$$P(X>n+m)=P(X>n)P(X>m).$$

若把 n 换成 $n-1$ 仍有

$$P(X>n+m-1)=P(X>n-1)P(X>m).$$

上两式相减可得

$$P(X=n+m)=P(X=n)P(X>m).$$

若取 $n=m=1$,并设 $P(X=1)=p$,则有

$$P(X=2)=p(1-p).$$

若取 $n=2,m=1$,可得

$$P(X=3)=P(X=2)P(X>1)=p(1-p)^2.$$

若令 $P(X=k)=p(1-p)^{k-1}$,则用数学归纳法可推得

$$P(X=k+1)=P(X=k)P(X>1)=p(1-p)^k,\quad k=0,1,\cdots.$$

这表明 X 的分布就是几何分布.

§2.5　常用连续分布

内容概要

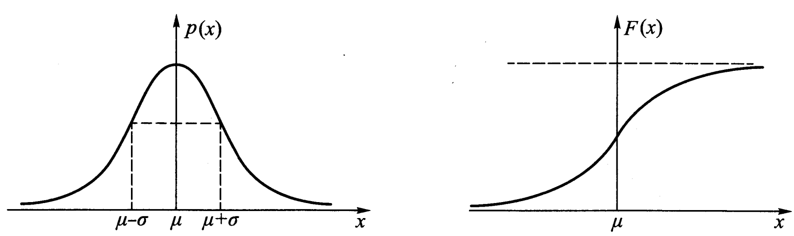
1. 正态分布

(1) 若随机变量 X 的密度函数和分布函数(如图 2.10) 分别为

$$p(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\qquad -\infty<x<\infty,$$

$$F(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{x}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t,\qquad -\infty<x<\infty,$$

则称 X 服从正态分布,记作 $X\sim N(\mu,\sigma^2)$,其中参数 $-\infty<\mu<\infty,\sigma>0$.



(a) 密度函数 $p(x)$　　(b) 分布函数 $F(x)$

图 2.10

(2) 背景:一个变量若是由大量微小的、独立的随机因素的叠加结果,则此变量一定是正态变量(服从正态分布的变量).测量误差就是由量具偏差、测量环境的影响、测量技术的影响、测量人员的心理影响等随机因素叠加而成的,所以测量误差常认为服从正态分布.

(3) 关于参数 μ

• μ 是正态分布的数学期望,即 $E(X)=\mu$,称 μ 为正态分布的位置参数.

• μ 是正态分布的对称中心,在 μ 的左侧和 $p(x)$ 下的面积为 0.5;在 μ 的右侧和 $p(x)$ 下的面积也为 0.5,所以 μ 也是正态分布的中位数(见后面 §2.7).

• 若 $X \sim N(\mu,\sigma^2)$,则 X 在离 μ 愈近取值的可能性愈大,离 μ 愈远取值的可能性愈小.

关于参数 σ

• σ^2 是正态分布的方差,即 $\mathrm{Var}(X)=\sigma^2$.

• σ 是正态分布的标准差,σ 愈小,正态分布愈集中;σ 愈大,正态分布愈分散.σ 又称为正态分布的尺度参数.

• 若随机变量 $X \sim N(\mu,\sigma^2)$,则其密度函数 $p(x)$ 在 $\mu \pm \sigma$ 处有两个拐点.

(4) 称 $\mu=0,\sigma=1$ 时的正态分布 $N(0,1)$ 为**标准正态分布**.记 U 为标准正态变量,$\varphi(u)$ 和 $\Phi(u)$ 为标准正态分布的密度函数和分布函数.$\varphi(u)$ 和 $\Phi(u)$ 满足:

• $\varphi(-u)=\varphi(u)$;

• $\Phi(-u)=1-\Phi(u)$.对 $u>0,\Phi(u)$ 的值有表可查.

(5) 标准化变换:若随机变量 $X \sim N(\mu,\sigma^2)$,则 $U=(X-\mu)/\sigma \sim N(0,1)$,其中 $U=(X-\mu)/\sigma$ 称为 X 的标准化变换.

(6) 若随机变量 $X \sim N(\mu,\sigma^2)$,则对任意实数 a 与 b,有

$$P(X \leqslant b)=\Phi\left(\frac{b-\mu}{\sigma}\right),$$

$$P(a<X)=1-\Phi\left(\frac{a-\mu}{\sigma}\right),$$

$$P(a<X \leqslant b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right),$$

可见,涉及正态变量的概率计算,一般是化为标准正态变量再查表获得.

(7) 正态分布的 3σ 原则:设随机变量 $X \sim N(\mu,\sigma^2)$,则

$$P(|X-\mu|<k\sigma)=\Phi(k)-\Phi(-k)=\begin{cases}0.682\,6, & k=1,\\ 0.954\,5, & k=2,\\ 0.997\,3, & k=3.\end{cases}$$

2. 均匀分布

(1) 若随机变量 X 的密度函数和分布函数(如图 2.11)分别为

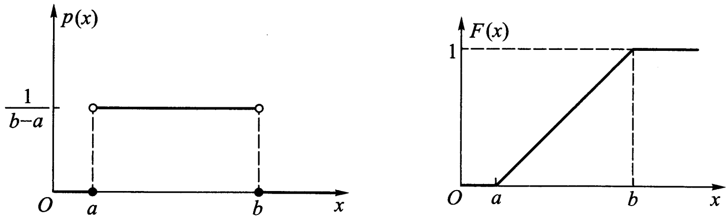
$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其他},\end{cases} \qquad F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a \leqslant x<b,\\ 1, & x \geqslant b.\end{cases}$$

则称 X 服从区间 (a,b) 上的均匀分布,记作 $X \sim U(a,b)$.

(2) 背景:向区间 (a,b) 随机投点,落点坐标 X 一定服从均匀分布 $U(a,b)$.这里"随机投点"是指:点落在任意相等长度的小区间上的可能性是相等的.

(3) 均匀分布 $U(a,b)$ 的数学期望和方差分别是

$$E(X)=\frac{a+b}{2}, \qquad \mathrm{Var}(X)=\frac{(b-a)^2}{12}.$$



(a) 密度函数 $p(x)$　　(b) 分布函数 $F(x)$

图 2.11

（4）称区间$(0,1)$上的均匀分布 $U(0,1)$ 为标准均匀分布，它是导出其他分布随机数的桥梁.

3. 指数分布

（1）若随机变量 X 的密度函数（如图 2.12）和分布函数分别为

$$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x \geqslant 0,\\ 0, & x<0,\end{cases}\qquad F(x)=\begin{cases}1-e^{-\lambda x}, & x \geqslant 0,\\ 0, & x<0,\end{cases}$$

则称 X 服从指数分布，记作 $X \sim Exp(\lambda)$，其中参数$\lambda > 0$.

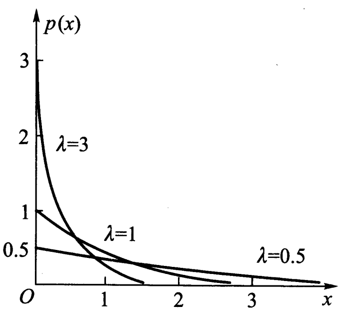


图 2.12

（2）背景：若一个元器件（或一台设备，或一个系统）遇到外来冲击时即告失效，则首次冲击来到的时间 X（寿命）服从指数分布.很多产品的寿命可认为服从或近似服从指数分布.

（3）指数分布 $Exp(\lambda)$ 的数学期望和方差分别为

$$E(X)=\frac{1}{\lambda},\qquad \mathrm{Var}(X)=\frac{1}{\lambda^2}.$$

（4）指数分布的无记忆性：　若随机变量 $X \sim Exp(\lambda)$，则对任意 $s>0, t>0$，有

$$P(X>s+t \mid X>s)=P(X>t).$$

4. 伽马分布

（1）**伽马函数**　称 $\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}e^{-x}dx$ 为伽马函数，其中参数 $\alpha>0$.伽马函数具有如下性质：

（a）$\Gamma(1)=1$；

（b）$\Gamma(1/2)=\sqrt{\pi}$；

（c）$\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$；

（d）$\Gamma(n+1)=n\Gamma(n)=n!$　（n 为自然数）.

（2）**伽马分布**　若随机变量 X 的密度函数（如图 2.13）为

$$p(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x \geqslant 0,\\ 0, & x<0,\end{cases}$$

则称 X 服从伽马分布，记作 $X \sim Ga(\alpha,\lambda)$，其中 $\alpha>0$ 为形状参数，$\lambda>0$ 为尺度参数.

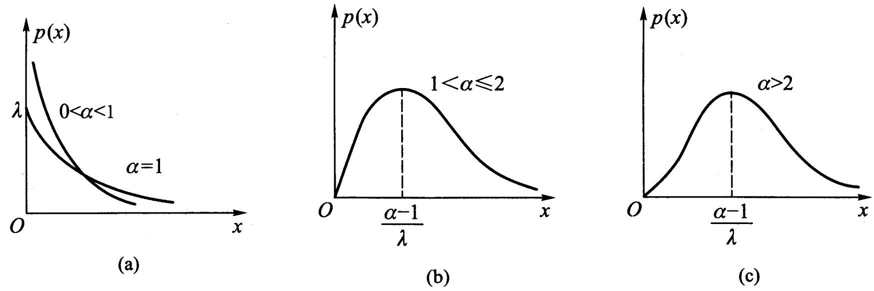


图 2.13

（3）背景：若一个元器件（或一台设备，或一个系统）能抵挡一些外来冲击，但遇到第 k 次冲击时即告失效，则第 k 次冲击来到的时间 X（寿命）服从形状参数为 k 的伽马分布 $Ga(k,\lambda)$.

（4）伽马分布 $Ga(\alpha,\lambda)$ 的数学期望和方差分别为

$$E(X)=\frac{\alpha}{\lambda},\qquad \mathrm{Var}(X)=\frac{\alpha}{\lambda^2}.$$

（5）伽马分布的两个特例：

（a）$\alpha=1$ 时的伽马分布就是指数分布，即 $Ga(1,\lambda)=Exp(\lambda)$.

