرکزی داریم:

$$P[1<\overline{X}<3]\approx P\left[\frac{1-\lambda}{\sqrt{\frac{\lambda}{n}}}<\frac{\overline{X}-\lambda}{\sqrt{\frac{\lambda}{n}}}<\frac{3-\lambda}{\sqrt{\frac{\lambda}{n}}}\right]=P\left[\frac{1-1.5}{\sqrt{\frac{1.5}{48}}}<Z<\frac{3-1.5}{\sqrt{\frac{1.5}{48}}}\right]$$

$$=P[-2.94<Z<8.82]=P[Z<8.82]-P[Z<-2.94]$$

$$=1-[1-P[Z<2.94]]=P[Z<2.94]=0.9984$$

۱۳ـ سوابق طبی نشان می‌دهند که در شهری یک نفر از هر ده نفر دچار بیماری تیروئید است. اگر بیست نفر را در این شهر به تصادف انتخاب کنیم و مورد آزمایش قرار دهیم احتمال این که حداقل یکی از آن‌ها دچار تیروئید باشد چقدر است؟

الف) با استفاده از جدول توزیع دو جمله‌ای

ب) با استفاده از جدول توزیع پواسن

ج) با استفاده از تقریب نرمال

پاسخ: الف) اگر X دارای توزیع دو جمله‌ای با پارامترهای $p=\frac{1}{10}$ و $n=20$ باشد داریم:

$$P(X\ge 1)=1-P(X<1)=1-P[X=0]=1-\left[\binom{20}{0}\left(\frac{1}{10}\right)^{0}\left(\frac{9}{10}\right)^{20}\right]$$

$$=1-0/1216=0/8784$$

ب) اگر X دارای توزیع پواسن با پارامتر $\lambda=np=20\times\frac{1}{10}=2$ باشد داریم:

$$P(X\ge 1)=1-P(X<1)=1-P[X=0]=1-\left[\frac{e^{-2}2^{0}}{0!}\right]=1-0/1353=0/8647$$

ج) با استفاده از تقریب نرمال $\mu=np$ و $\sigma^2=npq$ است.

$$P(X\ge 1)=1-P(X<1)$$

با توجه به این که توزیع دوجمله‌ای گسسته و توزیع نرمال پیوسته می‌باشد با استفاده از تصحیح پیوستگی داریم:

$$=1-P(X<1/5)=1-P\left[\frac{X-np}{\sqrt{npq}}<\frac{1/5-np}{\sqrt{npq}}\right]$$

$$=1-P\left[Z<\frac{1/5-2}{\sqrt{1/8}}\right]=1-P[Z<-0/37]=1-[1-P[Z<0/37]]$$

$$=P[Z<0/37]=0/6443$$

در این تمرین هر چه مقدار n بزرگ‌تر باشد و p در همسایگی $\frac{1}{2}$ باشد تقریب نرمال برای توزیع دوجمله‌ای مناسب تر می‌باشد.

۱۴ـ اگر $X_1, X_2, \dots, X_{25}$ دارای توزیع نرمال با میانگین ۱۵۰۰ و واریانس ۱۰۰ باشند مطلوب است:

$$P\left[\bar{X}<150 \ , \ S^2>164\right]$$

پاسخ: با توجه به این که $n=25$ و $\mu=1500$ و $\sigma^2=100$ است داریم:

$$P\left[\bar{X}<150 \ , \ S^2>164\right]=P\left[\bar{X}<150\right]P\left[S^2>164\right]$$

طبق قضیه‌ی ۵ـ۷ـ۴، $\bar{X}$ و S^2 از هم مستقل‌اند. پس:

$$P(\bar{X}<150)P(S^2>164)=P\left[\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}<\frac{150-\mu}{\frac{\sigma}{\sqrt{n}}}\right].P\left[\frac{(n-1)S^2}{\sigma^2}>\frac{164(n-1)}{\sigma^2}\right]$$

$$=P\left[Z<\frac{150-1500}{2}\right].P\left[\frac{(n-1)S^2}{\sigma^2}>\frac{164(24)}{100}\right]$$

$$=P(Z<-675).P\left[\frac{(n-1)S^2}{\sigma^2}>39/36\right]$$

با توجه به این که متغیر $\frac{(n-1)S^2}{\sigma^2}$ دارای توزیع کی دو با ۲۴ درجه‌ی آزادی است پس داریم:

$$=0\times 0/025=0$$

۱۵ـ یک فرایند تولید یاتاقان‌های معین تحت کنترل است. اگر قطر یاتاقان‌ها با میانگین ۰/۵ سانتی‌متر باشد و یک نمونه‌ی ۱۰ تایی از این یاتاقان‌ها دارای میانگین ۰/۵۰۶ سانتی‌متر و انحراف معیار ۰/۰۰۴ باشند مقدار $T=\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ را حساب کنید.

پاسخ: با توجه به اطلاعات تمرین داریم:

$$\mu=0/5 \quad , \quad n=10 \quad , \quad \bar{X}=0/506 \quad , \quad S=0/004$$

$$T=\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}=\frac{0/506-0/5}{\frac{0/004}{\sqrt{10}}}=\frac{0/006}{0/001}=6$$

۱۶ـ مقادیر زیر را پیدا کنید.

الف) $F_{0/05}$ با درجه‌ی آزادی ۷ و ۳

ب) $F_{0/05}$ با درجه‌ی آزادی ۵ و ∞

ج) $F_{0/01}$ با درجه‌ی آزادی ۳۰ و ۱۲۰

پاسخ: الف) $P\left[F_{(7,3)}>c\right]=0/05\to c=8/8868$

ب) $P\left[F_{(5,\infty)}>c\right]=0/05\to c=2/2141$

ج) $P\left[F_{(30,120)}>c\right]=0/01\to c=1/8600$

۱۷ـ یک تست کنکور استاندارد در زمینه‌ی ریاضی به ۲۵ پسر و ۱۶ دختر داده شده است. میانگین نمره‌ی پسرها ۸۲ و انحراف ۸ و حال آن که میانگین نمره‌ی دخترها ۷۸ و انحراف معیار ۷ شده است. اگر نسبت واریانس نمرات پسرها در جامعه نسبت به دخترها $\frac{۱}{۲}$ باشد. مقدار آماره‌ی F را به دست آورید.

پاسخ: واریانس نمرات پسرها را با $S_1^2 = ۶۴$ و واریانس نمرات دخترها را با $S_2^2 = ۴۹$ نشان می‌دهیم می‌دانیم $n_1 = ۲۵$ و $n_2 = ۱۶$ می‌باشد، همچنین می‌دانیم نسبت واریانس نمرات پسرها در جامعه نسبت به دخترها $\left[\frac{\sigma_1}{\sigma_2}\right]^2 = \frac{۱}{۲}$ می‌باشد پس داریم:

$$\left[\frac{\sigma_2}{\sigma_1}\right]^2 = ۲$$

بنابراین:

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \left[\frac{\sigma_2}{\sigma_1}\right]^2 \frac{S_1^2}{S_2^2}$$

$$F = ۲ \times \frac{۶۴}{۴۹} = \frac{۱۲۸}{۴۹} = ۲/۶۱۲$$

خودآزمایی فصل ششم

۱- برای توابع چگالی احتمال زیر فضای نمونه را بنویسید.

الف) $f(x)=\theta(1-\theta)^x \qquad x=0,1,2,\ldots \quad , \quad 0<\theta<1$

ب) $f(x)=\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} \qquad 0<x<1 \qquad 0<\alpha<1 \qquad 0<\beta<1$

ج) $f(x)=\dfrac{1}{\theta-7} \qquad 7<x<\theta$

پاسخ: الف) اگر تعداد شکست‌ها را برای رسیدن به اولین موفقیت در نظر بگیریم، داریم:

$$S=\{\underset{x=0}{\underset{\downarrow}{p}}\ ,\ \underset{x=1}{\underset{\downarrow}{pq}}\ ,\ \underset{x=2}{\underset{\downarrow}{pqq}}\ ,\ \underset{x=3}{\underset{\downarrow}{pqqq}}\ ,\ \underset{x=4}{\underset{\downarrow}{pqqqq}}\ ,\ \ldots\}$$

ب) توزیع مربوط به توزیع بتاست و بین صفر و یک قابل تغییر می‌باشد. اگر شکل آن را رسم کنیم متوجه خواهیم شد فضای نمونه بین صفر و بیش‌ترین مقدار قابل تغییر است.

ج) فضای نمونه در $(0, +\infty)$ قابل تغییر می‌باشد زیرا:

$7<x<\theta$

اگر $\theta \to 7$ آن گاه $f(x)\to+\infty$

اگر $\theta \to +\infty$ آن گاه $f(x)\to 0$

۲- در یک ایستگاه هواشناسی ده روز در فصل بارندگی، روزها از نظر بارانی بودن یا نبودن به صورت زیر ثبت شده‌اند. اگر روزهای بارانی را از هم مستقل فرض کنیم، چه قالب احتمالی می‌توان ارائه داد؟ پارامتر مدل ارائه شده را برآورد کنید.

۱ ۰ ۱ ۰ ۰ ۱ ۱ ۰ ۱ ۰

پاسخ: متغیر تصادفی X دارای تابع احتمال برنولی با پارامتر p است و تابع چگالی احتمال آن به صورت زیر است:

$$f(x)=p^x(1-p)^{1-x} \qquad x=0,1$$

که می‌توان به صورت زیر نیز نوشت:

$$f(x)=\begin{cases} p & x=1 \quad \text{بارانی بودن} \\ (1-p) & x=0 \quad \text{بارانی نبودن} \end{cases}$$

برای برآورد پارامتر به روش گشتاورها داریم:

$$\mu_1 = E(X) = p$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$$

از برابری دو رابطه‌ی اخیر داریم:

$$\hat{p} = \bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}\times 5 = 0/5$$

۳ـ برآوردگرهای گشتاورها را بر اساس یک نمونه‌ی تصادفی $X_1, X_2, \ldots, X_n$ برای هر یک از توابع چگالی احتمال زیر پیدا کنید.

$$f(x) = \theta x^{\theta-1} \qquad 0 < x < 1$$

$$f(x) = (\theta+1)x^{-(\theta+2)} \qquad x > 1$$

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{1-\beta} \qquad 0 < x < 1$$

$$f(x) = \theta(1-\theta)^{x-1} \qquad x = 1, 2, 3, \ldots$$

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

پاسخ: اگر $x_1, x_2, \ldots, x_n$ مشاهدات نمونه تصادفی $X_1, X_2, \ldots, X_n$ باشند داریم:

$$* f(x) = \theta x^{\theta-1} \qquad 0 < x < 1$$

$$\mu_1 = E(X) = \int_0^1 x \cdot \theta x^{\theta-1} dx = \theta \int_0^1 x^{\theta} dx = \theta \cdot \left[\frac{x^{\theta+1}}{\theta+1}\right]_0^1 = \frac{\theta}{\theta+1}$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

از برابری دو رابطه‌ی اخیر داریم:

$$\frac{\hat\theta}{\hat\theta+1} = \bar{x} \rightarrow (\hat\theta+1)\bar{x} = \hat\theta \rightarrow \hat\theta\bar{x} + \bar{x} = \hat\theta$$

$$\rightarrow \bar{x} = \hat\theta - \hat\theta\bar{x} \rightarrow \bar{x} = \hat\theta(1-\bar{x}) \rightarrow \hat\theta = \frac{\bar{x}}{1-\bar{x}}$$

$$* f(x) = (\theta+1)x^{-(\theta+2)} \qquad x > 1$$

$$\mu_1 = E(X) = \int_1^{\infty} x \cdot (\theta+1)x^{-(\theta+2)} dx = \int_1^{\infty} (\theta+1)x^{-\theta-1} dx$$

$$= (\theta+1)\int_1^{\infty} x^{-\theta-1} dx = (\theta+1)\left[\frac{1}{-\theta x^{\theta}}\right]_1^{\infty} = (\theta+1)\left[0 - \frac{1}{-\theta}\right] = \frac{\theta+1}{\theta}$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

از برابری دو رابطه‌ی اخیر داریم:

$$\frac{\hat\theta+1}{\hat\theta} = \bar{x} \rightarrow \hat\theta + 1 = \hat\theta\bar{x} \rightarrow 1 = \hat\theta\bar{x} - \hat\theta$$

$$\rightarrow 1 = \hat\theta(\bar{x}-1) \rightarrow \hat\theta = \frac{1}{\bar{x}-1}$$

$$* f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{1-\beta} \qquad 0 < x < 1$$

اگر در صورت مسئله توان $(1-x)$، $\beta-1$ می‌باشد. پس تابع چگالی بتا به صورت زیر است:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad 0 < x < 1$$

$$\mu_1 = E(X) = \frac{\alpha}{\alpha+\beta}$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

چون مقدار دو پارامتر α و β را می‌خواهیم برآورد کنیم پس باید μ_2 و m'_2 را نیز به دست آوریم:

$$\mu_2 = E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$m'_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \overline{x^2}$$

از حل معادلات $\begin{cases}\mu_1 = m'_1 \\ \mu_2 = m'_2\end{cases}$ برآوردگرهای گشتاوری را به دست می‌آوریم:

رابطه‌ی (*)

$$\begin{cases} \bar{x} = \dfrac{\alpha}{\alpha+\beta} & \rightarrow (\alpha+\beta)\bar{x} = \alpha \rightarrow \beta = \dfrac{\alpha(1-\bar{x})}{\bar{x}} \\ \overline{x^2} = \dfrac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} & \rightarrow (\alpha+\beta)(\alpha+\beta+1)\overline{x^2} = \alpha(\alpha+1) \end{cases}$$

در رابطه‌ی دوم به جای $(\alpha+\beta) = \frac{\alpha}{\bar{x}}$ و ساده می‌کنیم پس داریم:

$$\left[\frac{\alpha}{\bar{x}}\right]\left[\frac{\alpha}{\bar{x}}+1\right]\overline{x^2} = \alpha(\alpha+1) \rightarrow \alpha(\overline{x^2} - \bar{x}^2) = (\bar{x}^2 - \bar{x}\,\overline{x^2})$$

بنابراین داریم:

$$\hat\alpha = \frac{\bar{x}^2 - \bar{x}\,\overline{x^2}}{\overline{x^2} - \bar{x}^2}$$

با جایگزین کردن $\hat\alpha$ در رابطه‌ی * مقدار β نیز به دست می‌آید:

$$\hat\beta = \frac{(\bar{x} - \overline{x^2} - \bar{x}^2 + \bar{x}\,\overline{x^2})}{(\overline{x^2} - \bar{x}^2)}$$

$$* f(x) = \theta(1-\theta)^{x-1} \qquad x = 1, 2, 3, \ldots$$

تابع چگالی احتمال داده شده مربوط به شکل دیگری از توزیع هندسی است. قبلاً در فصل ۴ خواندیم که تابع احتمال هندسی به صورت زیر است:

$$f(x)=p(1-p)^{x} \qquad x=0,1,2,\ldots$$

که $E(X)=\frac{1-p}{p}$ و $V(X)=\frac{1-p}{p^{2}}$ است.

ولی برای این شکل تابع چگالی مقدار امید ریاضی به صورت زیر است: (پارامتر را θ در نظر می‌گیریم)

$$\mu_1=E(X)=\frac{1}{\theta}$$

$$m'_1=\frac{1}{n}\sum_{i=1}^{n} x_i=\bar{x}$$

از برابری دو رابطه‌ی بالا داریم:

$$\frac{1}{\theta}=\bar{x}\rightarrow\hat{\theta}=\frac{1}{\bar{x}}$$

$$*f(x)=\frac{1}{b-a} \qquad a<x<b$$

تابع چگالی احتمال داده شده مربوط به توزیع یکنواخت پیوسته می‌باشد که برای این تابع احتمال داریم:

$$\mu_1=E(X)=\int_a^b x.\frac{1}{b-a}\,dx=\frac{1}{b-a}\int_a^b x\,dx=\frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b=\frac{b^2-a^2}{2(b-a)}$$

$$=\frac{(b-a)(b+a)}{2(b-a)}=\frac{a+b}{2}$$

$$m'_1=\frac{1}{n}\sum_{i=1}^{n} x_i=\bar{x}$$

چون مقدار دو پارامتر a و b را می‌خواهیم برآورد کنیم باید μ_2 و m'_2 را نیز به دست آوریم:

$$\mu_2=E(X^2)=\int_a^b x^2.\frac{1}{b-a}\,dx=\frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b=\frac{b^3-a^3}{3(b-a)}$$

$$=\frac{(b-a)(a^2+b^2+ab)}{3(b-a)}=\frac{(a^2+b^2+ab)}{3}$$

$$m'_2=\frac{1}{n}\sum_{i=1}^{n} x_i^2=\overline{x^2}$$

از حل معادلات $\begin{cases}\mu_1=m'_1\\ \mu_2=m'_2\end{cases}$ برآوردگرهای گشتاوری را به دست می‌آوریم:

$$\begin{cases}a+b=2\bar{x}\\ a^2+b^2+ab=3\overline{x^2}\end{cases}\Rightarrow\begin{cases}\hat{a}=\bar{x}+\sqrt{3(\overline{x^2}-\bar{x}^2)}\\ \hat{b}=\bar{X}-\sqrt{3(\overline{x^2}-\bar{x}^2)}\end{cases}$$

رابطه‌ی (*)

۴ـ فرض کنید برای پارامتر θ در تابع احتمال سری لگاریتمی مشاهدات به صورت جدول فراوانی زیر داده شده باشد، اگر $\bar{X}$ یک براورد نقطه‌ای برای θ باشد، θ را براورد کنید.

x	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
f	۱۰۶۲	۲۶۳	۱۲۰	۵۰	۲۲	۷	۶	۲	۰	۱	۱

پاسخ: تابع احتمال سری لگاریتمی به صورت زیر می‌باشد:

$$f(x)=\frac{-1}{\operatorname{Ln}(1-\alpha)}\;\frac{\alpha^{x}}{x} \qquad x=1,2,\ldots$$

چون $\bar{X}$ را براورد نقطه‌ای پارامتر θ در نظر گرفته‌ایم پس داریم:

$$\hat{\theta}=\bar{X}=\frac{\sum_{i=1}^{k} f_i x_i}{n}$$

$$\hat{\theta}=\bar{X}=\frac{(1\times 1062)+(2\times 263)+\ldots+(11\times 1)}{1534}$$

$$=\frac{1062+526+360+200+110+42+42+16+0+10+11}{1534}$$

$$=\frac{2379}{1534}=1/551$$

بنابراین:

$$\hat{\theta}=\bar{X}=1/551$$

۵ـ تعداد ساعاتی که یک لامپ الکترونی کار می‌کند یک متغیر تصادفی نمایی با پارامتر θ است. θ را بر اساس نمونه‌ی تصادفی $X_1, X_2, \ldots, X_n$ به روش گشتاورها براورد کنید.