（b）称 $\alpha=n/2,\lambda=1/2$ 时的伽马分布为自由度为 n 的 **χ^2（卡方）分布**，记为 $\chi^2(n)$，其密度函数为

$$p(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}e^{-\frac{x}{2}}x^{\frac{n}{2}-1}, & x\geqslant 0,\\ 0, & x<0.\end{cases}$$

$\chi^2(n)$ 分布的期望和方差分别为 $E(X)=n,\ \mathrm{Var}(X)=2n.$

（6）若形状参数为整数 k，则伽马变量可以表示成 k 个独立同分布的指数变量之和，即若 $X\sim Ga(k,\lambda)$，则 $X=X_1+X_2+\cdots+X_k$，其中 $X_1,X_2,\cdots,X_k$ 是相互独立且都服从指数分布 $Exp(\lambda)$ 的随机变量，见图 2.14，其中“×”表示冲击来到的时间.

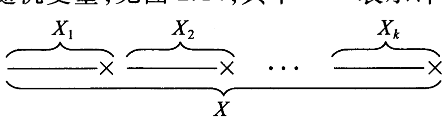


图 2.14

5. 贝塔分布

（1）**贝塔函数**　称 $\mathrm{B}(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x$ 为贝塔函数，其中参数 $a>0,b>0$. 贝塔函数具有如下性质：

（a）$\mathrm{B}(a,b)=\mathrm{B}(b,a)$；　（b）$\mathrm{B}(a,b)=\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

（2）**贝塔分布**　若随机变量 X 的密度函数（如图 2.15）为

$$p(x)=\begin{cases}\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

则称 X 服从贝塔分布,记作 $X\sim Be(a,b)$,其中 $a>0,b>0$ 都是形状参数.

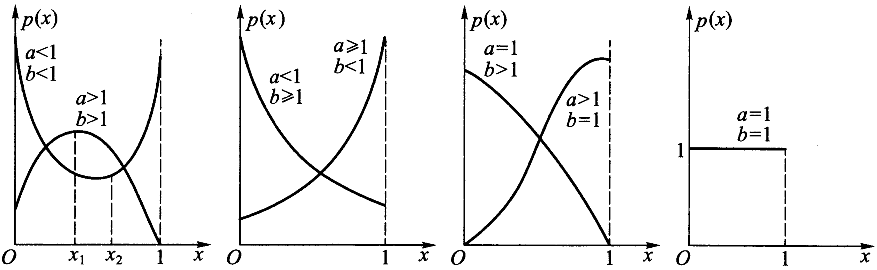


图 2.15

(3) 背景:很多比率,如产品的不合格品率、机器的维修率、某商品的市场占有率、射击的命中率等都是在区间(0,1)上取值的随机变量,贝塔分布 $Be(a,b)$ 可供描述这些随机变量之用,而在应用中,可调节 a 与 b 以适应实际中的要求.

(4) 贝塔分布 $Be(a,b)$ 的数学期望和方差分别为

$$E(X)=\frac{a}{a+b},\qquad \mathrm{Var}(X)=\frac{ab}{(a+b)^2(a+b+1)}.$$

(5) $a=b=1$ 时的贝塔分布就是区间(0,1)上的均匀分布,即 $Be(1,1)=U(0,1)$.

6. 常用连续分布表

名称与记号	密度函数 $p(x)$	期望	方差
正态分布 $N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},-\infty<x<\infty$	μ	σ^2
均匀分布 $U(a,b)$	$\frac{1}{b-a},a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $Exp(\lambda)$	$\lambda e^{-\lambda x},x\geqslant 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
伽马分布 $Ga(\alpha,\lambda)$	$\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x},x\geqslant 0$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\chi^2(n)$ 分布	$\frac{x^{n/2-1}e^{-x/2}}{\Gamma(n/2)2^{n/2}},x\geqslant 0$	n	$2n$
贝塔分布 $Be(a,b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},0<x<1$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$
对数正态分布 $LN(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\},x>0$	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$
柯西分布 $Cau(\mu,\lambda)$	$\frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2},-\infty<x<\infty$	不存在	不存在

续表

名称与记号	密度函数 $p(x)$	期望	方差
韦布尔分布 $W(m,\eta)$	$p(x)=F'(x),F(x)=1-\exp\left\{-\left(\frac{x}{\eta}\right)^m\right\}$ 或 $p(x)=\frac{m}{\eta}\left(\frac{x}{\eta}\right)^{m-1}\exp\left\{-\left(\frac{x}{\eta}\right)^m\right\},x>0$	$\eta\Gamma\left(1+\frac{1}{m}\right)$	$\eta^2\left[\Gamma\left(1+\frac{2}{m}\right)-\Gamma^2\left(1+\frac{1}{m}\right)\right]$

7. 要记住常用分布的数学期望与方差，还要弄清它们与分布中参数的关系，因为不少分布可由其数学期望与方差确定.而一个分布就是一个概率模型.

习题与解答 2.5

1. 设随机变量 X 服从区间 $(2,5)$ 上的均匀分布，求对 X 进行 3 次独立观测中，至少有 2 次的观测值大于 3 的概率.

解 在一次观测中，观测值大于 3 的概率为

$$p=P(X>3)=\int_3^5\frac{1}{3}\mathrm{d}x=\frac{2}{3}.$$

设 Y 为此种观测 $(X>3)$ 的次数，则 $Y\sim b(3,2/3)$，由此得

$$P(Y\geqslant 2)=1-P(Y=0)-P(Y=1)=1-\left(\frac{1}{3}\right)^3-3\times\frac{2}{3}\times\left(\frac{1}{3}\right)^2=\frac{20}{27}.$$

2. 在 $(0,1)$ 上任取一点记为 X，试求 $P\left(X^2-\frac{3}{4}X+\frac{1}{8}\geqslant 0\right)$.

解 由 $x^2-3x/4+1/8=0$ 解得 $x_1=0.25,x_2=0.5$.因为 $X\sim U(0,1)$，又因为二次函数 $y=x^2-3x/4+1/8$ 是开口向上的，故有

$$\left\{X^2-\frac{3}{4}X+\frac{1}{8}\geqslant 0\right\}=\{0\leqslant X\leqslant 0.25\}\cup\{0.5\leqslant X\leqslant 1\},$$

所以

$$P\left(X^2-\frac{3}{4}X+\frac{1}{8}\geqslant 0\right)=\int_0^{0.25}\mathrm{d}x+\int_{0.5}^1\mathrm{d}x=0.75.$$

3. 设 K 服从 $(1,6)$ 上的均匀分布，求方程 $x^2+Kx+1=0$ 有实根的概率.

解 方程 $x^2+Kx+1=0$ 有实根的充要条件是

$$\{K^2-4\geqslant 0\}=\{K\leqslant -2\}\cup\{K\geqslant 2\}.$$

而 $K\sim U(1,6)$，因此所求概率为

$$P(K\leqslant -2)+P(K\geqslant 2)=0+\int_2^6\frac{1}{5}\mathrm{d}x=\frac{4}{5}.$$

4. 若随机变量 $K\sim N(\mu,\sigma^2)$，而方程 $x^2+4x+K=0$ 无实根的概率为 0.5，试求 μ.

解 方程 $x^2+4x+K=0$ 无实根等价于 $16-4K<0$，所以由题意知

$$0.5=P(16-4K<0)=P(K>4)=1-\Phi\left(\frac{4-\mu}{\sigma}\right).$$

由此得知 $\mu=4$.

5. 设流经一个2 Ω电阻上的电流强度I是一个随机变量,它均匀分布在9 A至11 A之间.试求此电阻上消耗的平均功率,其中功率 $W = 2I^2$.

解　因为 $I \sim U(9,11)$,所以平均功率为

$$E(W) = E(2I^2) = \int_9^{11} 2x^2 \frac{1}{2}\mathrm{d}x = \frac{1}{3}x^3 \Big|_9^{11} = \frac{602}{3}.$$

6. 某种圆盘的直径在区间(a,b)上服从均匀分布,试求此种圆盘的平均面积.

解　记 X 为圆盘的直径,则圆盘的面积为 $Y = \pi X^2/4$,所以平均面积为

$$E(Y) = \frac{\pi}{4}E(X^2) = \frac{\pi}{4}\left[\frac{(b-a)^2}{12} + \frac{(a+b)^2}{4}\right] = \frac{\pi}{12}(a^2 + b^2 + ab).$$

7. 设某种商品每周的需求量 X 服从区间$(10,30)$上均匀分布,而商店进货数为区间$(10,30)$中的某一整数,商店每销售 1 单位商品可获利 500 元;若供大于求则降价处理,每处理 1 单位商品亏损 100 元;若供不应求,则可从外部调剂供应,此时每 1 单位商品仅获利 300 元.为使商店所获利润期望值不少于 9 280 元,试确定最少进货量.

解　设进货量为 a,则利润为

$$g(X) = \begin{cases} 500X - 100(a - X), & 10 \leqslant X \leqslant a, \\ 500a + 300(X - a), & a < X \leqslant 30 \end{cases}$$
$$= \begin{cases} 600X - 100a, & 10 \leqslant X \leqslant a, \\ 300X + 200a, & a < X \leqslant 30. \end{cases}$$

所以平均利润为

$$E(g(X)) = \int_{10}^{30} g(x)\frac{1}{20}\mathrm{d}x = \frac{1}{20}\int_{10}^{a}(600x - 100a)\mathrm{d}x + \frac{1}{20}\int_{a}^{30}(300x + 200a)\mathrm{d}x$$
$$= -7.5a^2 + 350a + 5\,250.$$

按照题意要求有

$$-7.5a^2 + 350a + 5\,250 \geqslant 9\,280 \quad \text{即} \quad -7.5a^2 + 350a - 4\,030 \geqslant 0,$$

解得

$$20\frac{2}{3} \leqslant a \leqslant 26,$$

因此最少进货为 21 单位.