پاسخ: متغیر تصادفی X دارای تابع چگالی احتمال نمایی با پارامتر θ به صورت زیر است:

$$f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}} \qquad x>0\;,\;\theta>0$$

که این توزیع دارای میانگین θ و واریانس θ^2 می‌باشد.

حال برای براورد پارامتر θ داریم:

$$\mu_1=E(X)=\theta$$

$$m'_1=\frac{1}{n}\sum_{i=1}^{n} X_i=\bar{X}$$

حال از تساوی قرار دادن رابطه‌های بالا داریم:

$$\mu_1=m'_1\rightarrow\hat{\theta}=\bar{X}$$

۶ـ اگر متغیر تصادفی X روی بازه‌ی $(\theta-1 , \theta+2)$ تعریف شده باشد. θ را به روش گشتاورها برآورد کنید.

پاسخ: متغیر تصادفی X دارای توزیع یکنواخت می‌باشد که در آن $a=\theta-1$ و $b=\theta+2$ است. برای توزیع یکنواخت داریم:

$$\mu_1=E(X)=\frac{a+b}{2}$$

که با در نظر گرفتن $a=\theta-1$ و $b=\theta+2$ داریم:

$$\mu_1=E(X)=\frac{\theta-1+\theta+2}{2}=\frac{2\theta+1}{2}$$

همچنین داریم:

$$m'_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\bar{X}$$

حال با برابر قرار دادن رابطه‌های بالا داریم:

$$\mu_1=m'_1\to\frac{2\theta+1}{2}=\bar{X}\to 2\theta+1=2\bar{X}\to 2\theta=2\bar{X}-1\to\hat{\theta}=\frac{2\bar{X}-1}{2}$$

۷ـ برآوردگرهای درستنمایی ماکزیمم را براساس یک نمونه‌ی تصادفی X_1، X_2، ...، X_n برای هر یک از توابع چگالی احتمال زیر پیدا کنید.

$$f(x)=\theta e^{-\theta x}\qquad x>0$$

$$f(x)=\begin{bmatrix}x+r-1\\r-1\end{bmatrix}\theta^r(1-\theta)^x\qquad x=0,1,2,\dots$$

$$f(x)=\frac{1}{\Gamma(n)}x^{n-1}e^{-x}\qquad x>0$$

$$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{1-\beta}\qquad 0<x<1$$

$$f(x)=\beta x^{\beta-1}e^{-x^{\beta}}\qquad x>0$$

پاسخ: اگر $x_1 , x_2 , \dots , x_n$ مشاهدات نمونه تصادفی $X_1 , X_2 , \dots , X_n$ باشند، داریم:

* $f(x)=\theta e^{-\theta x}\qquad x>0$

$$L(\theta)=\prod_{i=1}^{n}f_\theta(x_i)=\prod_{i=1}^{n}\theta e^{-\theta x_i}=\theta^n e^{-\theta\sum_{i=1}^{n}x_i}$$

$$\operatorname{Ln}L(\theta)=n\ln\theta-\theta\sum_{i=1}^{n}x_i$$

$$\frac{d\operatorname{Ln}L(\theta)}{d\theta}=\frac{n}{\theta}-\sum_{i=1}^{n}x_i=0\to\frac{n-\theta\sum_{i=1}^{n}x_i}{\theta}=0$$

صورت کسر بالا را برابر صفر قرار می‌دهیم:

$$n-\theta\sum_{i=1}^{n}x_i=0$$

$$\to n=\theta\sum_{i=1}^{n}x_i\to\hat{\theta}=\frac{n}{\sum_{i=1}^{n}x_i}$$

که $\hat{\theta}$ عکس رابطه‌ی $\bar{X}$ می‌باشد پس داریم:

$$\hat{\theta}=\frac{n}{\sum_{i=1}^{n}x_i}=\frac{1}{\bar{X}}$$

* $f(x)=\begin{bmatrix}x+r-1\\r-1\end{bmatrix}\theta^r(1-\theta)^x\qquad x=0,1,2,\dots$

تابع احتمال داده شده مربوط به توزیع دوجمله‌ای منفی با پارامترهای r و θ می‌باشد.

$$L(\theta)=\prod_{i=1}^{n}f_\theta(x_i)=\prod_{i=1}^{n}\begin{bmatrix}x+r-1\\r-1\end{bmatrix}\theta^r(1-\theta)^x=\begin{bmatrix}x+r-1\\r-1\end{bmatrix}\theta^{nr}.(1-\theta)^{\sum_{i=1}^{n}x_i}$$

$$\operatorname{Ln}L(\theta)=\operatorname{Ln}\begin{bmatrix}x+r-1\\r-1\end{bmatrix}+nr\ln\theta+\sum_{i=1}^{n}x_i\operatorname{Ln}(1-\theta)$$

$$\frac{d\operatorname{Ln}L(\theta)}{d\theta}=\frac{nr}{\theta}-\frac{\sum_{i=1}^{n}x_i}{1-\theta}=0$$

$$\Rightarrow\frac{nr}{\theta}=\frac{\sum_{i=1}^{n}x_i}{1-\theta}\to nr(1-\theta)=\theta\sum_{i=1}^{n}x_i\to nr-nr\theta=\theta\sum_{i=1}^{n}x_i$$

$$\to nr=\theta\sum_{i=1}^{n}x_i+nr\theta\to nr=\theta\left[\sum_{i=1}^{n}x_i+nr\right]\to\hat{\theta}=\frac{nr}{\sum_{i=1}^{n}x_i+nr}$$

$$\to\hat{\theta}=\frac{nr}{n\bar{x}+nr}=\frac{nr}{n(\bar{x}+r)}=\frac{r}{\bar{x}+r}$$

* $f(x)=\frac{1}{\Gamma(n)}x^{n-1}e^{-x}\qquad x>0$

با مقایسه‌ی تابع چگالی احتمال بالا با تابع گاما به این نتیجه می‌رسیم که:

$$f(x)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}\qquad x>0$$ (تابع گاما با پارامترهای α و β)

تابع چگالی بالا همان توزیع گاما است که در آن $\alpha=n$ و $\beta=1$ می‌باشد حال می‌خواهیم مقدار $\alpha=n$ را برآورد کنیم. (پارامتر توزیع n می‌باشد)

$$L(n)=\prod_{i=1}^{n}f(x_i)=\prod_{i=1}^{n}\frac{1}{\Gamma(n)}x^{n-1}e^{-x}=\left[\Gamma(n)\right]^{-n}\left(\prod_{i=1}^{n}x_i^{n-1}\right)e^{-\sum_{i=1}^{n}x_i}$$

$$\operatorname{Ln}L(n)=-n\operatorname{Ln}\Gamma(n)+(n-1)\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}x_i$$

برای معادله‌ی تابع درستنمایی داریم:

$$\frac{d\,\mathrm{Ln}\,L(n)}{dn} = -n\frac{\Gamma'(n)}{\Gamma(n)} + \sum_{i=1}^{n} \mathrm{Ln}\,x_i = ۰$$

یعنی:

$$\frac{\Gamma'(n)}{\Gamma(n)} = \frac{۱}{n}\sum_{i=1}^{n} \mathrm{Ln}\,x_i$$

برای حل معادله‌ی فوق به جداول مربوط به $\frac{\Gamma'(n)}{\Gamma(n)}$ یا روش‌های آنالیز عددی نیاز داریم.

* $f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-۱}(۱-x)^{۱-\beta} \qquad ۰<x<۱$

تذکر: در صورت سوال توان $(۱-x)$، $\beta-۱$ صحیح می‌باشد.

$$L(\theta)=\prod_{i=1}^{n} f(x_i)=\prod_{i=1}^{n}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-۱}(۱-x)^{\beta-۱}$$

$$=\left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right]^n \prod_{i=1}^{n}(x_i)^{\alpha-۱}(۱-x_i)^{\beta-۱}$$

$$\mathrm{Ln}\,L(\theta)=n\,\mathrm{Ln}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}+(\alpha-۱)\sum_{i=1}^{n}\mathrm{Ln}(x_i)+(\beta-۱)\sum_{i=1}^{n}\mathrm{Ln}(۱-x_i)$$

در ادامه کافی است یک بار نسبت به α با ثابت گرفتن مقدار β مشتق گرفته و مساوی صفر قرار دهیم، یعنی:

$$\frac{d\,\mathrm{Ln}\,L(\theta)}{d\alpha}=۰$$

و یک بار با ثابت گرفتن مقدار α نسبت به β مشتق گرفته و برابر صفر قرار دهیم یعنی:

$$\frac{d\,\mathrm{Ln}\,L(\theta)}{d\beta}=۰$$

بدین ترتیب مقادیر $\hat{\alpha}$ و $\hat{\beta}$ به دست می‌آید.

نکته: $L(\theta)$ همان $L(\alpha,\beta)$ می‌باشد که به طور خلاصه با $L(\theta)$ نشان می‌دهیم.

* $f(x)=\beta x^{\beta-۱}e^{-x^{\beta}} \qquad x>۰$

اگر متغیر تصادفی X دارای توزیع وایبل با پارامترهای α و λ باشد تابع چگالی احتمال آن به صورت زیر است:

$$f_{\alpha,\lambda}(x)=\alpha\lambda x^{\alpha-۱}e^{-\lambda x^{\alpha}} \qquad x>۰\ ,\ \alpha>۰\ ,\ \lambda>۰$$

با مقایسه‌ی تابع چگالی احتمال داده شده با توزیع وایبل به این نتیجه می‌رسیم که تابع چگالی داده شده همان تابع چگالی وایبل می‌باشد که در آن $\alpha=\beta$ و $\lambda=۱$ می‌باشد حال برای به دست آوردن برآوردگر درستنمایی ماکزیمم بر اساس پارامتر β داریم:

$$L(\beta)=\prod_{i=1}^{n} f(x_i)=\prod_{i=1}^{n}\beta x^{\beta-۱}e^{-x^{\beta}}=\beta^n\Big(\prod_{i=1}^{n}x_i^{\beta-۱}\Big)e^{-\beta\sum_{i=1}^{n}x_i}$$

$$\mathrm{Ln}\,L(\beta)=n\,\mathrm{Ln}\beta+(\beta-۱)\sum_{i=1}^{n}\mathrm{Ln}\,x_i-\beta\sum_{i=1}^{n}x_i$$

$$\frac{d\,\mathrm{Ln}\,L(\beta)}{d\beta}=\frac{n}{\beta}+\sum_{i=1}^{n}\mathrm{Ln}\,x_i-\sum_{i=1}^{n}x_i=۰$$

$$\frac{n}{\beta}=-\sum_{i=1}^{n}\mathrm{Ln}\,x_i+\sum_{i=1}^{n}x_i\rightarrow n=\beta\Big(\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\mathrm{Ln}\,x_i\Big)$$

$$\hat{\beta}=\frac{n}{\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\mathrm{Ln}\,x_i}\Rightarrow\hat{\beta}=\frac{۱}{\bar{x}}-\frac{n}{\sum_{i=1}^{n}\mathrm{Ln}\,x_i}$$

نکته: با توجه به خواص توان‌ها داریم: $(a^b)^c=a^{bc}$

۸ـ برای پارامتر θ فواصل اطمینان زیر پیشنهاد شده است. کدام یک بهترین فاصله‌ی اطمینان برای θ است؟

$(-۱\,,\,۲)\quad,\quad(۰/۵\,,\,۱)\quad,\quad(۰\,,\,۲)\quad,\quad(۰/۷۵\,,\,۱/۲۵)$

پاسخ: ابتدا درک این نکته ضروری است که فاصله‌های اطمینان پارامتری مفروض یکتا نیستند و همچنین برای به دست آوردن فاصله‌های اطمینان، باید بنا بر خواص آماری مختلف آن‌ها داوری شود. به عنوان مثال یک خاصیت مطلوب آن است که طول فاصله‌ی اطمینان $(۱-\alpha)$۱۰۰٪ تا سر حدامکان کوچک گرفته شود، خاصیت مطلوب دیگر آن است که امید این طول $E(\hat{\theta}_۲-\hat{\theta}_۱)$ تا سر حد امکان کوچک باشد. پس بنابر نکاتی که گفته شد بین فواصل اطمینان داده شده، بازه‌های اطمینان $(۰/۵\,,\,۱)$ و $(۰/۷۵\,,\,۱/۲۵)$ به دو مورد دیگر ترجیح داده می‌شود.

۹ـ یک مطالعه‌ی الودگی هوا در یک ایستگاه آزمایشی انجام شده است. مقادیر مواد آلی محلول در بنزین که در هوا معلق می‌باشد در ۸ نمونه مختلف هوا به صورت ۲/۲، ۱/۸، ۳/۱، ۲/۰، ۲/۴، ۲/۵، ۲/۱، ۱/۲ بوده است. فرض کنید جامعه‌ای که از آن نمونه گرفته شده نرمال باشد. یک فاصله اطمینان ۰/۹۵ برای میانگین واقعی به دست آورید.

پاسخ:

$$n=۸\quad,\quad ۱-\alpha=۰/۹۵\quad,\quad\alpha=۰/۰۵$$

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{۱۷/۳}{۸}=۲/۱۶$$

$$S^۲=\frac{۱}{n-۱}\sum_{i=1}^{n}(x_i-\bar{x})^۲=۰/۳۰\qquad,\qquad S=۰/۵۵$$

$$\frac{\alpha}{۲}=۰/۰۲۵\qquad t(۰/۰۲۵\,,\,۷)=۲/۳۶۵$$

$$\bar{x}-t_{(\frac{\alpha}{۲},n-۱)}\frac{S}{\sqrt{n}}<\mu<\bar{x}+t_{(\frac{\alpha}{۲},n-۱)}\frac{S}{\sqrt{n}}$$

$$۲/۱۶-۲/۳۶۵\frac{۰/۵۵}{\sqrt{۸}}<\mu<۲/۱۶+۲/۳۶۵\frac{۰/۵۵}{\sqrt{۸}}$$

$$۱/۷۰<\mu<۲/۶۲$$

۱۰ـ فرض کنید که ۱۲ پرتقال شهسواری بخرید، آن‌ها را وزن کنید و به دست آورید.

$$\sum_{i=1}^{12} x_i = 1.98 \quad , \quad \sum_{i=1}^{12} x_i^2 = 0.345$$

اگر وزن پرتقال‌ها (بر حسب کیلوگرم) دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشند یک فاصله‌ی اطمینان ۰/۹۵ درصد برای μ و σ^2 به دست آورید.

پاسخ:

$$\bar{x} = \frac{\sum_{i=1}^{12} x_i}{12} = \frac{1.98}{12} = 0.165$$

$$(n-1)S^2 = \sum_{i=1}^{12} x_i - n\bar{x}^2 = 0.345 - 12(0.165)^2 = 0.018$$

$$S^2 = \frac{0.018}{11} = 0.002 \quad , \quad S = 0.045$$

فاصله اطمینان برای μ

$$n = 12 \quad , \quad \bar{x} = 0.165 \quad , \quad S = 0.045$$

$$1-\alpha = 0.95 \ , \ \alpha = 0.05 \ , \ \frac{\alpha}{2} = 0.025 \ , \ t(0.025, 11) = 2.201$$

$$\bar{x} - t_{(\frac{\alpha}{2}, n-1)} \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{(\frac{\alpha}{2}, n-1)} \frac{S}{\sqrt{n}}$$

$$0.165 - 2.201 \frac{0.045}{\sqrt{12}} < \mu < 0.165 + 2.201 \frac{0.045}{\sqrt{12}}$$

$$0.136 < \mu < 0.194$$

فاصله اطمینان برای σ^2

$$(n-1)S^2 = 0.0180$$

$$\chi^2_{(\frac{\alpha}{2}, n-1)} = \chi^2_{(0.025, 11)} = 21.920$$

$$\chi^2_{(\frac{\alpha}{2}, n-1)} = \chi^2_{(0.975, 11)} = 3.816$$

$$\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2}, n-1)}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2}, n-1)}}$$

$$\frac{0.0180}{21.920} \le \sigma^2 \le \frac{0.018}{3.816}$$

$$0.001 \le \sigma^2 \le 0.005$$

۱۱ـ استادی یک امتحان نهایی برای درس مرکز کنترل اجرا و طراحی رایانه برای گروه بزرگی از دانشجویان در نظر گرفته است. تصور او راجع به نمره‌ی آن‌ها به طور ذهنی یک توزیع نرمال $\mu = 67.2$ و $\sigma = 1.5$ می‌باشد.