8. 统计调查表明,英格兰在1875年至1951年期间在矿山发生10人或10人以上死亡的两次事故之间的时间 T(以日计)服从均值为 241 的指数分布.试求 $P(50 < T < 100)$.

解　$P(50 < T < 100) = F(100) - F(50) = \mathrm{e}^{-50/241} - \mathrm{e}^{-100/241} = 0.152\,3.$

9. 若一次电话通话时间X(以 min 计)服从参数为0.25 的指数分布,试求一次通话的平均时间.

解　因为 $X \sim Exp(\lambda)$,其中 $\lambda = 0.25$.所以 $E(X) = 1/\lambda = 1/0.25 = 4$ min.

10. 某种设备的使用寿命 X(以年计)服从指数分布,其平均寿命为 4 年.制造此种设备的厂家规定,若设备在使用一年之内损坏,则可以予以调换.如果设备制造厂每售出一台设备可赢利 100 元,而调换一台设备制造厂需花费 300 元.试求每台设备的平均利润.

解　令

$$Y=\begin{cases}0, & \text{设备在使用一年之内不损坏},\\ 1, & \text{设备在使用一年之内损坏},\end{cases}$$

即 Y 是一台设备在使用一年之内损坏的台数,显然 $Y\sim b(1,p)$,其中

$$p=P(\text{设备在使用一年之内损坏})=P(X\leqslant 1)$$
$$=\int_0^1\frac{1}{4}\mathrm{e}^{-\frac{x}{4}}\mathrm{d}x=1-\mathrm{e}^{-0.25}=0.221\ 2.$$

因为每台设备的利润为 $Z=100-300Y$,所以每台设备的平均利润为

$$E(Z)=100-300E(Y)=100-300\times 0.221\ 2=33.64(\text{元})$$

11. 设顾客在某银行的窗口等待服务的时间 X(以 min 计)服从指数分布,其密度函数为

$$p(x)=\begin{cases}\frac{1}{5}\mathrm{e}^{-\frac{x}{5}}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

某顾客在窗口等待服务,若超过 10 min 他就离开.他一年要到银行 5 次,以 Y 表示一年内他未等到服务而离开窗口的次数,试求 $P(Y\geqslant 1)$.

解 因为 $Y\sim b(5,p)$,其中 $p=P(X>10)=\mathrm{e}^{-2}$,所以得

$$P(Y\geqslant 1)=1-P(Y=0)=1-(1-p)^5=1-(1-\mathrm{e}^{-2})^5=0.516\ 7.$$

12. 某仪器装了3个独立工作的同型号电子元件,其寿命 X(以 h 计)都服从同一指数分布,密度函数为

$$p(x)=\begin{cases}\frac{1}{600}\mathrm{e}^{-\frac{x}{600}}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

试求:此仪器在最初使用的 200 h 内,至少有一个此种电子元件损坏的概率.

解 设 Y 为仪器在最初使用的 200 h 内,损坏的元件个数,则 $Y\sim b(3,p)$,其中

$$p=\int_0^{200}\frac{1}{600}\mathrm{e}^{-\frac{x}{600}}\mathrm{d}x=1-\mathrm{e}^{-\frac{1}{3}},$$

所以至少有一个电子元件损坏的概率为

$$P(Y\geqslant 1)=1-P(Y=0)=1-(1-p)^3=1-[1-(1-\mathrm{e}^{-\frac{1}{3}})]^3$$
$$=1-\mathrm{e}^{-1}=0.632\ 1.$$

13. 设随机变量 X 的密度函数为

$$p(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0\end{cases}\qquad(\lambda>0).$$

试求 k,使得 $P(X>k)=0.5$.

解 因为 $0.5=P(X>k)=\int_k^{\infty}\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x=\mathrm{e}^{-\lambda k}$,由此解得 $k=\dfrac{\ln 2}{\lambda}$.

14. 设随机变量 X 的密度函数为

$$p(x)=\begin{cases}1/3, & 0\leqslant x\leqslant 1,\\ 2/9, & 3\leqslant x\leqslant 6,\\ 0, & \text{其他}.\end{cases}$$

若 $P(X \geqslant k)=2/3$,试求 k 的取值范围.

解　由题设条件 $2/3=P(X \geqslant k)=1-P(X<k)$,知 $F(k)=1/3$.又由 $p(x)$ 得分布函数如下

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x}{3}, & 0 \leqslant x<1,\\ \dfrac{1}{3}, & 1 \leqslant x<3,\\ \dfrac{2}{9}x-\dfrac{1}{3}, & 3 \leqslant x<6,\\ 1, & x \geqslant 6.\end{cases}$$

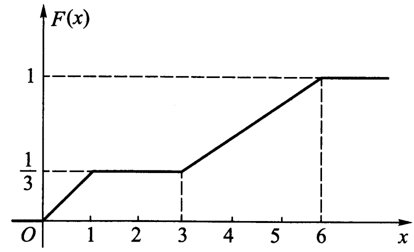


图 2.16

$F(x)$ 的图形如图 2.16.

由此得 $1 \leqslant k \leqslant 3$.

15. 写出以下正态分布的均值和标准差:

$$p_1(x)=\frac{1}{\sqrt{\pi}}e^{-(x^2+4x+4)},\quad p_2(x)=\sqrt{\frac{2}{\pi}}e^{-2x^2},\quad p_3(x)=\frac{1}{\sqrt{\pi}}e^{-x^2}.$$

解　对 $p_1(x)$ 有

$$p_1(x)=\frac{1}{\sqrt{2\pi}(1/\sqrt{2})}\exp\left\{-\frac{(x+2)^2}{2(1/\sqrt{2})^2}\right\},$$

所以 $p_1(x)$ 的均值 $\mu_1=-2$,标准差 $\sigma_1=1/\sqrt{2}$.

对 $p_2(x)$ 有

$$p_2(x)=\frac{1}{\sqrt{2\pi}(1/2)}\exp\left\{-\frac{(x-0)^2}{2(1/2)^2}\right\},$$

所以 $p_2(x)$ 的均值 $\mu_2=0$,标准差 $\sigma_2=1/2$.

对 $p_3(x)$ 有

$$p_3(x)=\frac{1}{\sqrt{2\pi}(1/\sqrt{2})}\exp\left\{-\frac{(x-0)^2}{2(1/\sqrt{2})^2}\right\},$$

所以 $p_3(x)$ 的均值 $\mu_3=0$,标准差 $\sigma_3=1/\sqrt{2}$.

16. 某地区 18 岁女青年的血压 X(收缩压,以 mmHg 计)服从 $N(110,12^2)$.试求该地区 18 岁女青年的血压在 100 至 120 的可能性有多大?

解

$$\begin{aligned}P(100<X<120)&=\Phi\left(\frac{120-110}{12}\right)-\Phi\left(\frac{100-110}{12}\right)\\&=\Phi\left(\frac{5}{6}\right)-\Phi\left(-\frac{5}{6}\right)\\&=2\Phi\left(\frac{5}{6}\right)-1=0.595\ 0.\end{aligned}$$

其中 $\Phi(5/6)=\Phi(0.833)=0.797\ 5$ 是用内插法得到的.

17. 某地区成年男子的体重 X(以 kg 计)服从正态分布 $N(\mu,\sigma^2)$.若已知 $P(X \leqslant 70)=0.5$,$P(X \leqslant 60)=0.25$.

(1) 求 μ 与 σ 各为多少?

(2) 若在这个地区随机地选出5名成年男子,问其中至少有两人体重超过65 kg的概率是多少?

解 (1) 由

$$0.5=P(X \leqslant 70)=\Phi\left(\frac{70-\mu}{\sigma}\right),$$

知

$$\frac{70-\mu}{\sigma}=0,$$

由此解得 $\mu=70$.又由

$$0.25=P(X \leqslant 60)=\Phi\left(\frac{60-70}{\sigma}\right)=\Phi\left(\frac{-10}{\sigma}\right)=1-\Phi\left(\frac{10}{\sigma}\right),$$

即 $0.75=\Phi(10/\sigma)$,查表知 $10/\sigma=0.675$,由此解得 $\sigma=14.81$.

(2) 记 Y 为选出的 5 名成年男子中体重超过 65 kg 的人数,则 $Y \sim b(5,p)$,其中

$$p=P(X>65)=\Phi\left(\frac{70-65}{14.81}\right)=\Phi(0.337\ 6)=0.632\ 4,$$

所以“5 名中至少有两人体重超过 65 kg”的概率为

$$P(Y \geqslant 2)=1-0.367\ 6^5-5\times 0.367\ 6^4\times 0.632\ 4=0.94.$$

18. 由某机器生产的螺栓的长度(以 cm 计)服从正态分布 $N(10.05,0.06^2)$,若规定长度在范围(10.05 ± 0.12)cm 内为合格品,求螺栓不合格的概率.

解 记螺栓的长度为 X,则

$$\begin{aligned}P(\text{螺栓不合格})&=1-P(10.05-0.12 \leqslant X \leqslant 10.05+0.12)\\&=2-2\Phi(0.12/0.06)=2-2\times 0.977\ 2=0.045\ 6.\end{aligned}$$

19. 某地抽样调查结果表明,考生的外语成绩(以百分制计)近似地服从 $\mu=72$ 的正态分布,已知 96 分以上的人数占总数的 2.3%,试求考生的成绩大于等于 60 分的概率.

解 记 X 为考生的外语成绩,由题设条件知 $X \sim N(72,\sigma^2)$,其中 σ 未知,但由题设条件知

$$0.023=P(X>96)=1-\Phi\left(\frac{96-72}{\sigma}\right),\qquad \text{即}\quad \Phi\left(\frac{24}{\sigma}\right)=0.977,$$

因此查表知 $24/\sigma=2$,由此解得 $\sigma=12$.从而得 $X \sim N(72,12^2)$,由此所求概率为

$$P(X \geqslant 60)=\Phi(1)=0.841\ 3.$$

20. 设随机变量 $X \sim N(3,2^2)$,(1) 求 $P(2<X \leqslant 5)$;(2) 求 $P(|X|>2)$;(3) 确定 c 使得 $P(X>c)=P(X<c)$.