الف) چه احتمال پیشین اختصاص دهد تا نمره‌ی متوسط ۲۵ نفر در فاصله‌ی ۶۵ تا ۷۰ قرار گیرد.

ب) اگر امتحان در یک گروه ۴۰ نفری از دانشجویان به عنوان یک نمونه‌ی تصادفی انجام شود و میانگین و انحراف معیار نمره‌ها به ترتیب ۷۴/۹ و ۷/۴ باشد مقادیر μ و σ^2 را برآورد کنید و خطای حاصل برآوردها را به دست آورید.

پاسخ: الف)

$$\mu = 67.2 \quad , \quad \sigma = 1.5 \quad , \quad n = 25$$

$$\bar{X} \sim N(67.2 , \frac{1.5}{\sqrt{25}}) \to \bar{X} \sim N(67.2 , 0.3)$$

$$P(65 < \bar{X} < 70) = P\left(\frac{65-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{70-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= P\left(\frac{65-67.2}{0.3} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{70-67.2}{0.3}\right) = P(-7.33 < Z < 9.33)$$

$$= P(Z < 9.33) - P(Z < -7.33) = P(Z < 9.33) - [1 - P(Z < 7.33)]$$

$$= P(Z < 9.33) + P(Z < 7.33) - 1$$

ب) فاصله‌ی اطمینان برای μ

چون مقدار α را مشخص نکرده است بنا به قرار داد $\alpha = 0.05$ در نظر گرفته و یک فاصله‌ی اطمینان ۹۵ درصد برای μ و σ^2 به دست می‌آوریم.

نکته: برای تعیین فاصله‌ی اطمینان برای μ وقتی σ^2 مجهول است دو حالت در نظر می‌گیریم:

حالت اول: حجم نمونه کوچک است $(n<30)$

فاصله‌ی اطمینان $100(1-\alpha)$ درصد برای μ در چنین حالتی عبارت است از:

$$\bar{X} - t_{(\frac{\alpha}{2}, n-1)} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{(\frac{\alpha}{2}, n-1)} \frac{S}{\sqrt{n}}$$

حالت دوم: حجم نمونه زیاد است $(n \ge 30)$

چون به ازای $n \ge 30$ توزیع T تقریباً با توزیع نرمال استاندارد برابر است پس خواهیم داشت:

$$T = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim N(0,1) \qquad n \ge 30$$

پس فاصله‌ی اطمینان $100(1-\alpha)$ درصد برای μ در حالتی که $n \ge 30$ است:

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

با توجه به نکته‌ی گفته شده چون $n \geq 30$ می‌باشد پس داریم:

$$n = 40 \quad , \quad \bar{x} = 74/9 \quad , \quad S = 7/4$$

$$\alpha = 0/05 \quad , \quad 1-\alpha = 0/95 \quad , \quad \frac{\alpha}{2} = 0/025 \quad , \quad Z_{\frac{\alpha}{2}} = Z_{0/025} = 1/96$$

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{x} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$74/9 - 1/96 \frac{7/4}{\sqrt{40}} < \mu < 74/9 + 1/96 \frac{7/4}{\sqrt{40}}$$

$$72/61 < \mu < 77/19$$

فاصله‌ی اطمینان برای σ^2

$$\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2}, n-1)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2}, n-1)}}$$

چون مقدار χ^2_{39} را در جدول توزیع کی دو نداریم به جای آن به طور تقریب χ^2_{40} را محاسبه می‌کنیم.

$$\chi^2_{(0/025, 39)} \simeq \chi^2_{(0/025, 40)} = 59/3$$

$$\chi^2_{(0/975, 39)} \simeq \chi^2_{(0/975, 40)} = 24/4$$

$$\frac{39(7/4)^2}{59/36} \leq \sigma^2 \leq \frac{39(7/4)^2}{24/4}$$

$$36/01 \leq \sigma^2 \leq 87/53$$

و فاصله‌ی اطمینان برای σ برابر است با:

$$6/00 \leq \sigma^2 \leq 9/36$$

نکته: تعیین خطا در برآورد میانگین

در برآورد نقطه‌ای میانگین می‌دانیم غالباً $\bar{X}$ با μ برابر نیست، چرا که با نمونه‌گیری‌های مختلف مقادیر مختلفی برای $\bar{X}$ خواهیم داشت، لذا برآورد نقطه‌ای دارای خطا است. می‌خواهیم میزان این خطا را با استفاده از حدود فاصله‌ی اطمینان تعیین کنیم. یعنی کران‌هایی را برای این خطا یعنی $|\bar{X} - \mu|$ در حالت‌های مختلف به دست آوریم.

الف) اگر واریانس جامعه σ^2 معلوم باشد، آن‌گاه:

$$|\bar{X} - \mu| < \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}}$$

ب) اگر واریانس جامعه σ^2 مجهول باشد آن‌گاه:

۱ـ $n \geq 30$

$$|\bar{X} - \mu| < \frac{S}{\sqrt{n}} Z_{\frac{\alpha}{2}}$$

۲ـ $n < 30$

$$|\bar{X} - \mu| < \frac{S}{\sqrt{n}} t_{(\frac{\alpha}{2}, n-1)}$$

بنابراین برای برآورد خطا حاصل از میانگین با توجه به این که σ^2 مجهول و $n \geq 30$ است داریم:

$$|\bar{X} - \mu| < \frac{S}{\sqrt{n}} Z_{\frac{\alpha}{2}} \rightarrow |\bar{X} - \mu| < \frac{7/4}{\sqrt{40}} \times 1/96 = 2/29$$

پس خطای حاصل از میانگین ۲/۲۹ می‌باشد.

۱۲ـ فرض کنید متغیر Y دارای توزیع دو جمله‌ای با پارامتر P و $n=300$ باشد. اگر $y=75$ مقدار مشاهده برای تغییر Y باشد یک فاصله‌ی اطمینان ۹۰ درصد برای P به دست آورید.

پاسخ:

$$n = 300 \quad , \quad y = 75 \quad , \quad \hat{P} = \frac{y}{n} = \frac{75}{300} = 0/25 \quad , \quad \alpha = 0/1$$

$$\frac{\alpha}{2} = 0/05 \quad , \quad Z_{\frac{\alpha}{2}} = Z_{0/05} = 1/645$$

$$\frac{y}{n} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\frac{y}{n}(1-\frac{y}{n})}{n}} < P < \frac{y}{n} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\frac{y}{n}(1-\frac{y}{n})}{n}}$$

$$\frac{75}{300} - 1/645 \sqrt{\frac{(0/25)(0/75)}{300}} < P < \frac{75}{300} + 1/645 \sqrt{\frac{(0/25)(0/75)}{300}}$$

$$0/2097 < P < 0/2903$$

۱۳ـ اگر $\bar{X}$ و $\bar{Y}$ میانگین‌های دو نمونه‌ی تصادفی n تایی به ترتیب از توزیع $N(\mu_1, \sigma^2)$ و $N(\mu_2, \sigma^2)$ باشند n را طوری پیدا کنید که:

$$P\left[\bar{X} - \bar{Y} - \frac{\sigma}{2} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + \frac{\sigma}{2}\right] = 0/9$$

پاسخ: اگر دو جامعه دارای واریانس مشترک σ^2 و حجم نمونه‌ها مساوی n باشند فاصله‌ی اطمینان برای $\mu_1 - \mu_2$ برابر است با:

$$(\bar{X} - \bar{Y}) - Z_{\frac{\alpha}{2}} \sqrt{\frac{2\sigma^2}{n}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + Z_{\frac{\alpha}{2}} \sqrt{\frac{2\sigma^2}{n}}$$

با توجه به رابطه‌ی داده شده و مقایسه‌ی آن با رابطه‌ی بالا داریم:

$$\frac{\sigma}{2} = Z_{\frac{\alpha}{2}} \sqrt{\frac{2\sigma^2}{n}}$$

از طرفی $1-\alpha = 0/9$ پس داریم:

$$1-\alpha = 0/9 \quad , \quad \alpha = 0/1 \quad , \quad \frac{\alpha}{2} = 0/05 \quad , \quad Z_{\frac{\alpha}{2}} = Z_{0/05} = 1/645$$

بنابراین:

$$\frac{\sigma}{2} = 1/645 \, \sigma \sqrt{\frac{2}{n}} \rightarrow \frac{1}{3/29} = \sqrt{\frac{2}{n}}$$

دو طرف را به توان ۲ می‌رسانیم، داریم:

$$\frac{2}{n} = (0/304)^2 \rightarrow 2 = n(0/092)$$

$$\rightarrow n = \frac{2}{0/092} = 21/739 \simeq 22$$

مقدار $n=22$ در نظر می‌گیریم.

۱۴ـ در یک شرکت پردازش، نمونه‌ای ۱۲ تایی از رایانه‌ها دارای عمر متوسط ۴/۵ سال و در شرکت دیگر عمر متوسط برای نمونه‌ی ۱۵ تایی ۳/۴ سال به دست آمده است. اگر توزیع مشاهدات نرمال با واریانس مشترک ۱ باشد یک فاصله‌ی اطمینان ۹۵ درصد برای $\mu_1-\mu_2$ به دست آورید.

پاسخ:
$$m=12\ ,\quad \bar{x}=4.5\ ,\quad n=15\ ,\quad \bar{y}=3.4$$
$$\sigma_1^2=\sigma_2^2=1\quad \alpha=0.05\ ,\quad Z_{\frac{\alpha}{2}}=Z_{0.025}=1.96$$
$$(\bar{x}-\bar{y})-Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{m}+\frac{\sigma^2}{n}}<\mu_1-\mu_2<(\bar{x}-\bar{y})+Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{m}+\frac{\sigma^2}{n}}$$
$$(4.5-3.4)-1.96\sqrt{\frac{1}{12}+\frac{1}{15}}<\mu_1-\mu_2<(4.5-3.4)+1.96\sqrt{\frac{1}{12}+\frac{1}{15}}$$
$$1.1-0.759<\mu_1-\mu_2<1.1+0.759$$
$$0.341<\mu_1-\mu_2<1.859$$

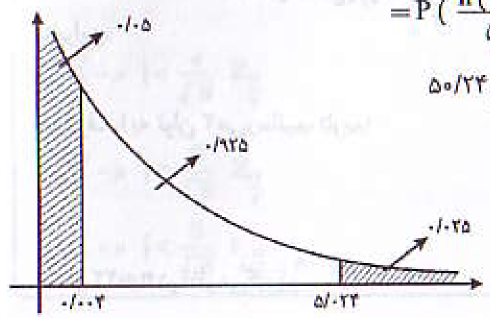
۱۵ـ اگر $\bar{X}$ میانگین یک نمونه‌ی تصادفی از توزیع نرمال با میانگین μ و واریانس σ^2 باشد احتمال این که فاصله‌ی تصادفی $\left[\frac{n(\bar{X}-\mu)^2}{0.004}\ ,\ \frac{n(\bar{X}-\mu)^2}{5.024}\right]$، σ^2 را در برداشته باشد چقدر است؟

پاسخ: می‌خواهیم از رابطه‌ی زیر مقدار احتمال را به دست آوریم داریم:
$$P\left(\frac{n(\bar{X}-\mu)^2}{5.024}<\sigma^2<\frac{n(\bar{X}-\mu)^2}{0.004}\right)=?$$

با توجه به این که $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{n-1}$ می‌باشد و همچنین می‌دانیم اگر $X_1, X_2, \dots, X_n$ یک نمونه‌ی n تایی از توزیع نرمال با میانگین μ و واریانس σ^2 باشد آن گاه داریم:
$$\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2\sim\chi_1^2\rightarrow\frac{n(\bar{X}-\mu)^2}{\sigma^2}\sim\chi_1^2$$

پس می‌خواهیم نشان دهیم که:
$$P\left(0.004<\frac{n(\bar{X}-\mu)^2}{\sigma^2}<5.024\right)=P\left(\frac{1}{0.004}>\frac{\sigma^2}{n(\bar{X}-\mu)^2}>\frac{1}{5.024}\right)$$
$$=P\left(\frac{n(\bar{X}-\mu)^2}{5.024}<\sigma^2<\frac{n(\bar{X}-\mu)^2}{0.004}\right)=0.925$$

در جدول χ^2 با یک درجه‌ی آزادی مقادیر ۰/۰۰۴ و ۵/۰۲۴ را پیدا می‌کنیم. با توجه به شکل داریم:



۱۶ـ زمان لازم برای انجام یک کار به دو روش I و II توسط کاربران رایانه بر حسب ثانیه به صورت زیر به دست آمده است:

I : ۲۰ ۱۶ ۲۶ ۲۷ ۲۳ ۲۲

I : ۲۷ ۳۲ ۴۲ ۳۵ ۳۲ ۳۴ ۳۸

الف) یک فاصله‌ی اطمینان ۹۰ درصد برای واریانس روش اول به دست آورید.

ب) یک فاصله‌ی اطمینان ۹۰ درصد برای نسبت واریانس‌ها به دست آورید.

پاسخ: الف)
$$n=6\ ,\quad \alpha=0.1\ ,\quad 1-\alpha=0.9\ ,\quad \frac{\alpha}{2}=0.05$$
$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2=16.265\ ,\quad S=4.033$$
$$\chi^2_{(\frac{\alpha}{2},n-1)}=\chi^2_{(0.05,5)}=11.070$$
$$\chi^2_{(1-\frac{\alpha}{2},n-1)}=\chi^2_{(0.95,5)}=1.145$$
$$\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},n-1)}}\le\sigma^2\le\frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},n-1)}}$$
$$\frac{5(16.265)}{11.070}\le\sigma^2\le\frac{5(16.265)}{1.145}$$
$$7.346\le\sigma^2\le71.026$$

ب)
$$n_1=6\ ,\quad S_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i-\bar{x})^2=16.265$$
$$n_2=7\ ,\quad S_2^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_i-\bar{y})^2=22.906$$
$$\alpha=0.1\ ,\quad \frac{\alpha}{2}=0.05\ ,\quad 1-\frac{\alpha}{2}=0.95$$
$$f\left(\frac{\alpha}{2},n_1-1,n_2-1\right)=f_{(0.05,5,6)}=4.3874$$
$$f\left(1-\frac{\alpha}{2},n_1-1,n_2-1\right)=f_{(0.95,5,6)}=\frac{1}{f_{(0.05,6,5)}}=\frac{1}{4.9503}=0.202$$
$$\frac{1}{f(\frac{\alpha}{2},n_1-1,n_2-1)}\frac{S_1^2}{S_2^2}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{1}{f(1-\frac{\alpha}{2},n_1-1,n_2-1)}\frac{S_1^2}{S_2^2}$$
$$\frac{1}{4.3874}\cdot\frac{16.265}{22.906}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{1}{0.202}\cdot\frac{16.265}{22.906}$$
$$0.162\le\frac{\sigma_1^2}{\sigma_2^2}\le3.515$$

۱۷ـ اعداد زیر مقادیر سختی به دست آمده برای نمونه‌های دو آلیاژ منیزیم می‌باشد.

آلیاژ I: ۶۶/۳ ۶۳/۵ ۶۴/۹ ۶۱/۸ ۶۴/۳ ۶۴/۷ ۶۵/۱ ۶۴/۵ ۶۸/۴ ۶۳/۲

آلیاژ II: ۷۱/۳ ۶۰/۴ ۶۲/۶ ۶۳/۹ ۶۸/۸ ۷۰/۱ ۶۴/۸ ۶۸/۹ ۶۵/۸ ۶۶/۲

با فرض نرمال بودن داده‌های دو نمونه یک فاصله‌ی اطمینان ۹۵ درصد برای نسبت واریانس‌ها به دست آورید.

پاسخ:

$$n_1 = 10 \quad , \quad S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2 = 3/222$$

$$n_2 = 10 \quad , \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2 = 12/138$$

$$\alpha = 0/05 \quad , \quad \frac{\alpha}{2} = 0/025 \quad , \quad 1 - \frac{\alpha}{2} = 0/975$$

$$f\left(\frac{\alpha}{2} , n_1 - 1 , n_2 - 1\right) = f(0/025 , 9 , 9) = 4/03$$

$$f\left(1 - \frac{\alpha}{2} , n_1 - 1 , n_2 - 1\right) = f(0/975 , 9 , 9) = \frac{1}{f(0/025 , 9 , 9)} = \frac{1}{4/03} = 0/248$$

$$\frac{1}{f\left(\frac{\alpha}{2} , n_1 - 1 , n_2 - 1\right)} \frac{S_1^2}{S_2^2} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{1}{f\left(1 - \frac{\alpha}{2} , n_1 - 1 , n_2 - 1\right)} \frac{S_1^2}{S_2^2}$$

$$\frac{1}{4/03} \frac{3/222}{12/138} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{1}{0/248} \frac{3/222}{12/138}$$

$$0/066 \le \frac{\sigma_1^2}{\sigma_2^2} \le 1/070$$

خودآزمایی فصل هفتم

۱ـ الف) فرق بین فرض ساده و مرکب چیست؟

ب) خطای نوع اول و دوم را تعریف کنید.

پاسخ: الف) هر فرض آماری که توزیع جامعه را کاملاً مشخص کند فرض ساده و فرضی که توزیع جامعه را مشخص نکند فرض مرکب گویند.

ب) در انجام آزمون فرض، اگر فرض آماری درست باشد و ما آن را ردکنیم مرتکب خطای نوع اول و اگر فرض آماری راکه پذیرفتیم نادرست باشد مرتکب خطای نوع دوم می‌شویم.

H_0	درست	نادرست
پذیرفتن	بدون خطا	خطای نوع دوم
رد کردن	خطای نوع اول	بدون خطا

خطای نوع اول و دوم را به زبان آماری به صورت زیر می‌نویسیم:

خطای نوع اول: ردکردن H_0 وقتی که H_0 درست است. $[RH_0 \mid H_0 \text{ درست}]$

خطای نوع دوم: قبول H_0 وقتی که H_0 غلط است. $[AH_0 \mid H_0 \text{ غلط}]$

احتمال ارتکاب خطای نوع اول را سطح معنی‌دار آزمون گویند و آن را به α نمایش می‌دهند.

$$\alpha=P[RH_0 \mid H_0 \text{ درست}]$$

احتمال ارتکاب خطای نوع دوم را با β نمایش می‌دهند.

$$\beta=P[AH_0 \mid H_0 \text{ غلط}]$$

۲ـ اگر متغیر تصادفی X دارای توزیع دوجمله‌ای با پارامترهای P و $n=5$ باشد کدام یک از فرض‌های زیر در مورد P ساده وکدام مرکب است؟

پاسخ:

$H_0: P=\frac{1}{2}$ فرض ساده $\qquad f(x)=\binom{5}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{5-x}$

$H_0: P\neq\frac{1}{2}$ فرض مرکب

$H_0: P=0.9$ فرض ساده $\qquad f(x)=\binom{5}{x}(0.9)^x(0.1)^{5-x}$

$H_0: P>\frac{1}{5}$ فرض مرکب

$H_0: P\leq\frac{1}{4}$ فرض مرکب

$H_0: P=\ 0\leq P\leq 0.2$ فرض مرکب

فرضی ساده است که به وسیله‌ی آن توزیع متغیر کاملاً مشخص شود ولی در فرض مرکب توزیع متغیر تصادفی مشخص نمی‌شود.

۳ـ فرض کنید متغیر تصادفی X دارای توزیع پواسن با پارامتر λ باشد برای آزمون فرض $H_0: \lambda=2$ در مقابل $\lambda=1.5$ اگر ناحیه‌ی بحرانی $C=\{x \mid x\leq 1\}$ باشد α و β و توان آزمون را حساب کنید.

پاسخ: ناحیه‌ی بحرانی $C=\{x \mid x\leq 1\}$

فضای بحرانی $C=\{0, 1\}$ $\qquad$ ناحیه‌ی قبول H_0 $\quad C'=\{2,3,\dots\}$

اگر فرض‌های آزمون به صورت زیر باشد، داریم:

$$\begin{cases} H_0: \lambda=2 \\ H_1: \lambda=1.5 \end{cases}$$

توزیع X تحت فرض H_0: $\quad X\sim P(2)\rightarrow f(x)=\frac{e^{-2}2^x}{x!}$

توزیع X تحت فرض H_1: $\quad X\sim P(1.5)\rightarrow f(x)=\frac{e^{-1.5}1.5^x}{x!}$

$$\alpha=P(RH_0 \mid H_0 \text{ درست})=P(X\leq 1 \mid \lambda=2)$$

$$=P(X=0 \text{ یا } X=1 \mid \lambda=2)=\frac{e^{-2}2^0}{0!}+\frac{e^{-2}2^1}{1!}$$

$$=e^{-2}+2e^{-2}=3e^{-2}=0.406$$

$$\beta=P(AH_0 \mid H_0 \text{ نادرست})=P(AH_0 \mid H_1 \text{ درست})=1-P(RH_0 \mid H_1 \text{ درست})$$

$$=1-P(X\leq 1 \mid \lambda=1.5)=1-P(X=0 \text{ یا } X=1 \mid \lambda=1.5)$$

$$=1-\left(\frac{e^{-1.5}1.5^0}{0!}+\frac{e^{-1.5}1.5^1}{1!}\right)=1-(e^{1.5}+1.5e^{-1.5})$$

$$=1-(2.5e^{-1.5})=1-0.558=0.442$$

$$\text{توان آزمون}=1-\beta=1-0.442=0.558$$

۴ـ فرض کنیم X در فاصله‌ی θ تا $-\theta$ به طور یکنواخت توزیع شود. یک مقدار x را مشاهده می‌کنیم و می‌خواهیم فرض $H_0: \theta=1$ را در مقابل $H_1: \theta=1.5$ آزمون کنیم. تصمیم می‌گیریم که اگر مقدار نمونه از 0.99 تجاوز کند H_0 را رد کنیم α و β و $1-\beta$ را حساب کنید.

پاسخ: اگر فرض‌های آزمون به صورت زیر باشد، داریم:

$$\begin{cases} H_0: \theta=1 \\ H_1: \theta=1.5 \end{cases}$$

می‌دانیم اگر $X\sim U(a, b)$ باشد تابع چگالی آن به صورت زیر است:

$$f(x)=\frac{1}{b-a} \qquad a<x<b$$

حال برای $X\sim U(-\theta, \theta)$ می‌باشد تابع چگالی به صورت زیر است:

$$f(x)=\frac{1}{2\theta} \qquad -\theta<x<\theta$$

توزیع X تحت فرض H_0: $X \sim U(-۱ , ۱) \to f(x) = \frac{۱}{۲}$

توزیع X تحت فرض H_1: $X \sim U(-۱/۵ , ۱/۵) \to f(x) = \frac{۱}{۳}$

$$\alpha = P(RH_0 \mid H_0 \text{ درست}) = P(X > ۰/۹۹ \mid \theta = ۱) = \int_{۰/۹۹}^{۱} \frac{۱}{۲}\, dx = ۰/۰۰۵$$

$$\beta = P(AH_0 \mid H_0 \text{ نادرست}) = P(AH_0 \mid H_1 \text{ درست}) = P(X < ۰/۹۹ \mid \theta = ۱/۵)$$

$$= \int_{-۱/۵}^{۰/۹۹} \frac{۱}{۳}\, dx = \frac{۱}{۳} x \Big|_{-۱/۵}^{۰/۹۹} = \frac{۱}{۳}(۰/۹۹ + ۱/۵) = \frac{۲/۴۹۰}{۳} = ۰/۸۳$$

توان آزمون $= ۱ - \beta = ۱ - ۰/۸۳ = ۰/۱۷$

۵ـ فرض کنید $\bar X$ میانگین یک نمونه‌ی تصادفی با حجم $n = ۳۶$ از جامعه‌ی نرمال با میانگین μ و واریانس $\sigma^2 = ۹$ باشد. برای آزمون فرض $H_0 : \mu = ۵۰$ در مقابل $H_1 : \mu = ۵۰$ اگر $\bar X \ge ۵۰/۸$ باشد فرض H_0 رد می‌شود. احتمال ارتکاب خطای نوع اول را حساب کنید.

پاسخ: $\bar X \ge ۵۰/۸ \to RH_0$ (ناحیه‌ی بحرانی)

$$\bar X \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(\mu, \frac{۹}{۳۶}\right)$$

توزیع $\bar X$ تحت فرض H_0: $\bar X \sim N\left(۵۰ , \frac{۹}{۳۶}\right)$

$$\alpha = P(RH_0 \mid H_0 \text{ درست}) = P(\bar X \ge ۵۰/۸ \mid \mu = ۵۰)$$

$$= P\left(\frac{\bar X - ۵۰}{\frac{۳}{۶}} \ge \frac{۵۰/۸ - ۵۰}{\frac{۳}{۶}}\right) = P(Z \ge ۱/۶) = ۱ - P(Z < ۱/۶)$$

$$= ۱ - ۰/۹۴۵۲ = ۰/۰۵۴۸$$

۶ـ اگر مقدار تشعشع کیهانی که یک خلبان در مدت پرواز با یک جت در معرض آن قرار می‌گیرد دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشد با استفاده از نمونه‌ی زیر فرض $H_0 : \mu = ۳/۳$ را در مقابل $H_1 : \mu < ۳/۳$ در سطح ۵ درصد آزمون کنید.

۳/۶۵ ۱/۴۳ ۱/۱۲ ۳/۴۰ ۲/۵۸ ۴/۴۷ ۴/۷۵ ۴/۲۲ ۲/۴۵

پاسخ: $H_0 : \mu = ۳/۳ \qquad H_1 : \mu < ۳/۳$

$n = ۹$, $\alpha = ۰/۰۵$, درجه‌ی آزادی $(n-1) = ۸$

$$\bar x = \frac{۱}{n} \sum_{i=۱}^{۹} x_i = ۳/۱۱۹$$

$$S^2 = \frac{۱}{n-۱} \sum_{i=۱}^{۹} (x_i - \bar x)^2 = ۱/۷۰۶ \quad , \quad S = ۱/۳۰۶$$

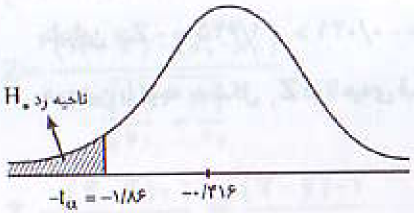
آماره‌ی آزمون

$$T_0 = \frac{\bar X - \mu_0}{\frac{S}{\sqrt n}} = \frac{۳/۱۱۹ - ۳/۳}{\frac{۱/۳۰۶}{۳}} = -۰/۴۱۶$$

مقدار جدول $P[T < -t_\alpha] = \alpha$

$P[T < -t_\alpha] = ۰/۰۵ \to -t_\alpha = -۱/۸۶$

اگر $T_0 < -t_\alpha$ باشد فرض H_0 را رد می‌کنیم. چون $T_0 = -۰/۴۱۶ > -t_\alpha$ می‌باشد پس فرض H_0 را می‌پذیریم.



۷ـ اگر $\bar X$ میانگین یک نمونه‌ی ۱۶ تایی از توزیع نرمال با میانگین μ و انحراف معیار $\sigma = ۸$ باشد. برای آزمون $H_0 : \mu = ۳۵$ در مقابل $H_1 : \mu > ۳۵$ فرض H_0 را رد می‌کنیم اگر $\bar X > ۳۶/۵$ باشد. مقدار P- را حساب کنید.

پاسخ: $n = ۱۶$, $\sigma = ۸$

$$\text{مقدار } P- = P[\bar X > ۳۶/۵ \mid H_0] = P\left[\frac{\bar X - \mu_0}{\frac{\sigma}{\sqrt n}} > \frac{۳۶/۵ - \mu_0}{\frac{\sigma}{\sqrt n}}\right]$$

$$= P\left[Z > \frac{۳۶/۵ - ۳۵}{\frac{۸}{۴}}\right] = P[Z > ۰/۷۵] = ۱ - P(Z < ۰/۷۵) = ۱ - ۰/۷۷۳۴ = ۰/۲۲۶۶$$

۸ـ در یک کارخانه، تجربه نشان داده است که ۱۵ درصد از کارگران به علت وجود ذرات سمی معلق در کارخانه پس از ۲۰ سال مبتلا به بیماری خاص می‌شوند. مدیریت جدید با ایجاد شرایط ادعا می‌کند که از ۱۴۰ کارگر حداکثر ۱۹ کارگر مبتلا به بیماری می‌شوند. آیا در سطح ۵ درصد ادعای مدیریت جدید پذیرفته می‌شود؟

پاسخ: می‌خواهیم فرض‌های زیر را آزمون کنیم.

$$\begin{cases} H_0 : P = ۰/۱۵ \\ H_1 : P < ۰/۱۵ \end{cases}$$

$n = ۱۴۰$, $x = ۱۹$, $\alpha = ۰/۰۵$, $P = ۰/۱۵$

برای انجام آزمون فرض‌های فوق از کمیت محوری $Z = \dfrac{\frac{X}{n} - P}{\sqrt{\frac{X}{n}\left(۱ - \frac{X}{n}\right)}}$ که برای n بزرگ تحت فرض H_0 دارای توزیع نرمال استاندارد است استفاده می‌کنیم.

آماره‌ی آزمون

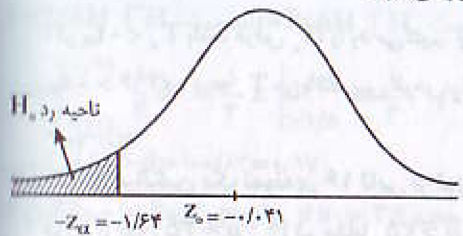
$$Z_0 = \frac{\frac{X}{n} - P}{\sqrt{\frac{X}{n}\left(۱ - \frac{X}{n}\right)}} = \frac{\frac{۱۹}{۱۴۰} - ۰/۱۵}{\sqrt{\frac{۱۹}{۱۴۰}\left(۱ - \frac{۱۹}{۱۴۰}\right)}} = -۰/۰۴۱$$

مقدار جدول $P[Z<-Z_\alpha]=0/05 \rightarrow -Z_\alpha=-1/645$

اگر $Z_0<-Z_\alpha$ فرض H_0 را رد می‌کنیم.

بنابراین $-Z_\alpha=-1/645<-0/041=Z_0$ پس فرض H_0 پذیرفته می‌شود.

همچنین با توجه به شکل Z_0 در ناحیه‌ی قبول H_0 قرار می‌گیرد.



۹ـ اگر $\bar X$ و $\bar Y$ به ترتیب میانگین‌های نمونه‌های تصادفی به حجم $n_1=14$ و $n_2=18$ از جامعه‌های نرمال با میانگین μ_1، واریانس $\sigma_1^2=26$ و میانگین μ_2، واریانس $\sigma_2^2=21$ باشند چگونگی آزمون $H_0: \mu_1-\mu_2=C$ را در مقابل $H_1: \mu_1-\mu_2<C$ در سطح ۰/۰۲۵ برای $C=0$ و $C=1$ شرح دهید.

پاسخ: ابتدا با توجه به این که $C=0$ باشد، فرض‌های زیر را آزمون می‌کنیم.

$$\begin{cases}H_0: \mu_1-\mu_2=0\\ H_1: \mu_1-\mu_2<0\end{cases}$$

کمیت محوری برای تفاضل میانگین دو جامعه در صورتی که واریانس‌های دو جامعه معلوم باشند به صورت زیر است:

$$Z=\frac{(\bar X-\bar Y)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

که Z تحت فرض H_0 دارای توزیع نرمال استاندارد خواهد بود. Z_0 مشاهده شده براساس مشاهده نمونه از دو جامعه برابر است با:

$$Z_0=\frac{\bar x-\bar y}{\sqrt{\frac{26}{14}+\frac{21}{18}}}=\frac{\bar x-\bar y}{1/739}$$

برای آزمون‌های $H_0: \mu_1-\mu_2=0$ در مقابل $H_1: \mu_1-\mu_2<0$ مراحل زیر را انجام می‌دهیم.

۱ـ آماره‌ی آزمون را تحت فرض H_0 و مشاهدات محاسبه می‌کنیم. (Z_0)

۲ـ مقدار Z_α را از رابطه‌ی $P(Z<-Z_\alpha)=\alpha$ به دست می‌آوریم که برای این تمرین داریم:

$$P(Z<-Z_\alpha)=0/025 \rightarrow -Z_\alpha=-1/96$$

۳ـ اگر $Z_0<-1/96$ باشد فرض H_0 رد می‌شود.

برای $C=1$ آزمون فرض برای تفاضل میانگین دو جامعه به صورت زیر است:

$$\begin{cases}H_0: \mu_1-\mu_2=1\\ H_1: \mu_1-\mu_2<1\end{cases}$$

که آماره‌ی آزمون برای فرض‌های بالا به صورت زیر است:

$$Z=\frac{(\bar X-\bar Y)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

که Z تحت فرض H_0 به صورت زیر می‌باشد:

$$Z_0=\frac{(\bar x-\bar y)-1}{\sqrt{\frac{26}{14}+\frac{21}{18}}}=\frac{(\bar x-\bar y)-1}{1/739}$$

که برای آزمون فرض‌های گفته شده تمام مراحل مانند قبل می‌باشد و اگر $Z_0<-Z_\alpha$ باشد فرض H_0 رد می‌شود.

۱۰ـ دو تولیدکننده‌ی لامپ ادعا می‌کنند که بر اساس فرایند جدید در تولید، عمر لامپ‌های تولیدی نسبت به گذشته افزایش یافته است. در یک بررسی نمونه از این دو تولیدکننده اطلاعات زیر به دست آمده است. آیا در سطح ۵ درصد می‌توان فرض $H_0: \mu_1=\mu_2$ در مقابل $H_1: \mu_1\neq\mu_2$ را پذیرفت.