解 (1) $P(2<X \leqslant 5)=\Phi(1)-\Phi(-0.5)=\Phi(1)-1+\Phi(0.5)=0.532\ 8$.

(2) $$\begin{aligned}P(|X|>2)&=P(X>2)+P(X<-2)=1-\Phi(-0.5)+\Phi(-2.5)\\&=\Phi(0.5)+1-\Phi(2.5)=0.697\ 7.\end{aligned}$$

(3) 因为 $1=P(X>c)+P(X<c)$,所以由题设条件 $P(X>c)=P(X<c)$ 得 $P(X<c)=0.5$,进而有 $(c-3)/2=0$,由此得 $c=3$.

21. 设随机变量 $X\sim N(4,3^2)$.

(1) 求 $P(-2<X\leqslant 10)$;(2) 求 $P(X>3)$;(3) 设 d 满足 $P(X>d)\geqslant 0.9$,问 d 至多为多少?

解　(1) $P(-2<X\leqslant 10)=\Phi(2)-\Phi(-2)=2\Phi(2)-1$

$$=2\times 0.977\,2-1=0.954\,4.$$

(2) $P(X>3)=1-\Phi(-1/3)=\Phi(1/3)=0.630\,4$.

(3) 由 $0.9\leqslant P(X>d)=\Phi((4-d)/3)$ 查表得 $(4-d)/3\geqslant 1.282$,由此解得 $d\leqslant 0.154$, 故 d 至多取 0.154.

22. 测量到某一目标的距离时,发生的随机误差 X(以 m 计)具有密度函数

$$p(x)=\frac{1}{40\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-20)^2}{3\,200}},\qquad -\infty<x<\infty.$$

求在三次测量中,至少有一次误差的绝对值不超过 30 m 的概率.

解　记 Y 为三次测量中误差的绝对值不超过 30 m 的次数,则 $Y\sim b(3,p)$,其中 p 为"一次测量中误差的绝对值不超过 30 m"的概率,由 $X\sim N(20,40^2)$,可知

$$p=P(-30\leqslant X\leqslant 30)=\Phi(10/40)-\Phi(-50/40)$$
$$=\Phi(0.25)-1+\Phi(1.25)=0.493\,1.$$

所以"三次测量中至少有一次误差的绝对值不超过 30 m"的概率为

$$P(Y\geqslant 1)=1-P(Y=0)=1-(1-p)^3=1-0.506\,9^3=0.869\,8.$$

23. 从甲地飞往乙地的航班,每天上午 10:10 起飞,飞行时间 X 服从均值是 4 h,标准差是 20 min 的正态分布.

(1) 该机在下午 2:30 以后到达乙地的概率是多少?

(2) 该机在下午 2:20 以前到达乙地的概率是多少?

(3) 该机在下午 1:50 至 2:30 之间到达乙地的概率是多少?

解　设时间单位为 min,则 $X\sim N(240,20^2)$.

(1) 所求概率为

$$P(X\geqslant 260)=1-\Phi((260-240)/20)=1-\Phi(1)=1-0.841\,3=0.158\,7.$$

(2) 所求概率为

$$P(X\leqslant 250)=\Phi((250-240)/20)=\Phi(0.5)=0.691\,5.$$

(3) 所求概率为

$$P(220\leqslant X\leqslant 260)=2\Phi(1)-1=2\times 0.841\,3-1=0.682\,6.$$

24. 在某场招聘人员的考试中,共有 10 000 人报考.假设考试成绩服从正态分布,且已知 90 分以上有 359 人,60 分以下有 1 151 人.现按考试成绩从高分到低分依次录用 2 500 人,试问被录用者中最低分为多少?

解　记 X 为考试成绩,则 $X\sim N(\mu,\sigma^2)$.由频率估计概率知

$$0.035\,9=P(X>90)=1-\Phi\left(\frac{90-\mu}{\sigma}\right),$$

$$0.1151 = P(X < 60) = \Phi\left(\frac{60-\mu}{\sigma}\right),$$

上面两式可改写为

$$0.9641 = \Phi\left(\frac{90-\mu}{\sigma}\right), \qquad 0.8849 = \Phi\left(\frac{\mu-60}{\sigma}\right),$$

再查表得

$$\frac{90-\mu}{\sigma} = 1.8, \qquad \frac{\mu-60}{\sigma} = 1.2,$$

由此解得 $\mu = 72, \sigma = 10$.设被录用者中最低分为 k,则由

$$0.25 = P(X \geqslant k) = 1 - \Phi\left(\frac{k-72}{10}\right), \quad 或 \quad 0.75 = \Phi\left(\frac{k-72}{10}\right),$$

查表得 $(k-72)/10 \geqslant 0.675$,从中解得 $k \geqslant 78.75$,因此取被录用者中最低分为 78.75 分即可.

注:当 $p < 0.5$ 时,满足等式 $\Phi(x)=p$ 的 x 在标准正态分布函数表上不易查得,故改写此式为 $\Phi(-x) = 1 - p > 0.5$,即可查得 $-x$.

25. 设随机变量 X 服从正态分布 $N(60,3^2)$,试求实数 a,b,c,d 使得 X 落在如下五个区间中的概率之比为 7∶24∶38∶24∶7.

$$(-\infty,a], \quad (a,b], \quad (b,c], \quad (c,d], \quad (d,\infty).$$

解　由题设条件知

$$P(X \leqslant a) = 0.07,\ P(X \leqslant b) = 0.31,\ P(X \leqslant c) = 0.69,\ P(X \leqslant d) = 0.93.$$

所以

(1) 由于 $P(X \leqslant a) = \Phi((a-60)/3) = 0.07$,即 $\Phi((60-a)/3) = 0.93$,因此查表得 $(60-a)/3 = 1.48$,由此得 $a = 55.56$.

(2) 由于 $P(X \leqslant b) = \Phi((b-60)/3) = 0.31$,即 $\Phi((60-b)/3) = 0.69$,因此查表得 $(60-b)/3 = 0.495$,由此得 $b = 58.5$.

(3) 由 $P(X \leqslant c) = 0.69$,查表得 $(c-60)/3 = 0.495$,由此得 $c = 61.5$.

(4) 由 $P(X \leqslant d) = 0.93$,查表得 $(d-60)/3 = 1.48$,由此得 $d = 64.44$.

26. 设随机变量 X 与 Y 均服从正态分布,X 服从正态分布 $N(\mu,4^2)$,Y 服从正态分布 $N(\mu,5^2)$,试比较以下 p_1 和 p_2 的大小:

$$p_1 = P(X \leqslant \mu - 4), \qquad p_2 = P(Y \geqslant \mu + 5).$$

解　因为

$$p_1 = P(X \leqslant \mu - 4) = \Phi(-1) = 1 - \Phi(1),$$
$$p_2 = P(Y \geqslant \mu + 5) = 1 - \Phi(1),$$

所以 p_1 与 p_2 一样大.

27. 设随机变量 X 服从正态分布 $N(0,\sigma^2)$,若 $P(|X| > k) = 0.1$,试求 $P(X < k)$.

解　由题设条件知

$$0.9 = P(-k \leqslant X \leqslant k) = \Phi(k/\sigma) - \Phi(-k/\sigma) = 2\Phi(k/\sigma) - 1,$$

由此得 $\Phi(k/\sigma) = 0.95$.所以 $P(X < k) = \Phi(k/\sigma) = 0.95$.

28. 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，试问：随着 σ 的增大，概率 $P(|X-\mu|<\sigma)$ 是如何变化的？

解　因为

$$P(|X-\mu|<\sigma)=P(-\sigma<X-\mu<\sigma)=\Phi(1)-\Phi(-1)=0.682\,6,$$

所以随着 σ 的增大,概率 $P(|X-\mu|<\sigma)$ 是不变的.

29. 设随机变量 X 服从参数为 $\mu=160$ 和 σ 的正态分布,若要求 $P(120<X\leqslant 200)\geqslant 0.90$,允许 σ 最大为多少?

解　由题设条件 $0.90\leqslant P(120<X\leqslant 200)=2\Phi(40/\sigma)-1$,得 $\Phi(40/\sigma)\geqslant 0.95$,从而查表得 $40/\sigma\geqslant 1.645$,或 $\sigma\leqslant 24.32$,这表明 σ 最大为 24.32.

30. 设随机变量 $X\sim N(\mu,\sigma^2)$,求 $E(|X-\mu|)$.

解　利用变换 $t=(x-\mu)/\sigma$ 及偶函数性质可得

$$E(|X-\mu|)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}|x-\mu|\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\mathrm{d}x$$

$$=\sigma\sqrt{\frac{2}{\pi}}\int_0^{\infty}\exp\left\{-\frac{t^2}{2}\right\}\mathrm{d}\left(\frac{t^2}{2}\right)=\sigma\sqrt{\frac{2}{\pi}}.$$

31. 设随机变量 $X\sim N(0,\sigma^2)$,证明 $E(|X|)=\sigma\sqrt{\dfrac{2}{\pi}}$.

证　在上题中令 $\mu=0$ 即可得结论.

32. 设随机变量 X 服从伽马分布 $Ga(2,0.5)$,试求 $P(X<4)$.

解　伽马分布 $Ga(2,0.5)$ 的密度函数为

$$p(x)=\begin{cases}\dfrac{0.5^2}{\Gamma(2)}x^{2-1}\mathrm{e}^{-0.5x}, & x\geqslant 0,\\ 0, & x<0.\end{cases}$$

由于 $\Gamma(2)=1$,因此所求概率为

$$P(X<4)=\frac{1}{4}\int_0^4 x\mathrm{e}^{-x/2}\mathrm{d}x=0.594\,0.$$

33. 某地区漏缴税款的比例 X 服从参数 $a=2,b=9$ 的贝塔分布,试求此比例小于 10% 的概率及平均漏缴税款的比例.