	تولیدکننده‌ی اول	تولیدکننده‌ی دوم
حجم نمونه	۱۰۰	۱۰۰
میانگین	۷۹۸	۸۲۶
واریانس	۷۹۸۲	۹۰۰۱

پاسخ: می‌خواهیم فرض‌های زیر را آزمون کنیم:

$$\begin{cases}H_0: \mu_1=\mu_2\\ H_1: \mu_1\neq\mu_2\end{cases}$$

که می‌توان به صورت زیر نیز نوشت:

$$\begin{cases}H_0: \mu_1-\mu_2=0\\ H_1: \mu_1-\mu_2\neq 0\end{cases}$$

$$m=n=100 \quad , \quad \bar x_1=798 \quad , \quad \bar x_2=826$$

$$S_1^2=7982 \quad , \quad S_2^2=9001 \quad , \quad \alpha=0/05$$

آماره‌ی آزمون تحت فرض H_0:

$$Z_0=\frac{\bar x_1-\bar x_2}{\sqrt{\frac{S_1^2}{m}+\frac{S_2^2}{n}}}$$

$$Z_0=\frac{798-826}{\sqrt{\frac{7982}{100}+\frac{9001}{100}}}=\frac{-28}{\sqrt{169/83}}=-2/149$$

مقدار جدول:

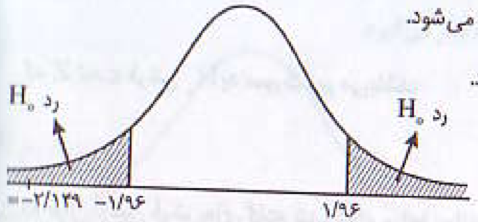
$$P(Z>Z_{\frac{\alpha}{2}})=۰/۰۲۵ \rightarrow Z_{\frac{\alpha}{2}}=۱/۹۶$$

$$P(Z<-Z_{\frac{\alpha}{2}})=۰/۰۲۵ \rightarrow Z_{\frac{\alpha}{2}}=-۱/۹۶$$

اگر $Z_0>Z_{\frac{\alpha}{2}}$ یا $Z_0<-Z_{\frac{\alpha}{2}}$ باشد فرض H_0 رد می‌شود.

چون $-۲/۱۴۹<-۱/۹۶$ پس فرض H_0 رد می‌شود.

با توجه به شکل $Z_0<-۱/۹۶$ می‌باشد و در ناحیه‌ی رد H_0 قرار می‌گیرد.



نکته: اگر در آزمون فرض برابری میانگین دو جامعه‌ی نرمال، واریانس‌ها مجهول و $n_1\ge ۳۰$ و $n_2\ge ۳۰$ باشد در این حالت در آماره‌ی آزمون به جای σ_1^2 و σ_2^2، از S_1^2 و S_2^2 (واریانس‌های نمونه‌ها) استفاده خواهیم کرد. بنابراین آماره‌ی آزمون به صورت زیر خواهد بود:

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{m}+\frac{S_2^2}{n}}}$$

۱۱ـ یک مهندس علاقه‌مند است قدرت استحکام دو نوع میلگرد از نوع A و B را با هم مقایسه کند. برای این منظور یک نمونه‌ی $n_1=۱۰$ از نوع A و یک نمونه‌ی $n_2=۱۲$ از نوع B انتخاب می‌کند و قدرت استحکام آن‌ها را اندازه‌گیری می‌کند و اطلاعات زیر را به دست می‌آورد:

$$n_1=۱۰ \;,\; n_2=۱۲ \;,\; \bar{x}_1=۸۲/۶ \;,\; \bar{x}_2=۷۸/۱ \;,\; S_1^2=۶/۵۲ \;,\; S_2^2=۷/۰۲$$

در سطح ۵ درصد $H_0:\mu_1=\mu_2$ را در مقابل $H_1:\mu_2<\mu_1$ آزمون کنید.

پاسخ: می‌خواهیم فرض‌های زیر را آزمون کنیم:

$$\begin{cases} H_0: & \mu_1-\mu_2=0 \\ H_1: & \mu_1-\mu_2>0 \end{cases}$$

برای آزمون فرض‌های بالا با توجه به این که واریانس‌ها دو جامعه‌ی مجهول و برابر هم یا نزدیک به هم باشند و حجم نمونه‌ها کوچک می‌باشد آماره‌ی آزمون تحت فرض H_0 دارای توزیع استودنت با n_1+n_2-2 درجه‌ی آزادی خواهد بود.

$$T_0=\frac{(\bar{x}_1-\bar{x}_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

که

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

برای محاسبه‌ی آماره‌ی آزمون داریم:

$$S_p^2=\frac{(۱۰-۱)۶/۵۲+(۱۲-۱)۷/۰۲}{۲۰}=۶/۷۹۵$$

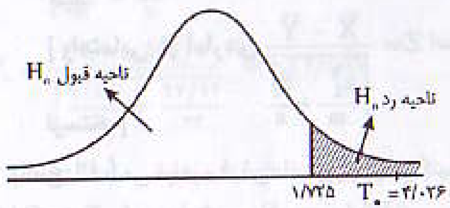
$$T_0=\frac{(۸۲/۶-۷۸/۱)}{\sqrt{۶/۷۹۵\left(\frac{1}{۱۰}+\frac{1}{۱۲}\right)}}=۴/۰۳۶$$

مقدار جدول با $n_1+n_2-2=۲۰$ درجه‌ی آزادی:

$$P[T>t_\alpha]=۰/۰۵\rightarrow t_\alpha=۱/۷۲۵$$

برای آزمون فرض‌های بالا اگر $T_0>t_{\alpha,(n_1+n_2-2)}$ باشد فرض H_0 رد می‌شود.

بنابراین $T_0=۴/۰۳۶>۱/۷۲۵=t_\alpha$ پس فرض H_0 رد می‌شود.



۱۲ـ در یک نمونه‌ی تصادفی از آزمون دانشجویان، نمرات ۱۰ نفر عبارتند از:

۹۵ ۸۲ ۴۰ ۵۲ ۶۰ ۸۰ ۸۲ ۵۸ ۶۵ ۵۰

در سطح ۵ درصد فرض $H_0:\sigma=۱۷$ را در مقابل $H_1:\sigma>۱۷$ آزمون کنید.

پاسخ: برای آزمون فرض‌های بالا می‌توان فرض‌های زیر را آزمون کرد:

$$\begin{cases} H_0: & \sigma^2=(۱۷)^2 \\ H_1: & \sigma^2>(۱۷)^2 \end{cases} \rightarrow \begin{cases} H_0: & \sigma^2=۲۸۹ \\ H_1: & \sigma^2>۲۸۹ \end{cases}$$

$$n=۱۰ \;,\; \alpha=۰/۰۵ \;,\; S^2=\frac{1}{n-1}\sum_{i=1}^{۱۰}(x_i-\bar{x})^2=۳۰۸/۴۸۹$$

آماره‌ی آزمون:

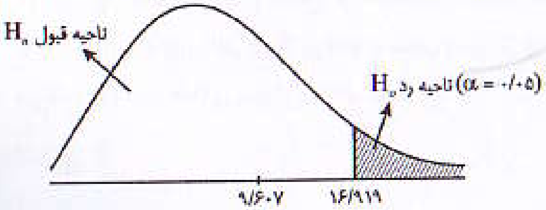
$$\chi_0^2=\frac{(n-1)S^2}{\sigma_0^2}=\frac{۹(۳۰۸/۴۸۹)}{۲۸۹}=۹/۶۰۷$$

مقدار جدول:

$$P\left[\chi^2>\chi^2_{(n-1,\alpha)}\right]=۰/۰۵ \rightarrow \chi^2_{(۹,\alpha)}=۱۶/۹۱۹$$

اگر $\chi_0^2>\chi^2_{(۹,۰/۰۵)}$ باشد فرض H_0 رد می‌شود.

حال با توجه به مقادیر به‌دست‌آمده چون $\chi_0^2=۹/۶۰۷<۱۶/۹۱۹=\chi^2_{(۹,۰/۰۵)}$ پس فرض H_0 پذیرفته می‌شود.



۱۳ـ برای بررسی میزان آلودگی هوای خانه در فاصله‌ی زمانی ۲۴ ساعته در یک نمونه‌ی ۱۶ تایی متوسط و واریانس میزان آلودگی در خانه‌هایی که فاقد افراد سیگاری است به‌ترتیب $\bar{x}=۶۷/۱$، و $S_x^۲=۷/۸۲$ در یک نمونه‌ی ۱۳ تایی متوسط و واریانس میزان آلودگی در خانه‌هایی که دارای حداقل یک فرد سیگاری است به ترتیب $\bar{y}=۱۳۲/۳$ و $S_y^۲=۲۴/۱۲$ به دست آمده است.

الف) فرض $H_۰:\sigma_۱^۲=\sigma_۲^۲$ را در مقابل $H_۱:\sigma_۱^۲<\sigma_۲^۲$ در سطح ۵ درصد آزمون کنید.

ب) اگر فرض $H_۰$ رد شد، فرض $H_۰:\mu_۱=\mu_۲$ را در مقابل $H_۱:\mu_۱<\mu_۲$ در سطح ۵ درصد آزمون کنید.

[راهنمایی: از آماره‌ی $Z=\dfrac{\bar{X}-\bar{Y}}{\sqrt{\dfrac{S_x^۲}{m}+\dfrac{S_y^۲}{n}}}$ استفاده کنید، علیرغم این که حجم نمونه‌ها زیاد بزرگ نیستند.]

پاسخ: الف) می‌خواهیم فرض‌های زیر را آزمون کنیم:

$$\begin{cases}H_۰:\sigma_۱^۲=\sigma_۲^۲\\ H_۱:\sigma_۱^۲<\sigma_۲^۲\end{cases}$$

$$m=۱۶\ ,\quad S_x^۲=۷/۸۲\ ,\quad n=۱۳\ ,\quad S_y^۲=۲۴/۱۲\ ,\quad \alpha=۰/۰۵$$

محاسبه‌ی مقدار آماره

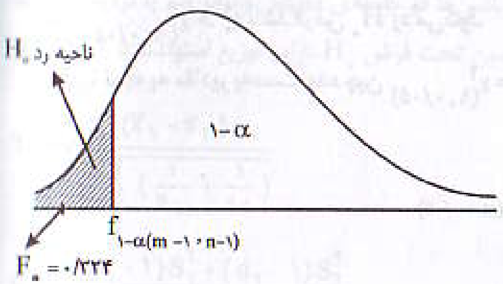
$$F_۰=\left[\frac{\sigma_۲}{\sigma_۱}\right]^۲\frac{S_۱^۲}{S_۲^۲}=۱\times\frac{۷/۸۲}{۲۴/۱۲}=۰/۳۲۴$$

مقدار جدول

$$f_{(m-۱,n-۱,۱-\alpha)}=f_{(۱۵,۱۲,۰/۹۵)}=\frac{۱}{f_{(۱۲,۱۵,۰/۰۵)}}=\frac{۱}{۲/۴۷۵۳}=۰/۴۰۴$$

اگر $F_۰<f_{۱-\alpha,(m-۱,n-۱)}$ باشد فرض $H_۰$ رد می‌شود.

بنابراین با توجه به مقدارهای به دست آمده داریم $F_۰=۰/۳۲۴<f_{(۱۵,۱۲,۰/۹۵)}=۰/۴۰۴$ پس فرض $H_۰$ رد می‌شود.



ب) می‌خواهیم فرض‌های زیر را آزمون کنیم:

$$\begin{cases}H_۰:\ \mu_۱-\mu_۲=۰\\ H_۱:\ \mu_۱-\mu_۲<۰\end{cases}$$

$$m=۱۶\ ,\quad n=۱۳\ ,\quad \bar{x}=۶۷/۱\ ,\quad S_x^۲=۷/۸۲\ ,\quad \bar{y}=۱۳۲/۳\ ,\quad S_y^۲=۲۴/۱۲$$

آماره‌ی آزمون تحت فرض $H_۰$ به صورت زیر است:

$$Z_۰=\frac{\bar{x}-\bar{y}}{\sqrt{\dfrac{S_x^۲}{m}+\dfrac{S_y^۲}{n}}}$$

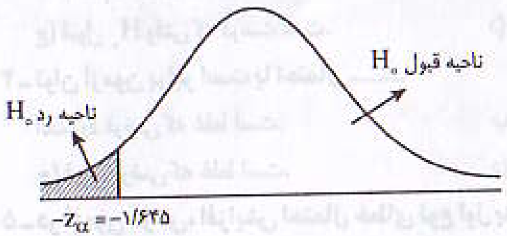
$$Z_۰=\frac{۶۷/۱-۱۳۲/۳}{\sqrt{\dfrac{۷/۸۲}{۱۶}+\dfrac{۲۴/۱۲}{۱۳}}}=-۴۲/۵۸۷$$

مقدار جدول

$$P[Z<-Z_\alpha]=۰/۰۵\rightarrow -Z_\alpha=-۱/۶۴۵$$

اگر $Z_۰<-Z_\alpha$ باشد فرض $H_۰$ رد می‌شود.

بنابراین $Z_۰=-۴۲/۵۸۷<-۱/۶۴۵=-Z_\alpha$ پس فرض $H_۰$ رد می‌شود.

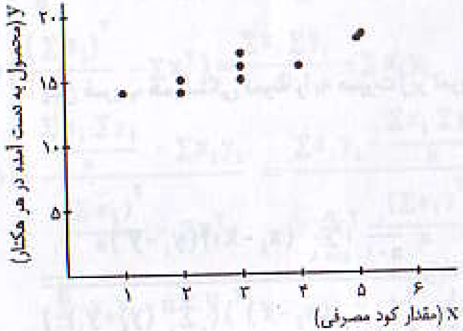


خودآزمایی فصل هشتم

۱- مقدار کود مصرفی (X) بر حسب تن در هکتار و محصول به دست آمده از هر هکتار (Y) در جدول زیر آمده است. نمودار پراکنش بین X و Y را رسم کنید و ضریب همبستگی نمونه‌ای را به دست آورید.

x	۱	۲	۵	۳	۲	۳	۳	۵	۴
y	۱۴	۱۴	۱۸	۱۵	۱۵	۱۷	۱۶	۱۸	۱۷

پاسخ:



برای به دست آوردن ضریب همبستگی نمونه داریم:

x_i	۱	۲	۵	۳	۲	۳	۳	۵	۴	۲۸
y_i	۱۴	۱۴	۱۸	۱۵	۱۵	۱۷	۱۶	۱۸	۱۷	۱۴۴
$x_i y_i$	۱۴	۲۸	۹۰	۴۵	۳۰	۵۱	۴۸	۹۰	۶۷	۴۶۴
$(x_i-\bar{x})^۲$	۴/۴۵	۱/۲۳	۳/۵۷	۰/۰۱	۱/۲۳	۰/۰۱	۰/۰۱	۳/۵۷	۰/۷۹	۱۴/۸۷
$(y_i-\bar{y})^۲$	۴	۴	۴	۱	۱	۱	۰	۴	۱	۲۰

$$n=۹ \ , \ \bar{x}=۳/۱۱ \ , \ \bar{y}=۱۶ \ , \ \sum_{i=۱}^{۹} x_i y_i = ۴۶۴$$

$$\sum_{i=۱}^{۹}(x_i-\bar{x})^۲ = ۱۴/۸۷ \ , \ \sum_{i=۱}^{۹}(y_i-\bar{y})^۲ = ۲۰$$

$$r=\frac{\sum_{i=۱}^{n} x_i y_i - n\bar{x}\,\bar{y}}{\sqrt{(\sum_{i=۱}^{n}(x_i-\bar{x})^۲)(\sum_{i=۱}^{n}(y_i-\bar{y})^۲)}}=\frac{۴۶۴-۹(۳/۱۱)(۱۶)}{\sqrt{(۱۴/۸۷)(۲۰)}}=۰/۹۴$$

با توجه به نمودار پراکنش و مقدار همبستگی، نتیجه می‌شود همبستگی مثبت است یعنی با افزایش کود مصرفی، مقدار محصول به دست آمده از هر هکتار نیز افزایش می‌یابد.

۲ـ نشان دهید که:

$$r=\frac{\sum x_iy_i-n\bar{x}\,\bar{y}}{\sum x_i^2-n\bar{x}^2}\quad,\quad r=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}}$$

پاسخ: در فصل‌های قبل دیدیم که ضریب همبستگی در جامعه را به صورت زیر تعریف می‌کنیم:

$$\rho=\frac{Cov(X,Y)}{\sqrt{V(X).V(Y)}}$$

حال اگر ضریب همبستگی نمونه را با r نشان دهیم داریم:

کوواریانس نمونه $S_{x,y}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$

واریانس نمونه x $S_x^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$

کوواریانس نمونه y $S_y^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$

پس ضریب همبستگی نمونه را به صورت زیر تعریف می‌کنیم:

$$r=\frac{S_{xy}}{\sqrt{S_x^2\,.\,S_y^2}}$$

$$=\frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}\right).\left(\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}\right)}}=\frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\frac{1}{n-1}\sqrt{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)\left(\sum_{i=1}^{n}(y_i-\bar{y})^2\right)}}$$

بنابراین داریم:

$$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)\left(\sum_{i=1}^{n}(y_i-\bar{y})^2\right)}}$$

* می‌خواهیم نشان دهیم که:

$$b=\frac{\sum x_iy_i-n\bar{x}\,\bar{y}}{\sum x_i^2-n\bar{x}^2}$$

پاسخ: برای به دست آوردن مقدار b باید $\sum_{i=1}^{n}e_i^2$ را مینیمم کرد.

$$e_i=y_i-a-bx_i$$

$$e_i^2=(y_i-a-bx_i)^2$$

$$\sum_{i=1}^{n}e_i^2=\sum_{i=1}^{n}(y_i-a-bx_i)^2$$

با فرض برابری این رابطه

$$\phi(a,b)=\sum_{i=1}^{n}e_i^2$$

تابع $\phi(a,b)$ تابعی از a و b خواهد بود. برای به دست آوردن a و b از تابع $\phi(a,b)$ یک بار نسبت به a و یک بار نسبت به b مشتق می‌گیریم و برابر صفر قرار می‌دهیم.

$$\begin{cases}\dfrac{\partial\phi(a,b)}{\partial a}=-2\sum_{i=1}^{n}(y_i-a-bx_i)=0\\[2mm]\dfrac{\partial\phi(a,b)}{\partial b}=-2\sum_{i=1}^{n}x_i(y_i-a-bx_i)=0\end{cases}$$

از حل دو معادله‌ی اخیر داریم:

$$\xrightarrow{\times(-\bar{x})}\begin{cases}\sum_{i=1}^{n}y_i-na-b\sum_{i=1}^{n}x_i=0\\\sum_{i=1}^{n}x_iy_i-a\sum_{i=1}^{n}x_i-b\sum_{i=1}^{n}x_i^2=0\end{cases}$$

پس داریم:

$$\frac{-\sum x_i\sum y_i}{n}+\sum x_iy_i+b\frac{(\sum x_i)^2}{n}-b\sum x_i^2=0$$

$$b\left(\frac{(\sum x_i)^2}{n}-\sum x_i^2\right)=\frac{\sum x_i\sum y_i}{n}-\sum x_iy_i$$

$$b=\frac{\frac{\sum x_i\sum y_i}{n}-\sum x_iy_i}{\frac{(\sum x_i)^2}{n}-\sum x_i^2}=\frac{\sum x_iy_i-\frac{\sum x_i\sum y_i}{n}}{\sum x_i^2-\frac{(\sum x_i)^2}{n}}$$

با توجه به این که $\sum_{i=1}^{n}x_i=n\bar{x}$ و $\sum_{i=1}^{n}y_i=n\bar{y}$ داریم:

$$b=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\,\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}$$

برای براورد a نیز داریم:

$$na=\sum_{i=1}^{n}y_i-b\sum_{i=1}^{n}x_i$$

$$a=\frac{\sum_{i=1}^{n}y_i}{n}-b\frac{\sum_{i=1}^{n}x_i}{n}\rightarrow a=\bar{y}-b\bar{x}$$

۳ـ اگر $b_{yx}=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{n\sum x_i^2-(\sum x_i)^2}$ **و** $b_{xy}=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{n\sum y_i^2-(\sum y_i)^2}$ **تعریف شوند. نشان دهید که r برابر با میانگین هندسی b_{xy} و b_{yx} است.**

پاسخ: می‌خواهیم نشان دهیم که:

$$r=\sqrt{b_{xy}\,.\,b_{yx}}$$

برای سهولت در محاسبات ابتدا b_{xy} و b_{yx} را ساده می‌کنیم داریم:

$$b_{xy}=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{n\sum y_i^2-(\sum y_i)^2}$$

با توجه به این که $\sum x_i = n\bar{x}$ و $\sum y_i = n\bar{y}$ پس داریم:

$$b_{xy} = \frac{n\sum x_i y_i - n^2 \bar{x}\,\bar{y}}{n\sum y_i^2 - n^2\bar{y}^2} = \frac{n(\sum x_i y_i - n\bar{x}\,\bar{y})}{n(\sum y_i^2 - n\bar{y}^2)} = \frac{\sum x_i y_i - n\bar{x}\,\bar{y}}{\sum (y_i - \bar{y})^2}$$

همچنین برای b_{yx} داریم:

$$b_{yx} = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2} = \frac{n\sum x_i y_i - n^2\bar{x}\,\bar{y}}{n\sum x_i^2 - n^2\bar{x}^2} = \frac{n(\sum x_i y_i - n\bar{x}\,\bar{y})}{n(\sum x_i^2 - n\bar{x}^2)}$$

$$= \frac{\sum x_i y_i - n\bar{x}\,\bar{y}}{\sum (x_i - \bar{x})^2}$$

حال می‌توان نشان داد که:

$$r = \sqrt{b_{xy}\,.\,b_{yx}} = \sqrt{\frac{\sum x_i y_i - n\bar{x}\,\bar{y}}{\sum (y_i - \bar{y})^2}\,.\,\frac{\sum x_i y_i - n\bar{x}\,\bar{y}}{\sum (x_i - \bar{x})^2}}$$

$$r = \sqrt{\frac{(\sum x_i y_i - n\bar{x}\,\bar{y})^2}{[\sum (y_i - \bar{y})^2][\sum (x_i - \bar{x})^2]}} = \frac{\sum x_i y_i - n\bar{x}\,\bar{y}}{\sqrt{[\sum (x_i - \bar{x})^2][\sum (y_i - \bar{y})^2]}}$$

۴ـ فرض کنید که $E[y|X=x]=\beta X$ و $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ یک نمونه‌ی تصادفی از Y با مقادیر X وابسته به آن‌ها باشد. فرض کنید $V(y_i)=\sigma_i^2$ نشان دهید که برآوردگر:

$$\hat{b} = \frac{\sum \frac{x_i y_i}{\sigma_i^2}}{\sum \frac{x_i^2}{\sigma_i^2}}$$

یک برآوردگر ناریب برای β است.

پاسخ: می‌دانیم $y_i \sim N(\beta x_i, \sigma_i^2)$ می‌باشد و برای این که $\hat{b}$ یک برآوردگر ناریب برای β باشد باید $E(\hat{b})=\beta$ شود پس داریم:

$$\hat{b} = \frac{\sum \frac{x_i y_i}{\sigma_i^2}}{\sum \frac{x_i^2}{\sigma_i^2}} = \sum c_i y_i$$

که در آن:

$$c_i = \frac{\frac{x_i}{\sigma_i^2}}{\sum \frac{x_i^2}{\sigma_i^2}}$$

حال نشان می‌دهیم که:

$$E(\hat{b}) = E[\sum c_i y_i] = \sum c_i E(y_i) = \sum c_i (\beta x_i) = \beta \sum c_i x_i = \beta \frac{\sum \frac{x_i^2}{\sigma_i^2}}{\sum \frac{x_i^2}{\sigma_i^2}} = \beta$$

پس $\hat{b}$ یک برآوردگر ناریب برای β است.

۵ـ اطلاعات زیر مربوط به متوسط فروش روزانه رایانه و هزینه‌ی تبلیغات روزانه‌ی یکی از نمایندگی‌هاست.

هزینه تبلیغات (هزار ریال)	X	۲	۱	۴	۳	۵	۷	۹	۶	۸	۱۰
فروش روزانه (ده هزار ریال)	Y	۲۸	۲۰	۳۲	۲۹	۳۶	۳۳	۲۶	۳۴	۳۳	۴۰

الف) ضریب همبستگی بین هزینه‌ی تبلیغات و فروش روزانه را به دست آورید.

ب) اگر $\hat{y}=\hat{a}+\hat{b}x$ خط برازش شده باشد $\hat{a}$ و $\hat{b}$ را به دست آورید.

ج) اگر ۱۱۰۰۰ ریال هزینه‌ی تبلیغات شود تعداد فروش چقدر است؟

پاسخ: الف)

هزینه‌ی تبلیغات	X_i	۲	۱	۴	۳	۵	۷	۹	۶	۸	۱۰	۵۵
فروش روزانه	Y_i	۲۸	۲۰	۳۲	۲۹	۳۶	۳۳	۲۶	۳۴	۳۳	۴۰	۳۱۱
	$x_i y_i$	۵۶	۲۰	۱۲۸	۸۷	۱۸۰	۲۳۱	۲۳۴	۲۰۴	۲۶۴	۴۰۰	۱۸۰۴
	$(x_i-\bar{x})^2$	۱۲/۲۵	۲۰/۲۵	۲/۲۵	۶/۲۵	۰/۲۵	۲/۲۵	۱۲/۲۵	۰/۲۵	۶/۲۵	۲۰/۲۵	۸۲/۵
	$(y_i-\bar{y})^2$	۹/۶۱	۱۲۳/۲۱	۰/۸۱	۴/۴۱	۲۴/۰۱	۳/۶۱	۲۶/۰۱	۸/۴۱	۳/۶۱	۷۹/۲۱	۲۸۲/۹۰

$$n=۱۰ \quad , \quad \bar{x}=۵/۵ \quad , \quad \bar{y}=۳۱/۱۰ \quad , \quad \sum_{i=1}^{10} x_i y_i = ۱۸۰۴$$

$$\sum_{i=1}^{10} (x_i-\bar{x})^2 = ۸۲/۵ \quad , \quad \sum_{i=1}^{10} (y_i-\bar{y})^2 = ۲۸۲/۹۰$$

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\,\bar{y}}{\sqrt{(\sum (x_i-\bar{x})^2)(\sum (y_i-\bar{y})^2)}} = \frac{۱۸۰۴ - ۱۰(۵/۵)(۳۱/۱۰)}{\sqrt{(۸۲/۵)(۲۸۲/۹۰)}} = ۰/۶۱$$

ب)

$$\hat{b} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\,\bar{y}}{\sum_{i=1}^{n} (x_i-\bar{x})^2} = \frac{۱۸۰۴ - ۱۰(۵/۵)(۳۱/۱۰)}{۸۲/۵} = ۱/۱۳$$

$$\hat{a} = \bar{y} - \hat{b}\,\bar{x} = ۳۱/۱۰ - (۱/۱۳)(۵/۵) = ۳۱/۱۰ - ۶/۲۲ = ۲۴/۸۸$$

خط برازش شده

$$\hat{y} = ۲۴/۸۸ + ۱/۱۳x$$

ج) می‌خواهیم، پیش‌بینی تعداد فروش را برای $x_0 = ۱۱$ (هزار ریال) به دست آوریم:

خط برازش داده شده: $\hat{y} = ۲۴/۸۸ + ۱/۱۳x$

برای $x_0 = ۱۱$ (هزار ریال): $\hat{y}_0 = ۲۴/۸۸ + ۱/۱۳(۱۱) = ۳۷/۳۱$

که می‌توان تعداد فروش را برای $x_0 = ۱۱$، تقریباً ۳۷ در نظر گرفت.

۶ـ اگر در یک نمونه‌ی ۱۰ تایی داشته باشیم:

$$\sum x_i = ۵۰۰ \quad , \quad \sum y_i = ۱۱۰۰ \quad , \quad \sum x_i y_i = ۶۱۸۰۰ \quad , \quad \sum x_i^2 = ۲۸۴۰۰$$

خط رگرسیون را به دست آورید.

پاسخ: اگر $\hat{y} = \hat{a} + \hat{b}x$ خط برازش شده باشد، برای به دست آوردن $\hat{a}$ و $\hat{b}$ داریم:

$$\hat{b} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

که در آن:

$$n = ۱۰ \; , \; \bar{x} = ۵۰ \; , \; \bar{y} = ۱۱۰ \; , \; \sum x_i y_i = ۶۱۸۰۰ \; , \; \sum x_i^2 = ۲۸۴۰۰$$

می‌باشد و با توجه به مقادیری که مسئله به ما داده است داریم:

$$\hat{b} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{۶۱۸۰۰ - ۱۰(۵۰)(۱۱۰)}{۲۸۴۰۰ - ۱۰(۵۰)^2} = \frac{۶۸۰۰}{۳۴۰۰} = ۲$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = ۱۱۰ - ۲(۵۰) = ۱۰$$

پس خط رگرسیون برازش شده به صورت زیر است:

$$\hat{y} = ۱۰ + ۲x$$

۷ـ با توجه به دو معادله‌ی رگرسیون نمونه‌ای، میانگین $\bar{x}$ و $\bar{y}$ و r را به دست آورید.

$$۲۵x = ۶y + ۷ \; , \; ۴y = ۹x + ۱۵$$

پاسخ: رگرسیون x نسبت به y: $۲۵x = ۶y + ۷ \to x = \frac{۶}{۲۵}y + \frac{۷}{۲۵}$

که برای این رگرسیون $b_{xy} = \frac{۶}{۲۵}$

رگرسیون y نسبت به x: $۴y = ۹x + ۱۵ \to y = \frac{۹}{۴}x + \frac{۱۵}{۴}$

که برای این معادله $b_{yx} = \frac{۹}{۴}$

با توجه به تمرین ۳، گفته شد که r برابر است با میانگین هندسی b_{xy} و b_{yx} پس داریم:

$$r = \sqrt{b_{xy} \cdot b_{yx}} = \sqrt{\frac{۶}{۲۵} \times \frac{۹}{۴}} = \frac{۳\sqrt{۶}}{۱۰} = \frac{۷/۳۵}{۱۰} = ۰/۷۳۵$$

حال می‌خواهیم مقادیر $\bar{x}$ و $\bar{y}$ را به دست آوریم: ابتدا دو معادله‌ی رگرسیون داده شده را ساده می‌کنیم بنابراین داریم:

معادله‌ی (۱): $۲۵x = ۶y + ۷ \to ۶y = ۲۵x - ۷ \to y = \frac{۲۵}{۶}x - \frac{۷}{۶}$

معادله‌ی (۲): $۴y = ۹x + ۱۵ \to y = \frac{۹}{۴}x + \frac{۱۵}{۴}$

با توجه به معادله‌ی (۱) داریم: $a_1 = -\frac{۷}{۶} \; , \; b_1 = \frac{۲۵}{۶}$

با توجه به معادله‌ی (۲) داریم: $a_2 = \frac{۱۵}{۴} \; , \; b_2 = \frac{۹}{۴}$

همچنین می‌دانیم که $\hat{a} = \bar{y} - b\bar{x}$ پس:

$$\begin{cases} \frac{-۷}{۶} = \bar{y} - \frac{۲۵}{۶}\bar{x} & \text{(۱) بر اساس معادله‌ی} \quad \times ۶ \\ \frac{۱۵}{۴} = \bar{y} - \frac{۹}{۴}\bar{x} & \text{(۲) بر اساس معادله‌ی} \quad \times ۴ \end{cases}$$

$$\begin{cases} -۷ = ۶\bar{y} - ۲۵\bar{x} & \times(-۲) \\ ۱۵ = ۴\bar{y} - ۹\bar{x} & \times ۳ \end{cases}$$

$$\to \begin{cases} ۱۴ = -۱۲\bar{y} + ۵۰\bar{x} \\ ۴۵ = ۱۲\bar{y} - ۲۷\bar{x} \end{cases} \to ۵۹ = ۲۳\bar{x} \to \bar{x} = \frac{۵۹}{۲۳} = ۲/۵۷$$

با توجه به این که $\bar{x} = ۲/۵۷$ داریم:

$$۱۵ = ۴\bar{y} - ۹(۲/۵۷) \to ۴\bar{y} = ۳۸/۱۳ \to \bar{y} = ۹/۵۳$$

حال می‌خواهیم میانگین $r = ۰/۷۳۵$، $\bar{x} = ۲/۵۷$ و $\bar{y} = ۹/۵۳$ را به دست آوریم:

$$\text{میانگین} = \frac{۰/۷۳۵ + ۲/۵۷ + ۹/۵۳}{۳} = ۴/۲۸$$

۸ـ نمره‌های میان ترم (X) و پایان ترم (Y) یک کلاس ۹ نفری به صورت جدول زیر است:

X	۷۷	۵۰	۷۱	۷۲	۸۱	۹۴	۹۶	۹۹	۶۷
Y	۸۲	۶۶	۷۸	۳۴	۴۷	۸۵	۹۹	۹۹	۶۸

الف) معادله‌ی خط رگرسیون را پیدا کنید.

ب) نمره‌ی پایان ترم دانشجویی که میان ترم ۸۵ گرفته است و در پایان ترم به علت بیماری نتوانسته است در امتحان پایان ترم شرکت کند تخمین بزنید.