解　贝塔分布 $Be(2,9)$ 的密度函数为

$$p(x)=\begin{cases}\dfrac{\Gamma(2+9)}{\Gamma(2)\Gamma(9)}x^{2-1}(1-x)^{9-1}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

因为 $\Gamma(2+9)=10!,\Gamma(2)=1,\Gamma(9)=8!$,所以 $\Gamma(2+9)/\Gamma(9)=90$,因此

$$P(X<0.1)=90\int_0^{0.1}x(1-x)^8\mathrm{d}x=0.263\,9.$$

$$E(X)=\frac{a}{a+b}=\frac{2}{11}=0.181\,8.$$

34. 某班级学生中数学成绩不及格的比率 X 服从 $a=1,b=4$ 的贝塔分布,试求 $P(X>E(X))$.

解 贝塔分布 $Be(1,4)$ 的密度函数为

$$p(x)=\begin{cases}4(1-x)^3, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

且由 $E(X)=a/(a+b)=1/5=0.2$，知

$$P(X>E(X))=\int_{0.2}^{1}4(1-x)^3\mathrm{d}x=0.409\ 6.$$

§2.6 随机变量函数的分布

内容概要

1. 设连续随机变量 X 的密度函数为 $p_X(x)$，$Y=g(X)$.

(1) 若 $y=g(x)$ 严格单调，其反函数 $h(y)$ 有连续导函数，则 $Y=g(X)$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X[h(y)]\,|h'(y)|, & a<y<b,\\ 0, & \text{其他},\end{cases}$$

其中 $a=\min\{g(-\infty),g(\infty)\}$，$b=\max\{g(-\infty),g(\infty)\}$.

(2) 若 $y=g(x)$ 在不相重叠的区间 $I_1,I_2,\cdots$ 上逐段严格单调，其反函数 $h_1(y),h_2(y),\cdots$ 有连续导函数，则 $Y=g(X)$ 的密度函数为

$$p_Y(y)=\sum_i p_X(h_i(y))\,|h_i'(y)|.$$

2. 正态变量的线性变换仍为正态变量 若随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则当 $a\neq 0$ 时，有 $Y=aX+b\sim N(a\mu+b,a^2\sigma^2)$.

3. 对数正态分布

(1) 若随机变量 X 的密度函数为

$$p(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\,x\sigma}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

则称 X 服从对数正态分布，记为 $X\sim LN(\mu,\sigma^2)$，其中 $-\infty<\mu<\infty$，$\sigma>0$. 对数正态分布的密度函数如图 2.17.

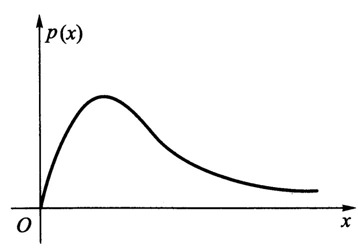


图 2.17

(2) 若 $X\sim LN(\mu,\sigma^2)$，则 $E(X)=\mathrm{e}^{\mu+\sigma^2/2}$，$\mathrm{Var}(X)=(\mathrm{e}^{\sigma^2}-1)\mathrm{e}^{2\mu+\sigma^2}$.

(3) 若 $X\sim LN(\mu,\sigma^2)$，则 $Y=\ln X\sim N(\mu,\sigma^2)$.

4. 若随机变量 $X\sim Ga(\alpha,\lambda)$，则当 $k>0$ 时，有 $Y=kX\sim Ga(\alpha,\lambda/k)$.

5. 若随机变量 X 的分布函数 $F_X(x)$ 为严格单调增的连续函数，其反函数 $F_X^{-1}(y)$ 存

在,则 $Y=F_X(X)$ 服从(0,1)上的均匀分布 $U(0,1)$.

习题与解答 2.6

1. 已知离散随机变量 X 的分布列为

X	-2	-1	0	1	3
P	1/5	1/6	1/5	1/15	11/30

试求 $Y=X^2$ 与 $Z=|X|$ 的分布列.

解

Y	4	1	0	1	9
P	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{11}{30}$

即

Y	0	1	4	9
P	$\frac{1}{5}$	$\frac{7}{30}$	$\frac{1}{5}$	$\frac{11}{30}$

Z	2	1	0	1	3
P	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{11}{30}$

即

Z	0	1	2	3
P	$\frac{1}{5}$	$\frac{7}{30}$	$\frac{1}{5}$	$\frac{11}{30}$

2. 已知随机变量 X 的密度函数为

$$p(x)=\frac{2}{\pi}\cdot\frac{1}{e^x+e^{-x}},\qquad -\infty<x<\infty.$$

试求随机变量 $Y=g(X)$ 的概率分布,其中

$$g(x)=\begin{cases}-1, & 当\ x<0,\\ 1, & 当\ x\geqslant 0.\end{cases}$$

解　因为 $p(x)$ 为偶函数,所以可得 $P(X<0)=P(X\geqslant 0)=0.5$.由此得

$$P(Y=-1)=P(X<0)=P(X\geqslant 0)=P(Y=1)=0.5.$$

所以 Y 的分布列为

Y	-1	1
P	0.5	0.5

3. 设随机变量 X 服从 $(-1,2)$ 上的均匀分布,记

$$Y=\begin{cases}1, & X\geqslant 0,\\ -1, & X<0.\end{cases}$$

试求 Y 的分布列.

解　因为 $P(Y=-1)=P(X<0)=1/3$,$P(Y=1)=P(X\geqslant 0)=2/3$.所以 Y 的分布列为

Y	-1	1
P	1/3	2/3

4. 设随机变量 $X \sim U(0,1)$,试求 $1-X$ 的分布.

解 X 的密度函数为

$$p_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

因为 $y=g(x)=1-x$ 在 $(0,1)$ 上为严格单调减函数,其反函数为 $x=h(y)=1-y$,且有 $h'(y)=-1$,所以 $Y=1-X$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X(1-y)\,|-1|, & 0<y<1,\\ 0, & \text{其他}\end{cases}=\begin{cases}1, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

这表明:当 $X \sim U(0,1)$ 时,$1-X$ 与 X 同分布.

5. 设随机变量 X 服从 $(-\pi/2,\pi/2)$ 上的均匀分布,求随机变量 $Y=\cos X$ 的密度函数 $p_Y(y)$.

解 X 的密度函数为

$$p_X(x)=\begin{cases}1/\pi, & -\pi/2<x<\pi/2,\\ 0, & \text{其他}.\end{cases}$$

由于 X 在 $(-\pi/2,\pi/2)$ 内取值,所以 $Y=\cos X$ 的可能取值区间为 $(0,1)$.在 Y 的可能取值区间外,$p_Y(y)=0$.

当 $0<y<1$ 时,使 $\{Y\leqslant y\}$ 的 x 取值范围为两个互不相交的区间 Δ_1 和 Δ_2,其中 $\Delta_1=(-\pi/2,-\arccos y)$,$\Delta_2=(\arccos y,\pi/2)$.如图 2.18.

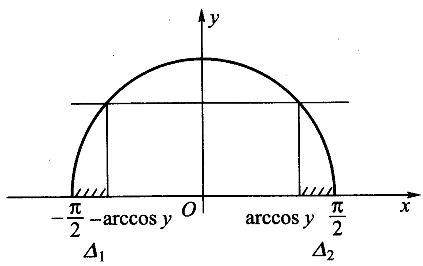


图 2.18

故

$$F_Y(y)=P(Y\leqslant y)=\int_{-\pi/2}^{-\arccos y}\frac{1}{\pi}\mathrm{d}x+\int_{\arccos y}^{\pi/2}\frac{1}{\pi}\mathrm{d}x.$$

在上式两端对 y 求导,得

$$p_Y(y)=\frac{1}{\pi\sqrt{1-y^2}}+\frac{1}{\pi\sqrt{1-y^2}}=\frac{2}{\pi\sqrt{1-y^2}},\qquad 0<y<1.$$

即

$$p_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

6. 设圆的直径服从区间 $(0,1)$ 上的均匀分布,求圆的面积的密度函数.

解 设圆的直径为 X,则圆的面积 $Y=\pi X^2/4$,而 X 的密度函数为

$$p_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

因为 $y=g(x)=\pi x^2/4$ 在区间 $(0,1)$ 上为严格单调增函数,其反函数为 $x=h(y)=\sqrt{4y/\pi}$,且 $h'(y)=1/\sqrt{\pi y}$,所以圆面积 $Y=\pi X^2/4$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X\left(\sqrt{\dfrac{4y}{\pi}}\right)\left|\dfrac{1}{\sqrt{\pi y}}\right|, & 0<y<\dfrac{\pi}{4}\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{\sqrt{\pi y}}, & 0<y<\dfrac{\pi}{4},\\ 0, & \text{其他}.\end{cases}$$

7. 设随机变量 X 服从区间(1,2)上的均匀分布,试求 $Y=e^{2X}$ 的密度函数.

解　X 的密度函数为

$$p_X(x)=\begin{cases}1, & 1<x<2,\\ 0, & \text{其他}.\end{cases}$$

由于 X 在(1,2)内取值,所以 $Y=e^{2X}$ 的可能取值区间为 (e^2,e^4),且 $y=g(x)=e^{2x}$ 在区间(1,2)上为严格单调增函数,其反函数为 $x=h(y)=(\ln y)/2$,且 $h'(y)=1/(2y)$,所以 $Y=e^{2X}$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X\left(\dfrac{\ln y}{2}\right)\left|\dfrac{1}{2y}\right|, & e^2<y<e^4,\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{2y}, & e^2<y<e^4,\\ 0, & \text{其他}.\end{cases}$$

8. 设随机变量 X 服从区间(0,2)上的均匀分布.

(1) 求 $Y=X^2$ 的密度函数;(2) 求 $P(Y<2)$.

解　X 的密度函数为

$$p_X(x)=\begin{cases}1/2, & 0<x<2,\\ 0, & \text{其他}.\end{cases}$$

(1) $Y=X^2$ 的可能取值区间为(0,4).因为 $y=g(x)=x^2$ 在区间(0,2)上为严格单调增函数,其反函数为 $x=h(y)=\sqrt{y}$,且 $h'(y)=1/(2\sqrt{y})$,所以 $Y=X^2$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X(\sqrt{y})\left|\dfrac{1}{2\sqrt{y}}\right|, & 0<y<4\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{4\sqrt{y}}, & 0<y<4,\\ 0, & \text{其他}.\end{cases}$$

(2)

$$P(Y<2)=\int_0^2\frac{1}{4\sqrt{y}}\mathrm{d}y=\frac{1}{2}\sqrt{y}\,\Big|_0^2=\frac{\sqrt{2}}{2}.$$

9. 设随机变量 X 服从区间(-1,1)上的均匀分布,求:

(1) $P\left(|X|>\dfrac{1}{2}\right)$;

(2) $Y=|X|$ 的密度函数.

解　(1) $$P\left(|X|>\frac{1}{2}\right)=P\left(X>\frac{1}{2}\right)+P\left(X<-\frac{1}{2}\right)=\int_{\frac{1}{2}}^{1}\frac{1}{2}\mathrm{d}x+\int_{-1}^{-\frac{1}{2}}\frac{1}{2}\mathrm{d}x=\frac{1}{4}+\frac{1}{4}=0.5.$$

(2) $F_Y(y)=P(|X|\leqslant y)$.当 $y<0$ 时,$F_Y(y)=0$;当 $y\geqslant 1$ 时,$F_Y(y)=1$;当 $0\leqslant y<1$ 时,

$$F_Y(y)=P(-y\leqslant X\leqslant y)=\int_{-y}^{y}\frac{1}{2}\mathrm{d}x=y.$$

所以得

$$p_Y(y)=\begin{cases}1, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

10. 设随机变量 X 服从(0,1)上的均匀分布,试求以下 Y 的密度函数:

(1) $Y=-2\ln X$;　　(2) $Y=3X+1$;

(3) $Y=e^X$;　　(4) $Y=|\ln X|$.

解　X 的密度函数为

$$p_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

(1) 因为 Y 的可能取值区间为 $(0,\infty)$,且 $y=g(x)=-2\ln x$ 在区间 $(0,1)$ 上为严格单调减函数,其反函数为 $x=h(y)=e^{-0.5y}$,且 $h'(y)=-0.5e^{-0.5y}$.所以 $Y=-2\ln X$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X(e^{-0.5y})\,|-0.5e^{-0.5y}|, & y>0,\\ 0, & y\leqslant 0\end{cases}=\begin{cases}0.5e^{-0.5y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(2) 因为 Y 的可能取值区间为 $(1,4)$,且 $y=g(x)=3x+1$ 在区间 $(0,1)$ 上为严格单调增函数,其反函数为 $x=h(y)=(y-1)/3$,且 $h'(y)=1/3$.所以 $Y=3X+1$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X((y-1)/3)\,|1/3|, & 1<y<4\\ 0, & \text{其他}\end{cases}=\begin{cases}1/3, & 1<y<4,\\ 0, & \text{其他}.\end{cases}$$

(3) 因为 Y 的可能取值区间为 $(1,e)$,且 $y=g(x)=e^x$ 在区间 $(0,1)$ 上为严格单调增函数,其反函数为 $x=h(y)=\ln y$,且 $h'(y)=1/y$.所以 $Y=e^X$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X(\ln y)\,|1/y|, & 1<y<e,\\ 0, & \text{其他}\end{cases}=\begin{cases}1/y, & 1<y<e,\\ 0, & \text{其他}.\end{cases}$$

(4) 因为 Y 的可能取值区间为 $(0,\infty)$,且 $y=|\ln x|$ 在区间 $(0,1)$ 上为严格单调减函数,其反函数为 $x=h(y)=e^{-y}$,且 $h'(y)=-e^{-y}$.所以 $Y=|\ln X|$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X(e^{-y})\,|-e^{-y}|, & y>0\\ 0, & y\leqslant 0\end{cases}=\begin{cases}e^{-y}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

11. 设随机变量 X 的密度函数为

$$p_X(x)=\begin{cases}\dfrac{3}{2}x^2, & -1<x<1,\\ 0, & \text{其他}.\end{cases}$$

试求下列随机变量的分布:

(1) $Y_1=3X$;　(2) $Y_2=3-X$;　(3) $Y_3=X^2$.

解　(1) 因为 Y_1 的可能取值区间为 $(-3,3)$,且 $y=g(x)=3x$ 在区间 $(-1,1)$ 上为严格单调增函数,其反函数为 $x=h(y)=y/3$,且 $h'(y)=1/3$,所以 $Y_1=3X$ 的密度函数为

$$p_{Y_1}(y)=\begin{cases}p_X(y/3)\,|1/3|, & -3<y<3,\\ 0, & \text{其他}\end{cases}=\begin{cases}y^2/18, & -3<y<3,\\ 0, & \text{其他}.\end{cases}$$

(2) 因为 Y_2 的可能取值区间为 $(2,4)$,且 $y=g(x)=3-x$ 在区间 $(-1,1)$ 上为严格单调减函数,其反函数为 $x=h(y)=3-y$,且 $h'(y)=-1$,所以 $Y_2=3-X$ 的密度函数为

$$p_{Y_2}(y)=\begin{cases}p_X(3-y)\,|-1|, & 2<y<4,\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{3}{2}(3-y)^2, & 2<y<4,\\ 0, & \text{其他}.\end{cases}$$

(3) 因为 Y_3 的可能取值区间为 $(0,1)$，所以在区间 $(0,1)$ 外，Y_3 的密度函数为 $p_{Y_3}(y)=0$. 而当 $0<y<1$ 时，Y_3 的分布函数为

$$F_{Y_3}(y)=P(Y_3\leqslant y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant\sqrt{y})=F_X(\sqrt{y})-F_X(-\sqrt{y}),$$

上式两边关于 y 求导，得

$$p_{Y_3}(y)=p_X(\sqrt{y})\frac{1}{2\sqrt{y}}+p_X(-\sqrt{y})\frac{1}{2\sqrt{y}}=\frac{3}{2}\sqrt{y},$$

即

$$p_{Y_3}(y)=\begin{cases}\dfrac{3}{2}\sqrt{y}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

这是贝塔分布 $Be(3/2,1)$.

12. 设 $X\sim N(0,\sigma^2)$，求 $Y=X^2$ 的分布.

解　因为 $Y=X^2$ 的可能取值区间为 $(0,\infty)$，所以当 $y\leqslant 0$ 时，Y 的密度函数为 $p_Y(y)=0$. 而当 $y>0$ 时，Y 的分布函数为

$$F_Y(y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant\sqrt{y})=F_X(\sqrt{y})-F_X(-\sqrt{y}),$$

对上式两边关于 y 求导，得

$$p_Y(y)=p_X(\sqrt{y})\frac{1}{2\sqrt{y}}+p_X(-\sqrt{y})\frac{1}{2\sqrt{y}}=\frac{1}{\sqrt{y}}\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{y}{2\sigma^2}\right\}.$$

即

$$p_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi y}\,\sigma}\exp\left\{-\dfrac{y}{2\sigma^2}\right\}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

这是伽马分布 $Ga\left(\dfrac{1}{2},\dfrac{1}{2\sigma^2}\right)$.

13. 设 $X\sim N(\mu,\sigma^2)$，求 $Y=e^X$ 的数学期望与方差.

解　因为 $Y=e^X$ 的可能取值范围为 $(0,\infty)$，且 $y=g(x)=e^x$ 为严格单调增函数，其反函数为 $x=h(y)=\ln y$，及 $h'(y)=1/y$，所以 Y 的密度函数为

$$p_Y(y)=\begin{cases}p_X(\ln y)\left|\dfrac{1}{y}\right|, & y>0,\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{\sqrt{2\pi}\,y\sigma}\exp\left\{-\dfrac{(\ln y-\mu)^2}{2\sigma^2}\right\}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

这是对数正态分布 $LN(\mu,\sigma^2)$. 为求其数学期望，采用线性变换 $t=(x-\mu)/\sigma$ 可得

$$E(Y)=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^x e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{\sigma t+\mu}e^{-t^2/2}\sigma\,dt$$

$$=e^{\mu}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-(t^2/2-\sigma t)}dt=e^{\mu+\sigma^2/2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-(t-\sigma)^2/2}dt=e^{\mu+\sigma^2/2}.$$

上式最后一个等式成立是因为积分中的被积函数是 $N(\sigma,1)$ 的密度函数之故.

为求 Y 的方差，先求 $E(Y^2)$. 施行相同的线性变换，可得

$$E(Y^2)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\mathrm{e}^{2x}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\mathrm{e}^{2\sigma t+2\mu}\mathrm{e}^{-t^2/2}\sigma\,\mathrm{d}t$$

$$=\mathrm{e}^{2\mu}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-(t^2/2-2\sigma t)}\mathrm{d}t=\mathrm{e}^{2\mu+2\sigma^2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-(t-2\sigma)^2/2}\mathrm{d}t=\mathrm{e}^{2\mu+2\sigma^2}.$$

上式最后一个等式成立是因为积分中的被积函数是 $N(2\sigma,1)$ 的密度函数之故.由此得

$$\mathrm{Var}(Y)=E(Y^2)-[E(Y)]^2=\mathrm{e}^{2\mu+2\sigma^2}-[\mathrm{e}^{\mu+\sigma^2/2}]^2=\mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2}-1).$$

14. 设随机变量 X 服从标准正态分布 $N(0,1)$,试求以下 Y 的密度函数:

(1) $Y=|X|$; (2) $Y=2X^2+1$.