پاسخ: الف)

x_i	۷۷	۵۰	۷۱	۷۲	۸۱	۹۴	۹۶	۹۹	۶۷	۷۰۷
yi	۸۲	۶۶	۷۸	۳۴	۴۷	۸۵	۹۹	۹۹	۶۸	۶۵۸
xiyi	۶۳۱۴	۳۳۰۰	۵۵۳۸	۲۴۴۸	۳۸۰۷	۷۹۹۰	۹۵۰۴	۹۸۰۱	۴۵۵۶	۵۳۲۵۸
$(x_i-\bar{x})^۲$	۲/۴۳	۸۱۵/۶۷	۵۷/۱۵	۴۲/۰۳	۵/۹۵	۲۲۸/۳۹	۳۰۴/۱۵	۴۱۷/۷۹	۱۳۲/۶۳	۲۰۱۸/۱۹

$$n=۹ \quad , \quad \bar{x}=۷۸/۵۶ \quad , \quad \bar{y}=۷۳/۱۱ \quad , \quad \sum_{i=۱}^{۹} x_i y_i = ۵۳۲۵۸$$

$$\sum_{i=۱}^{۹} (x_i-\bar{x})^۲ = ۲۰۱۸/۱۹$$

اگر معادله‌ی خط رگرسیون به صورت زیر باشد داریم:

$$\hat{y}=\hat{a}+\hat{b}x$$

$$\hat{b}=\frac{\sum_{i=۱}^{n} x_i y_i - n\bar{x}\,\bar{y}}{\sum_{i=۱}^{n}(x_i-\bar{x})^۲}=\frac{۵۳۲۵۸-۹(۷۸/۵۶)(۷۳/۱۱)}{۲۰۱۸/۱۹}=\frac{۱۵۹۹/۳۱}{۲۰۱۸/۱۹}=۰/۷۹$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}=۷۳/۱۱-۰/۷۹(۷۸/۵۶)=۷۳/۱۱-۶۲/۰۶=۱۱/۰۵$$

معادله‌ی خط برازش شده

$$\hat{y}=۱۱/۰۵+۰/۷۹x$$

ب) برای تخمین نمره‌ی پایان ترم دانشجویی که مقدار نمره‌ی میان ترم او $x_۰=۸۵$ است داریم:

$$x_۰=۸۵\rightarrow \hat{y}_۰=۱۱/۰۵+۰/۷۹\,(۸۵)=۱۱/۰۵+۶۷/۱۵=۷۸/۲۰$$

۹ـ اگر متغیرهای

$$T_\alpha=\frac{\hat{a}-\alpha}{\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}} \quad , \quad T_\beta=\frac{\hat{b}-\beta}{\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}}$$

کمیت‌های محوری باشند فاصله‌ی اطمینان $(۱-\alpha)$ درصد برای β و α را به دست آورید.

پاسخ:

نکته: در مدل رگرسیون ساده خطی با فرض $E_i\sim N(۰\,,\,\sigma^۲)$ که در آن $i=۱\,,\,۲\,,\,\ldots\,,\,n$ داریم:

$$\text{الف) } \hat{\beta}\sim N\left(\beta\,,\,\frac{\sigma^۲}{\sum_{i=۱}^{n}(X_i-\bar{X})^۲}\right) \quad \text{و} \quad \text{ب) } \hat{a}\sim N\left(\alpha\,,\,\sigma^۲\left(\frac{۱}{n}+\frac{\bar{X}^۲}{\sum_{i=۱}^{n}(X_i-\bar{X})^۲}\right)\right)$$

و اگر $\sigma^۲$ مجهول باشد از برآوردگر آن $\hat{\sigma}^۲_E=\frac{۱}{n-۱}\sum_{i=۱}^{n} e_i^۲$ استفاده می‌کنیم.

بنابراین با توجه به نکته‌ی گفته شده اگر $\sigma^۲$ مجهول باشد.

ب)

$$T_\alpha=\frac{\hat{a}-\alpha}{\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}}\sim t_{(n-۲)}$$

و

الف)

$$T_\beta=\frac{\hat{b}-\beta}{\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}}\sim t_{(n-۲)}$$

از طرفی می‌دانیم $P(-t_{\frac{\alpha}{۲}}<T<t_{\frac{\alpha}{۲}})=۱-\alpha$ با جایگذاری مقادیر T یک بار از قسمت الف و بار دیگر از قسمت (ب) می‌توان فاصله‌ی اطمینان برای β و α را به صورت زیر به دست آورد:

الف)

$$P\left(-t_{\frac{\alpha}{۲},n-۲}<\frac{\hat{b}-\beta}{\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}}<t_{\frac{\alpha}{۲},n-۲}\right)=۱-\alpha$$

$$\rightarrow P\left(-t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}<\hat{b}-\beta<t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}\right)=۱-\alpha$$

$$\rightarrow P\left(\hat{b}-t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}<\beta<\hat{b}+t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\frac{\hat{\sigma}^۲}{\sum(X_i-\bar{X})^۲}}\right)=۱-\alpha$$

ب) فاصله‌ی اطمینان برای α به صورت زیر است:

$$P\left(-t_{\frac{\alpha}{۲},(n-۲)}<\frac{\hat{a}-\alpha}{\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}}<t_{\frac{\alpha}{۲},(n-۲)}\right)=۱-\alpha$$

$$\rightarrow P\left(-t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}<\hat{a}-\alpha<t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}\right)=۱-\alpha$$

$$\rightarrow P\left(\hat{a}-t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}<\alpha<\hat{a}+t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n}+\frac{\bar{X}^۲}{\sum(X_i-\bar{X})^۲}\right]}\right)=۱-\alpha$$

۱۰ـ اگر $C_i = \dfrac{X_i - \bar{X}}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}$ باشد نشان دهید که:

$$\sum_{i=۱}^{n} C_i = ۰ \ , \ \sum_{i=۱}^{n} C_i X_i = ۱ \ , \ \sum_{i=۱}^{n} C_i^۲ = \frac{۱}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}$$

پاسخ:

$$* \sum_{i=۱}^{n} C_i = \frac{\sum_{i=۱}^{n} (X_i - \bar{X})}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}$$

با توجه به خواص میانگین حسابی داریم: $\sum_{i=۱}^{n} (X_i - \bar{X}) = ۰$ پس داریم:

$$\frac{\sum_{i=۱}^{n} (X_i - \bar{X})}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲} = ۰$$

$$* \sum_{i=۱}^{n} C_i X_i = \frac{\sum_{i=۱}^{n} X_i (X_i - \bar{X})}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲} = \frac{\sum_{i=۱}^{n} (X_i^۲ - X_i \bar{X})}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲} = \frac{\sum_{i=۱}^{n} X_i^۲ - \bar{X} \sum_{i=۱}^{n} X_i}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}$$

$$= \frac{\sum_{i=۱}^{n} X_i^۲ - n\bar{X}^۲}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲} = \frac{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲} = ۱$$

$$* \sum_{i=۱}^{n} C_i^۲ = \sum_{i=۱}^{n} \left[\frac{X_i - \bar{X}}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲} \right]^۲ = \frac{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}{\left[\sum_{i=۱}^{n} (X_i - \bar{X})^۲\right]^۲} = \frac{۱}{\sum_{i=۱}^{n} (X_i - \bar{X})^۲}$$

۱۱ـ اگر متغیرهای X و Y دارای تابع چگالی توأم زیر باشند نشان دهید که:

$$E[Y|X=x] = \frac{۲}{۳} - \frac{۲}{۳}x$$

$$f(x, y) = ۲۴xy \qquad x>۰ \ , \ y>۰ \ , \ x+y<۱$$

پاسخ: با انتگرال‌گیری نسبت به y، چگالی حاشیه‌ای x به صورت زیر به دست می‌آید:

$$f(x) = \int_۰^{۱-x} ۲۴xy\,dy = ۲۴ \int_۰^{۱-x} xy\,dy = ۲۴ \left[x \frac{y^۲}{۲} \right]_۰^{۱-x}$$

$$= ۲۴ \left[\frac{x(۱-x)^۲}{۲} \right] = ۱۲x(۱-x)^۲$$

بنابراین چگالی شرطی Y به شرط X=x به صورت زیر است:

$$f(y|x) = \frac{f(x, y)}{f(x)} = \frac{۲۴xy}{۱۲x(۱-x)^۲} = \frac{۲y}{(۱-x)^۲}$$

حال برای محاسبه‌ی $E[Y|X=x]$ داریم:

$$E[Y|X=x] = \int_۰^{۱-x} y \cdot \frac{۲y}{(۱-x)^۲}\,dy = \int_۰^{۱-x} \frac{۲y^۲}{(۱-x)^۲}\,dy = \left[\frac{۲y^۳}{۳(۱-x)^۲} \right]_۰^{۱-x}$$

$$= \left[\frac{۲(۱-x)^۳}{۳(۱-x)^۲} - ۰ \right] = \frac{۲}{۳}(۱-x) = \frac{۲}{۳} - \frac{۲}{۳}x$$

۱۲ـ جدول زیر مقادیر ارزیابی شده و قیمت‌های فروش ۸ خانه را نشان می‌دهد:

مقدارارزیابی‌شده (X)	۷۰/۳	۱۰۲	۶۲/۵	۷۴/۸	۵۷/۹	۸۱/۶	۱۱۰/۴	۸۸
(Y) قیمت فروش	۱۱۴/۴	۱۶۹/۳	۱۰۶/۲	۱۲۵	۹۹/۸	۱۳۲/۱	۱۷۴/۲	۱۴۳/۵

الف) یک خط رگرسیون کم‌ترین مربعات برای داده‌ها برازش دهید.

ب) فرض $H_۰ = \beta = ۱/۳$ را در مقابل $H_۱ = \beta \neq ۱/۳$ در سطح ۵ درصد آزمون کنید

پاسخ: الف)

x_i	۷۰/۳	۱۰۲	۶۲/۵	۷۴/۸	۵۷/۹	۸۱/۶	۱۱۰/۴	۸۸	۶۴۷/۵
y_i	۱۱۴/۴	۱۶۹/۳	۱۰۶/۲	۱۲۵	۹۹/۸	۱۳۲/۱	۱۷۴/۲	۱۴۳/۵	۱۰۶۴/۵
$x_i y_i$	۸۰۴۲/۳۲	۱۷۲۶۸/۶۰	۶۶۳۷/۵۰	۹۳۵۰	۵۷۷۸/۴۲	۱۰۷۷۹/۳۶	۱۹۲۳۱/۶۸	۱۲۶۲۸	۸۹۷۱۵/۸۸
$(x_i - \bar{x})^۲$	۱۱۳/۲۱	۴۴۳/۵۲	۲۴۰/۰۲	۳۷/۷۰	۵۳۰/۸۴	۰/۴۴	۸۶۷/۸۹	۴۹/۸۴	۲۳۸۳/۴۷
e_i	۲۷/۱۹	۴/۸۶	۰/۶۱	۱/۰۹	۱/۰۷	۱/۹۴-	۲/۷۶-	۰/۰۸-	
$e_i^۲$	۷۳۹/۳۰	۲۳/۶۲	۰/۳۷	۱/۱۹	۱/۱۴	۳/۷۶	۷/۶۲	۰/۰۱	۷۷۷/۰۱

$$n = ۸ \ , \ \bar{x} = ۸۰/۹۴ \ , \ \bar{y} = ۱۳۳/۰۶ \ , \ \sum_{i=۱}^{۸} x_i y_i = ۸۹۷۱۵/۸۸$$

$$\sum_{i=۱}^{۸} (x_i - \bar{x})^۲ = ۲۳۸۳/۴۷$$

$$\hat{b} = \frac{\sum_{i=۱}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=۱}^{n} (x_i - \bar{x})^۲} = \frac{۸۹۷۱۵/۸۸ - ۸(۸۰/۹۴)(۱۳۳/۰۶)}{۲۳۸۳/۴۷} = \frac{۳۵۵۶/۸۷}{۲۳۸۳/۴۷} = ۱/۴۹$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = ۱۳۳/۰۶ - (۱/۴۹)(۸۰/۹۴) = ۱۳۳/۰۶ - ۱۲۰/۶۰ = ۱۲/۴۶$$

خط برازش شده به صورت زیر می‌باشد:

$$\hat{y} = \hat{a} - \hat{b}x$$

$$\hat{y} = ۱۲/۴۶ + ۱/۴۹x$$

ب) می‌خواهیم فرض‌های زیر را آزمون کنیم:

$$\begin{cases} H_۰ : \beta = ۱/۳ \\ H_۱ : \beta \neq ۱/۳ \end{cases}$$

برای محاسبه‌ی آزمون چون $\sigma^۲$ مجهول است ابتدا مقدار $\hat{\sigma}^۲$ را محاسبه می‌کنیم:

$$\hat{\sigma}^۲ = \frac{۱}{n-۲} \sum_{i=۱}^{n} e_i^۲$$

$$e_i = y_i - \hat{y}_i = y_i - (۱۲/۴۶ + ۱/۴۹ x_i)$$

$$e_۱ = ۱۴۴/۴ - (۱۲/۴۶ + ۱/۴۹(۷۰/۳)) = ۲۷/۱۹$$

$$e_۲ = ۱۶۹/۳ - (۱۲/۴۶ + ۱/۴۹(۱۰۲)) = ۴/۸۶$$

$$e_۳ = ۱۰۶/۲ - (۱۲/۴۶ + ۱/۴۹(۶۲/۵)) = ۰/۶۱$$

$$e_۴ = ۱۲۵ - (۱۲/۴۶ + ۱/۴۹(۷۴/۸)) = ۱/۰۹$$

$$e_۵ = ۹۹/۸ - (۱۲/۴۶ + ۱/۴۹(۵۷/۹)) = ۱/۰۷$$

$$e_۶ = ۱۳۲/۱ - (۱۲/۴۶ + ۱/۴۹(۸۱/۶)) = -۱/۹۴$$

$$e_۷ = ۱۷۴/۲ - (۱۲/۴۶ + ۱/۴۹(۱۱۰/۴)) = -۲/۷۶$$

$$e_۸ = ۱۴۳/۵ - (۱۲/۴۶ + ۱/۴۹(۸۸)) = -۰/۰۸$$

$$\hat{\sigma}^۲ = \frac{۱}{n-۲} \sum_{i=۱}^{n} e_i^۲ = \frac{۱}{۶}(۷۷۷/۰۱) = ۱۲۹/۵$$

مقدار آماره‌ی آزمون برای آزمون فرض $H_۰ : \beta = ۱/۳$ در مقابل $H_۱ : \beta \neq ۱/۳$ به صورت زیر است:

$$T_۰ = \frac{\hat{b} - \beta}{\sqrt{\dfrac{\hat{\sigma}^۲}{\sum_{i=۱}^{n}(x_i - \bar{x})^۲}}} = \frac{۱/۴۹ - ۱/۳}{\sqrt{\dfrac{۱۲۹/۵}{۲۲۸۳/۴۷}}} = \frac{۰/۱۹}{۰/۲۳} = ۰/۸۳$$

مقدار جدول:

$$t(n-۲, \frac{\alpha}{۲}) = t(۶, ۰/۰۲۵) = ۲/۴۴۷$$

اگر $T_۰ > t(n-۲, \frac{\alpha}{۲})$ یا $T_۰ < -t(n-۲, \frac{\alpha}{۲})$ باشد فرض $H_۰$ رد می‌شود ولی در این جا چون $T_۰ = ۰/۸۳ < t(۶, ۰/۰۲۵) = ۲/۴۴۷$ پس فرض $H_۰$ رد نمی‌شود. یعنی $\beta = ۱/۳$ است.

۱۳ـ در جدول زیر نمرات تست هوش و برنامه‌نویسی ۱۲ دانشجو ثبت شده است:

الف) ضرایب خط رگرسیون نمونه را به دست آورید.

ب) اگر خطای مشاهدات دارای واریانس $\sigma_e^۲$ باشد آن را براورد کنید.

ج) فرض $H_۰ : \beta = ۰$ را در مقابل $H_۱ : \beta \neq ۰$ در سطح ۵ درصد آزمون کنید.

د) فرض $H_۰ : \alpha = ۰$ را در مقابل $H_۱ : \alpha \neq ۰$ با استفاده از فاصله‌ی اطمینان برای α آزمون کنید.