解 (1) $Y=|X|$的可能取值范围为$(0,\infty)$,所以当 $y\leqslant 0$ 时,Y 的密度函数为 $p_Y(y)=0$;当 $y>0$ 时,Y 的分布函数为

$$F_Y(y)=P(|X|\leqslant y)=P(-y\leqslant X\leqslant y)=F_X(y)-F_X(-y).$$

对上式两端关于 y 求导得

$$p_Y(y)=p_X(y)+p_X(-y)=\sqrt{\frac{2}{\pi}}\mathrm{e}^{-\frac{y^2}{2}},$$

所以 Y 的密度函数为

$$p_Y(y)=\begin{cases}\sqrt{\dfrac{2}{\pi}}\mathrm{e}^{-\frac{y^2}{2}}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

这个分布被称为半正态分布.

(2) $Y=2X^2+1$ 的可能取值范围为$(1,\infty)$,所以当 $y\leqslant 1$ 时,Y 的密度函数为 $p_Y(y)=0$;当 $y>1$ 时,Y 的分布函数为

$$F_Y(y)=P(2X^2+1\leqslant y)=P\left(-\sqrt{(y-1)/2}\leqslant X\leqslant\sqrt{(y-1)/2}\right)$$

$$=F_X\left(\sqrt{(y-1)/2}\right)-F_X\left(-\sqrt{(y-1)/2}\right).$$

对上式两端关于 y 求导得

$$p_Y(y)=p_X\left(\sqrt{(y-1)/2}\right)\frac{1}{4\sqrt{(y-1)/2}}+p_X\left(-\sqrt{(y-1)/2}\right)\frac{1}{4\sqrt{(y-1)/2}}$$

$$=\frac{1}{2\sqrt{(y-1)/2}}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y-1}{4}}=\frac{1}{2\sqrt{\pi(y-1)}}\mathrm{e}^{-\frac{y-1}{4}},$$

所以 Y 的密度函数为

$$p_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{\pi(y-1)}}\mathrm{e}^{-\frac{y-1}{4}}, & y>1,\\ 0, & \text{其他}.\end{cases}$$

15. 设随机变量 X 的密度函数为

$$p(x)=\begin{cases}\mathrm{e}^{-x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

试求以下 Y 的密度函数:

(1) $Y=2X+1$; (2) $Y=\mathrm{e}^X$; (3) $Y=X^2$.

解　(1) 因为 $Y=2X+1$ 的可能取值范围是 $(1,\infty)$，且 $y=g(x)=2x+1$ 是严格单调增函数，其反函数为 $x=h(y)=(y-1)/2$，及 $h'(y)=1/2$，所以 Y 的密度函数为

$$p_Y(y)=\begin{cases}p_X\left(\dfrac{y-1}{2}\right)\left|\dfrac{1}{2}\right|, & y>1,\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{2}e^{-\frac{y-1}{2}}, & y>1,\\ 0, & \text{其他}.\end{cases}$$

(2) 因为 $Y=e^X$ 的可能取值范围是 $(1,\infty)$，且 $y=g(x)=e^x$ 是严格单调增函数，其反函数为 $x=h(y)=\ln y$，及 $h'(y)=1/y$，所以 Y 的密度函数为

$$p_Y(y)=\begin{cases}p_X(\ln y)\left|\dfrac{1}{y}\right|, & y>1\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{y^2}, & y>1,\\ 0, & \text{其他}.\end{cases}$$

(3) 因为 $Y=X^2$ 的可能取值范围是 $(0,\infty)$，且 $y=g(x)=x^2$ 在 $(0,\infty)$ 上是严格单调增函数，其反函数为 $x=h(y)=\sqrt{y}$，及 $h'(y)=1/(2\sqrt{y})$，所以 Y 的密度函数为

$$p_Y(y)=\begin{cases}p_X(\sqrt{y})\left|\dfrac{1}{2\sqrt{y}}\right|, & y>0\\ 0, & \text{其他}\end{cases}=\begin{cases}\dfrac{1}{2\sqrt{y}}e^{-\sqrt{y}}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

这是韦布尔分布的特例.一般韦布尔分布(记为 $W(m,\eta)$)的密度函数为

$$p(y)=\begin{cases}\dfrac{m}{\eta}\left(\dfrac{y}{\eta}\right)^{m-1}\exp\left\{-\left(\dfrac{y}{\eta}\right)^{m}\right\}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

本题结论就是 $m=1/2,\eta=1$ 时的韦布尔分布 $W(1/2,1)$.

16. 设随机变量 X 服从参数为2的指数分布，试证：$Y_1=e^{-2X}$ 和 $Y_2=1-e^{-2X}$ 都服从区间 $(0,1)$ 上的均匀分布.

证　因为 X 的密度函数为

$$p_X(x)=\begin{cases}2e^{-2x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

又因为 Y_1 的可能取值范围是 $(0,1)$，且 $y_1=e^{-2x}$ 是严格单调减函数，其反函数为 $x=h(y_1)=-0.5\ln y_1$，$h'(y_1)=-0.5/y_1$，所以 Y_1 的密度函数为

$$p_{Y_1}(y)=\begin{cases}p_X(-0.5\ln y_1)\left|-0.5/y_1\right|, & 0<y_1<1\\ 0, & \text{其他}\end{cases}=\begin{cases}1, & 0<y_1<1,\\ 0, & \text{其他},\end{cases}$$

即 $Y_1\sim U(0,1)$.又由前面第4题知，$Y_2=1-e^{-2X}=1-Y_1$ 也服从区间 $(0,1)$ 上的均匀分布.结论得证.

17. 设随机变量 $X\sim LN(\mu,\sigma^2)$，试证：$Y=\ln X\sim N(\mu,\sigma^2)$.

证　因为 X 的密度函数为

$$p_X(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\,x\sigma}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

又因为 $Y=\ln X$ 的可能取值范围为 $(-\infty,\infty)$，且 $y=g(x)=\ln x$ 是区间 $(0,\infty)$ 上的严格单调增函数，其反函数为 $x=h(y)=e^y$，$h'(y)=e^y$，所以 Y 的密度函数为

$$p_Y(y)=p_X(\mathrm{e}^y)\left|\mathrm{e}^y\right|=\frac{1}{\sqrt{2\pi}\,\mathrm{e}^y\sigma}\exp\left\{-\frac{(\ln \mathrm{e}^y-\mu)^2}{2\sigma^2}\right\}\mathrm{e}^y$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\},\qquad -\infty<y<\infty.$$

这正是 $N(\mu,\sigma^2)$ 的密度函数.

18. 设随机变量 $Y\sim LN(5,0.12^2)$，试求 $P(Y<188.7)$.

解　$P(Y<188.7)=P(\ln Y<\ln 188.7)=\Phi\left(\dfrac{5.24-5}{0.12}\right)=\Phi(2)=0.977\ 2.$

§2.7　分布的其他特征数

内容概要

1. k 阶矩

(1) 称 $\mu_k=E(X^k)$ 为 X 的 k 阶**原点矩**.一阶原点矩就是数学期望；

(2) 称 $\nu_k=E(X-E(X))^k$ 为 X 的 k 阶**中心矩**.二阶中心矩就是方差；

(3) 前 k 阶中心矩可用原点矩表示，如

$$\nu_1=0,$$
$$\nu_2=\mu_2-\mu_1^2,$$
$$\nu_3=\mu_3-3\mu_2\mu_1+2\mu_1^3,$$
$$\nu_4=\mu_4-4\mu_3\mu_1+6\mu_2\mu_1^2-3\mu_1^4.$$

2. 变异系数　称比值 $C_v(X)=\dfrac{\sqrt{\mathrm{Var}(X)}}{E(X)}$ 为 X 的变异系数.变异系数是一个无量纲的量.

3. 分位数　设连续随机变量 X 的分布函数为 $F(x)$，密度函数为 $p(x)$.对任意 $p\in(0,1)$，

(1) 称满足条件

$$F(x_p)=\int_{-\infty}^{x_p}p(x)\,\mathrm{d}x=p$$

的 x_p 为此分布的 p **分位数**，又称**下侧 p 分位数**，它把密度函数下的面积一分为二，左侧面积恰好为 p；

(2) 称满足条件

$$1-F(x_p')=\int_{x_p'}^{\infty}p(x)\,\mathrm{d}x=p$$

的 x_p' 为此分布的**上侧 p 分位数**；

(3) 分位数与上侧分位数的转换公式： $x_p'=x_{1-p}$， $x_p=x_{1-p}'$；

(4) 称 $p=0.5$ 时的 p 分位数 $x_{0.5}$ 为此分布的**中位数**，即 $x_{0.5}$ 满足

$$F(x_{0.5})=\int_{-\infty}^{x_{0.5}}p(x)\,\mathrm{d}x=0.5;$$

（5）若随机变量 X 的密度函数 $p(x)$ 是偶函数，则此分布的 p 分位数 x_p 满足 $x_p=-x_{1-p}$. 中位数为分布对称中心；

（6）记标准正态分布的 p 分位数为 u_p. 因为标准正态密度函数是偶函数，所以 $u_p=-u_{1-p}$. 譬如 $u_{0.25}=-u_{0.75}=-0.675$；

（7）一般正态分布 $N(\mu,\sigma^2)$ 的 p 分位数 x_p 满足 $x_p=\mu+\sigma\times u_p$. 譬如 $N(10,2^2)$ 的 0.25 分位数为 $x_{0.25}=10+2u_{0.25}=8.65$；

（8）分布的矩有可能不存在，但连续分布的分位数总存在. p 分位数 x_p 总是 p 的增函数.

4. 偏度系数

（1）称比值

$$\beta_S=\frac{E(X-E(X))^3}{[\mathrm{Var}(X)]^{3/2}}$$

为 X 的分布的偏度系数，简称**偏度**；

（2）偏度系数刻画的是分布的不对称程度，$|\beta_S|$ 愈大，分布的对称性愈差；

（3）任一对称分布的偏度 $\beta_S=0$. 当 $\beta_S>0$ 时，分布为正偏（又称右偏）；当 $\beta_S<0$ 时，分布为负偏（又称左偏）.