دانشجو (i)	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	
نمره‌تست‌هوش x_i	۶۵	۵۰	۵۵	۶۵	۵۵	۷۰	۶۵	۷۰	۵۵	۷۰	۵۰	۵۵	۷۲۵
نمره‌برنامه‌نویسی y_i	۸۵	۷۴	۷۶	۹۰	۸۵	۸۷	۹۴	۹۸	۸۱	۹۱	۷۶	۷۴	۱۰۱۱
$x_i y_i$	۵۵۲۵	۳۷۰۰	۴۱۸۰	۵۸۵۰	۴۶۷۵	۶۰۹۰	۶۱۱۰	۶۸۶۰	۴۴۵۵	۶۳۷۰	۳۸۰۰	۴۰۷۰	۶۱۶۸۵
$(x_i - \bar{x})^۲$	۲۰/۹۸	۱۰۸/۵۸	۲۹/۳۸	۲۰/۹۸	۲۹/۳۸	۹۱/۷۸	۲۰/۹۸	۹۱/۷۸	۲۹/۳۸	۹۱/۷۸	۱۰۸/۵۸	۲۹/۳۸	۶۷۲/۹۶
e_i	−۳/۳۳	−۰/۹۸	−۳/۴۳	۱/۶۷	۵/۵۷	−۵/۷۸	۵/۶۷	۵/۲۲	۱/۵۷	−۱/۷۸	۱/۰۲	−۵/۴۳	
$e_i^۲$	۱۱/۰۹	۰/۹۶	۱۱/۷۶	۲/۷۹	۳۱/۰۲	۳۳/۴۱	۳۲/۱۵	۲۷/۲۵	۲/۴۶	۳/۱۷	۱/۰۴	۲۹/۴۸	۱۸۶/۵۸

پاسخ: الف) اگر نمره‌های تست هوش را با x و نمره‌ی برنامه‌نویسی را با y نشان دهیم داریم:

$$n = ۱۲ \quad , \quad \bar{x} = ۶۰/۴۲ \quad , \quad \bar{y} = ۸۴/۲۵ \quad , \quad \sum_{i=۱}^{۱۲} x_i y_i = ۶۱۶۸۵$$

$$\sum_{i=۱}^{۱۲} (x_i - \bar{x})^۲ = ۶۷۲/۹۶$$

برای به دست آوردن ضرایب خط رگرسیون داریم:

$$\hat{b} = \frac{\sum_{i=۱}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=۱}^{n}(x_i - \bar{x})^۲} = \frac{۶۱۶۸۵ - ۱۲(۶۰/۴۲)(۸۴/۲۵)}{۶۷۲/۹۶} = \frac{۶۰۰/۳۸}{۶۷۲/۹۶} = ۰/۸۹$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = ۸۴/۲۵ - ۰/۸۹(۶۰/۴۲) = ۸۴/۲۵ - ۵۳/۷۷ = ۳۰/۴۸$$

خط برازش شده به صورت زیر است:

$$\hat{y} = \hat{a} + \hat{b}x$$

$$\hat{y} = ۳۰/۴۸ + ۰/۸۹x$$

ب) برای براورد $\sigma_e^۲$ ابتدا باید $\sum_{i=۱}^{n} e_i^۲$ را محاسبه کنیم بنابراین داریم:

$$e_i = y_i - \hat{y}_i = y_i - (۳۰/۴۸ + ۰/۸۹ x_i)$$

$$\sum_{i=۱}^{۱۲} e_i^۲ = ۱۸۶/۵۸$$

$$\hat{\sigma}_e^۲ = \frac{۱}{n}\sum_{i=۱}^{۱۲} e_i^۲ = \frac{۱}{۱۰}(۱۸۶/۵۸) = ۱۸/۶۵۸$$

ج) می‌خواهیم فرض‌های زیر را آزمون کنیم:

$$\begin{cases} H_0: \beta = ۰ \\ H_1: \beta \neq ۰ \end{cases}$$

مقدار آماره‌ی آزمون برای فرض‌های بالا عبارت است از:

$$T_0 = \frac{\hat{b} - \beta}{\sqrt{\dfrac{\hat{\sigma}_e^۲}{\sum_{i=۱}^{n} (x_i - \bar{x})^۲}}} = \frac{۰/۸۹ - ۰}{\sqrt{\dfrac{۱۸/۶۵۸}{۶۷۲/۹۶}}} = \frac{۰/۸۹}{۰/۱۷} = ۵/۲۴$$

مقدار جدول:

$$t(n-۲, \frac{\alpha}{۲}) = t(۱۰, ۰/۰۲۵) = ۲/۲۲۸$$

اگر $T_0 > t(n-۲, \frac{\alpha}{۲})$ یا $T_0 < -t(n-۲, \frac{\alpha}{۲})$ باشد فرض H_0 رد می‌شود بنابراین چون $T_0 = ۵/۲۴ > ۲/۲۲۸$ پس فرض H_0 رد می‌شود. یعنی $\beta \neq ۰$ است.

د) می‌خواهیم فرض‌های $\begin{cases} H_0: \alpha = ۰ \\ H_1: \alpha \neq ۰ \end{cases}$ را با استفاده از فاصله‌ی اطمینان برای α آزمون کنیم اگر مقدار $\alpha = ۰$ در فاصله‌ی اطمینان قرار گیرد فرض H_0 را قبول می‌کنیم. بنابراین داریم:

$$P\left(-t_{\frac{\alpha}{۲},(n-۲)} < \frac{\hat{a} - \alpha}{\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n} + \frac{\bar{X}^۲}{\sum (X_i - \bar{X})^۲}\right]}} < t_{\frac{\alpha}{۲},(n-۲)}\right) = ۱ - \alpha$$

پس فاصله‌ی اطمینان α-۱ درصد برای α به صورت زیر است:

$$\hat{a} - t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n} + \frac{\bar{X}^۲}{\sum (X_i - \bar{X})^۲}\right]} < \alpha < \hat{a} + t_{\frac{\alpha}{۲},(n-۲)}\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n} + \frac{\bar{X}^۲}{\sum (X_i - \bar{X})^۲}\right]}$$

که در آن مقادیر تحت فرض H_0 به صورت زیر است:

$$\hat{a} = ۳۰/۴۸$$

$$t_{\frac{\alpha}{۲},(n-۲)} = t_{۰/۰۲۵, ۱۰} = ۲/۲۲۸$$

$$\sqrt{\hat{\sigma}^۲\left[\frac{۱}{n} + \frac{\bar{X}^۲}{\sum (X_i - \bar{X})^۲}\right]} = \sqrt{۱۸/۶۵۸\left[\frac{۱}{۱۲} + \frac{(۶۰/۴۲)^۲}{۶۷۲/۹۶}\right]} = \sqrt{۱۰۲/۶۲} = ۱۰/۱۳$$

پس فاصله‌ی اطمینان ۰/۹۵ درصد برای α به صورت زیر است:

$$۳۰/۴۸ - (۲/۲۲۸ \times ۱۰/۱۳) < \alpha < ۳۰/۴۸ + (۲/۲۲۸ \times ۱۰/۱۳)$$

$$۷/۹۱ < \alpha < ۵۳/۰۵$$

چون مقدار $\alpha = ۰$ در فاصله‌ی اطمینان نیست پس فرض H_0 رد می‌شود یعنی $\alpha \neq ۰$ می‌باشد.

۱۴ـ جدول زیر تعداد ثبت‌نام کنندگان در یک دانشگاه را در ۷ سال گذشته نشان می‌دهد. با استفاده از روش حداقل مربعات منحنی‌ای به شکل $Y = \alpha\beta^X$ را براورد کرده و تعداد ثبت نام را برای ۵ سال اینده پیش‌بینی کنید.

سال (X)	۱	۲	۳	۴	۵	۶	۷
تعداد ثبت‌نام شدگان (Y)	۳۰۴	۳۴۱	۳۹۳	۴۵۷	۵۴۸	۶۷۰	۸۸۲

پاسخ: معادله‌ی $y_* = \alpha\beta^{x_*}$ با گرفتن لگاریتم

$$\log(y_*) = \log(\alpha) + x_* \log(\beta)$$

به معادله‌ی $y = a + bx$ تبدیل می‌شود که در آن $y = \log(y_*)$، $a = \log(\alpha)$، $x = x_*$ و $b = \log(\beta)$ می‌باشد.

با توجه به تبدیلات انجام شده می‌خواهیم خط برازش شده را به دست آوریم بنابراین داریم:

سال (x)	۱	۲	۳	۴	۵	۶	۷	۲۸
$y = \log(y_*)$	۲/۴۸	۲/۵۳	۲/۵۹	۲/۶۶	۲/۷۴	۲/۸۳	۲/۹۵	۱۸/۷۸
$x_i y_i$	۲/۴۸	۵/۰۶	۷/۷۷	۱۰/۶۴	۱۳/۷۰	۱۶/۹۸	۲۰/۶۵	۷۷/۲۸
$(x_i - \bar{x})^۲$	۹	۴	۱	۰	۱	۴	۹	۲۸

$$n = ۷ \quad , \quad \bar{x} = ۴ \quad , \quad \bar{y} = ۲/۶۸ \quad , \quad \sum_{i=۱}^{۷} x_i y_i = ۷۷/۲۸ \quad , \quad \sum_{i=۱}^{۷} (x_i - \bar{x})^۲ = ۲۸$$

$$\hat{b} = \frac{\sum_{i=۱}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=۱}^{n} (x_i - \bar{x})^۲} = \frac{۷۷/۲۸ - ۷(۴)(۲/۶۸)}{۲۸} = ۰/۰۸$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = ۲/۶۸ - (۰/۰۸)(۴) = ۲/۳۶$$

معادله‌ی رگرسیون تبدیل یافته

$$\hat{y} = \hat{a} + \hat{b}x$$

$$\hat{y} = ۲/۳۶ + ۰/۰۸x$$

معادله‌ی رگرسیون قبل از تبدیل به صورت زیر است:

$$y_* = \alpha\beta^{x_*} \rightarrow \hat{y}_* = (۲۲۹/۰۹)(۱/۲)^{x_*}$$

که در آن:

$$\alpha = ۱۰^{\hat{a}} = ۱۰^{۲/۳۶} = ۲۲۹/۰۹$$

$$\beta = ۱۰^{\hat{b}} = ۱۰^{۰/۰۸} = ۱/۲$$

می‌خواهیم تعداد ثبت‌نام را برای ۵ سال آینده پیش‌بینی کنیم داریم:

اگر $x=۸$،

$$\hat{y}_* = (۲۲۹/۰۹)(۱/۲)^۸ = ۹۸۵/۰۵$$

اگر $x=۹$،

$$\hat{y}_* = (۲۲۹/۰۹)(۱/۲)^۹ = ۱۱۸۲/۰۵$$

اگر $x=۱۰$،

$$\hat{y}_* = (۲۲۹/۰۹)(۱/۲)^{۱۰} = ۱۴۱۸/۴۶$$

اگر $x=۱۱$،

$$\hat{y}_* = (۲۲۹/۰۹)(۱/۲)^{۱۱} = ۱۷۰۲/۱۶$$

اگر $x=۱۲$،

$$\hat{y}_* = (۲۲۹/۰۹)(۱/۲)^{۱۲} = ۲۰۴۲/۵۹$$

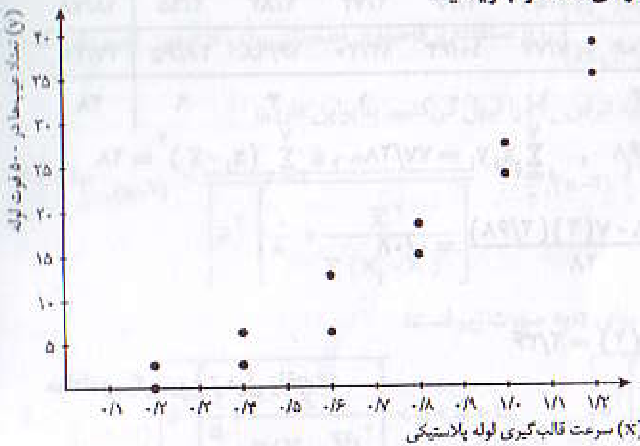
۱۵ـ یک مهندس تولید رابطه‌ی بین سرعت قالب‌گیری لوله‌ی پلاستیکی (X) و تعداد عیب‌ها در ۵۰۰ فوت لوله (Y) را که با ماشینی جدید تولید می‌شود بررسی کرده است. ۱۲ امتحان با ماشین انجام شده و برد سرعت‌های قالب‌گیری از ۰/۲ فوت تا ۱/۲ فوت در ثانیه بوده است. در هر نوبت امتحان، ۵۰۰ فوت لوله تولید شده است نتایج به شرح زیر بوده‌اند:

i	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
X_i	۰/۲	۰/۲	۰/۴	۰/۴	۰/۶	۰/۶	۰/۸	۰/۸	۱/۰	۱/۰	۱/۲	۱/۲
Y_i	۳	۰	۶	۳	۱۲	۷	۱۹	۱۵	۲۴	۲۹	۳۹	۳۴

الف) نمودار پراکنش را رسم کنید و ضرایب خط رگرسیون $\hat{y} = \hat{a} + \hat{b}x$ را به دست آورید.

ب) پس از تحلیل بیشتر، مهندس تصمیم می‌گیرد که از تبدیل $Y' = \sqrt{Y}$ استفاده کند. نمودار پراکنش Y' را در مقابل X رسم کنید. آیا در این جا تبدیل فوق در خطی کردن رابطه‌ی بین متغیرها موفق به نظر می‌رسد، خط برازش شده را بنویسید.

پاسخ: الف)



i	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	
x_i	۰/۲	۰/۲	۰/۴	۰/۴	۰/۶	۰/۶	۰/۸	۰/۸	۱/۰	۱/۰	۱/۲	۱/۲	۸۴
y_i	۳	۰	۶	۳	۱۲	۷	۱۹	۱۵	۲۴	۲۹	۳۹	۳۴	۱۹۱
$x_i y_i$	۰/۶	۰	۲/۴	۱/۲	۷/۲	۴/۲	۱۵/۲	۱۲	۲۴	۲۹	۴۶/۸	۴۰/۸	۱۸۳/۴
$(x_i-\bar{x})^2$	۰/۲۵	۰/۲۵	۰/۰۹	۰/۰۹	۰/۰۱	۰/۰۱	۰/۰۱	۰/۰۱	۰/۰۹	۰/۰۹	۰/۲۵	۰/۲۵	۱/۴

$n = ۱۲ \quad , \quad \bar{x} = ۰/۷ \quad , \quad \bar{y} = ۱۵/۹۲ \quad , \quad \sum_{i=۱}^{۱۲} x_i y_i = ۱۸۳/۴ \quad , \quad \sum_{i=۱}^{۱۲} (x_i - \bar{x})^2 = ۱/۴$

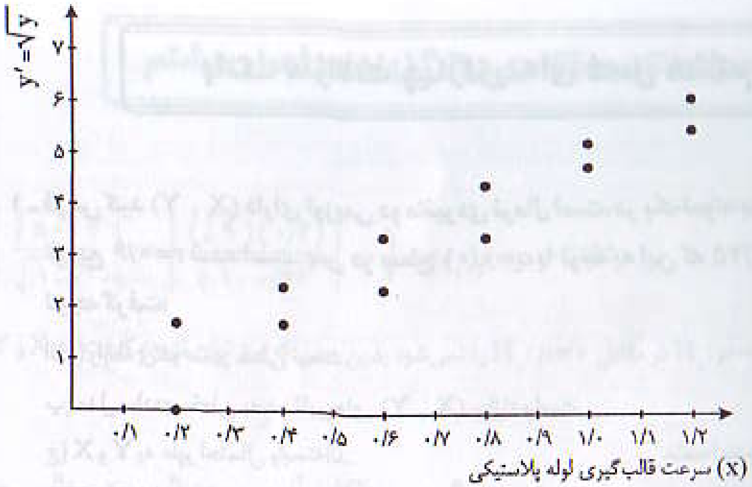
$$\hat{b} = \frac{\sum_{i=۱}^{۱۲} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=۱}^{n} (x_i - \bar{x})^2} = \frac{۱۸۳/۴ - ۱۲(۰/۷)(۱۵/۹۲)}{۱/۴} = ۳۵/۴۸$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = ۱۵/۹۲ - (۳۵/۴۸)(۰/۷) = ۱۵/۹۲ - ۲۴/۸۴ = -۸/۹۲$$

پس خط برازش شده به صورت مقابل است:

$$\hat{y} = -۸/۹۲ + ۳۵/۴۸x$$

ب) حال می‌خواهیم نمودار پراکنش $y' = \sqrt{y}$ را در مقابل x رسم کنیم.



i	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	
x_i	۰/۲	۰/۲	۰/۴	۰/۴	۰/۶	۰/۶	۰/۸	۰/۸	۱/۰	۱/۰	۱/۲	۱/۲	۸۴
$y'_i = \sqrt{y_i}$	۱/۷۳	۰	۲/۴۵	۱/۷۳	۳/۴۶	۲/۶۵	۴/۳۶	۳/۸۷	۴/۹	۵/۳۹	۶/۲۴	۵/۸۳	۴۲/۶۱
$x_i y'_i$	۰/۳۵	۰	۰/۹۸	۰/۶۹	۲/۰۸	۱/۵۹	۳/۴۹	۳/۱۰	۴/۹	۵/۳۹	۷/۴۹	۷	۳۷/۰۶
$(x_i-\bar{x})^2$	۰/۲۵	۰/۲۵	۰/۰۹	۰/۰۹	۰/۰۱	۰/۰۱	۰/۰۱	۰/۰۱	۰/۰۹	۰/۰۹	۰/۲۵	۰/۲۵	۱/۴

$n = ۱۲ \quad , \quad \bar{x} = ۰/۷ \quad , \quad \bar{y} = ۳/۵۵ \quad , \quad \sum_{i=۱}^{۱۲} x_i y_i = ۳۷/۰۶ \quad , \quad \sum_{i=۱}^{۱۲} (x_i - \bar{x})^2 = ۱/۴$

$$\hat{b} = \frac{\sum_{i=۱}^{۱۲} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=۱}^{n} (x_i - \bar{x})^2} = \frac{۳۷/۰۶ - ۱۲(۰/۷)(۳/۵۵)}{۱/۴} = ۵/۱۷$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = ۳/۵۵ - (۵/۱۷)(۰/۷) = ۳/۵۵ - ۳/۶۲ = -۰/۰۷$$

پس خط رگرسیون تبدیل یافته به صورت زیر است:

$$\hat{y} = -۰/۰۷ + ۵/۱۷x$$

با مقایسه کردن معادله‌ی رگرسیون اولیه و معادله‌ی رگرسیون تبدیل یافته و مقادیر $\hat{\alpha}$ و $\hat{\beta}$ متوجه خواهیم شد این تبدیل در خطی کردن رابطه بین متغیرها موفق به نظر می‌رسد.