5. 峰度系数

（1）称

$$\beta_k=\frac{E(X-E(X))^4}{[\mathrm{Var}(X)]^2}-3$$

为 X 的分布的峰度系数，简称峰度；

（2）峰度系数是刻画分布的尖峭性和尾部粗细的一个特征数；

（3）任一正态分布的峰度 $\beta_k=0$. 当 $\beta_k<0$ 时，分布比标准正态分布平坦；当 $\beta_k>0$ 时，分布比标准正态分布更尖峭.

6. 偏度与峰度都是描述分布（密度）形状的参数.

习题与解答 2.7

1. 设随机变量 $X\sim U(a,b)$，对 $k=1,2,3,4$，求 $\mu_k=E(X^k)$ 与 $\nu_k=E(X-E(X))^k$. 进一步求此分布的偏度系数和峰度系数.

解　因为

$$E(X^k)=\int_a^b\frac{x^k}{b-a}\mathrm{d}x=\frac{1}{b-a}\times\frac{b^{k+1}-a^{k+1}}{k+1},$$

所以

$$\mu_1=E(X)=\frac{a+b}{2},\qquad \mu_2=E(X^2)=\frac{1}{3}(a^2+ab+b^2),$$

$$\mu_3=E(X^3)=\frac{1}{4}(a^3+a^2b+ab^2+b^3),$$

$$\mu_4 = E(X^4) = \frac{1}{5}(a^4 + a^3 b + a^2 b^2 + ab^3 + b^4).$$

$$\nu_1 = E(X - E(X)) = 0, \qquad \nu_2 = E(X - E(X))^2 = \mathrm{Var}(X) = \frac{(b-a)^2}{12},$$

$$\nu_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = 0, \quad \nu_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4 = \frac{(b-a)^4}{80}.$$

偏度系数和峰度系数分别为

$$\beta_S = \frac{\nu_3}{\nu_2^{3/2}} = 0, \qquad \beta_k = \frac{\nu_4}{\nu_2^2} - 3 = \frac{(b-a)^4/80}{[(b-a)^2/12]^2} - 3 = -1.2.$$

注：上述 β_S, β_k 与 a, b 无关，这表明：任一均匀分布的偏度为 0，峰度为 -1.2.

2. 设随机变量 $X \sim U(0, a)$，求此分布的变异系数.

解 因为 $E(X) = a/2$，$\mathrm{Var}(X) = a^2/12$，所以此分布的变异系数为

$$C_v(X) = \frac{\sqrt{\mathrm{Var}(X)}}{E(X)} = \frac{\sqrt{a^2/12}}{a/2} = \frac{\sqrt{3}}{3} = 0.577\ 4.$$

3. 求以下分布的中位数：

(1) 区间 (a, b) 上的均匀分布；

(2) 正态分布 $N(\mu, \sigma^2)$；

(3) 对数正态分布 $LN(\mu, \sigma^2)$.

解 (1) 从 $0.5 = \int_a^{x_{0.5}} \frac{1}{b-a} \mathrm{d}x$ 中解得 $x_{0.5} = \frac{a+b}{2}$.

(2) 记 $X \sim N(\mu, \sigma^2)$，由 $P(X \leqslant \mu) = \Phi\left(\frac{\mu - \mu}{\sigma}\right) = 0.5$ 可得 $x_{0.5} = \mu$.

(3) 记 $Y \sim LN(\mu, \sigma^2)$，令 $X = \ln Y$，则 $X \sim N(\mu, \sigma^2)$. 又记 $x_{0.5}$ 为 X 的中位数，$y_{0.5}$ 为 Y 的中位数，则由 (2) 知 $x_{0.5} = \mu$，即

$$0.5 = P(X \leqslant \mu) = P(\ln Y \leqslant \mu) = P(Y \leqslant \mathrm{e}^{\mu}),$$

由此得 $y_{0.5} = \mathrm{e}^{\mu}$.

4. 设随机变量 $X \sim Ga(\alpha, \lambda)$，对 $k = 1, 2, 3$，求 $\mu_k = E(X^k)$ 与 $\nu_k = E(X - E(X))^k$.

解 因为

$$E(X^k) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{k+\alpha-1} \mathrm{e}^{-\lambda x} \mathrm{d}x = \frac{\Gamma(k+\alpha)}{\Gamma(\alpha)\lambda^k},$$

所以

$$\mu_1 = E(X) = \frac{\alpha}{\lambda}, \quad \mu_2 = E(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2},$$

$$\mu_3 = E(X^3) = \frac{\alpha(\alpha+1)(\alpha+2)}{\lambda^3},$$

$$\nu_1 = E(X - E(X)) = 0, \quad \nu_2 = \mathrm{Var}(X) = \frac{\alpha}{\lambda^2},$$

$$\nu_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \frac{2\alpha}{\lambda^3}.$$

5. 设随机变量 $X \sim Exp(\lambda)$，对 $k=1,2,3,4$，求 $\mu_k = E(X^k)$ 与 $\nu_k = E(X-E(X))^k$. 进一步求此分布的变异系数、偏度系数和峰度系数.

解　因为

$$E(X^k) = \lambda\int_0^{\infty} x^k e^{-\lambda x}dx = \frac{k!}{\lambda^k},$$

所以

$$\mu_1 = E(X) = \frac{1}{\lambda}, \qquad \mu_2 = E(X^2) = \frac{2}{\lambda^2},$$

$$\mu_3 = E(X^3) = \frac{6}{\lambda^3}, \qquad \mu_4 = E(X^4) = \frac{24}{\lambda^4},$$

$$\nu_1 = E(X-E(X)) = 0, \qquad \nu_2 = \mathrm{Var}(X) = \frac{1}{\lambda^2},$$

$$\nu_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \frac{2}{\lambda^3}, \qquad \nu_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4 = \frac{9}{\lambda^4}.$$

此分布的变异系数、偏度系数和峰度系数分别为

$$C_v(X) = \frac{\sqrt{\mathrm{Var}(X)}}{E(X)} = \frac{\sqrt{1/\lambda^2}}{1/\lambda} = 1,$$

$$\beta_S = \frac{\nu_3}{\nu_2^{3/2}} = \frac{2/\lambda^3}{[1/\lambda^2]^{3/2}} = 2, \qquad \beta_k = \frac{\nu_4}{\nu_2^2} - 3 = \frac{9/\lambda^4}{1/\lambda^4} - 3 = 6.$$

由此可见：指数分布的变异系数、偏度系数与峰度系数均与参数 λ 无关. 它永远是正偏尖峰.

6. 设随机变量 X 服从正态分布 $N(10,9)$，试求 $x_{0.1}$ 和 $x_{0.9}$.

解　一般正态分布 $N(\mu,\sigma^2)$ 的 p 分位数 x_p 与标准正态分布的 p 分位数 u_p 间满足关系式：$x_p = \mu + \sigma \times u_p$，所以

$$x_{0.1} = 10 + 3u_{0.1} = 10 + 3 \times (-1.282) = 6.154,$$

$$x_{0.9} = 10 + 3u_{0.9} = 10 + 3 \times 1.282 = 13.846.$$

7. 设随机变量 X 服从双参数韦布尔分布，其分布函数为

$$F(x) = 1 - \exp\left\{-\left(\frac{x}{\eta}\right)^m\right\}, \quad x > 0,$$

其中 $\eta > 0, m > 0$. 试写出该分布的 p 分位数 x_p 的表达式，且求出当 $m=1.5, \eta = 1\ 000$ 时的 $x_{0.1}, x_{0.5}, x_{0.8}$ 的值.

解　因为 p 分位数 x_p 满足

$$1 - \exp\left\{-\left(\frac{x_p}{\eta}\right)^m\right\} = p,$$

解之得

$$x_p = \eta[-\ln(1-p)]^{1/m}.$$

将 $m=1.5, \eta = 1\ 000$ 代入上式，可得

$$x_{0.1} = 1\,000(-\ln 0.9)^{1/1.5} = 223.08,$$
$$x_{0.5} = 1\,000(-\ln 0.5)^{1/1.5} = 783.22,$$
$$x_{0.8} = 1\,000(-\ln 0.2)^{1/1.5} = 1\,373.36.$$

8. 自由度为 2 的χ^2 分布的密度函数为
$$p(x) = \frac{1}{2}\mathrm{e}^{-\frac{x}{2}}, \qquad x > 0.$$
试求出其分布函数及分位数 $x_{0.1}, x_{0.5}, x_{0.8}$.

解 此分布的分布函数 $F(x)$ 为

当 $x \leqslant 0$ 时,$F(x)=0$;当 $x > 0$ 时,
$$F(x) = P(X \leqslant x) = \int_0^x \frac{1}{2}\mathrm{e}^{-\frac{t}{2}}\mathrm{d}t = 1 - \mathrm{e}^{-\frac{x}{2}}.$$
所以此分布的 p 分位数 x_p 满足 $p = F(x_p) = 1 - \mathrm{e}^{-\frac{x_p}{2}}$,从中解得 $x_p = -2\ln(1-p)$.由此得 $x_{0.1} = -2\ln 0.9 = 0.211$;$x_{0.5} = -2\ln 0.5 = 1.386$;$x_{0.8} = -2\ln 0.2 = 3.219$.

9. 设随机变量 X 的密度函数 $p(x)$ 关于直线 $x=c$ 是对称的,且 $E(X)$ 存在,试证:

(1) 这个对称中心 c 既是均值又是中位数,即 $E(X) = x_{0.5} = c$;

(2) 如果 $c=0$,则 $x_p = -x_{1-p}$.

证 (1) 由 $p(x)$ 关于直线 $x=c$ 对称可知 $p(c+x)=p(c-x)$, $-\infty < x < \infty$,因此
$$\begin{aligned}E(X-c) &= \int_{-\infty}^{\infty}(x-c)p(x)\mathrm{d}x = \int_{-\infty}^{\infty} tp(t+c)\mathrm{d}t\\ &= \int_{-\infty}^{\infty} tp(c-t)\mathrm{d}t = \int_{-\infty}^{\infty}(c-y)p(y)\mathrm{d}y\\ &= E(c-X),\end{aligned}$$
所以得 $E(X)=c$.又由
$$\begin{aligned}0.5 &= \int_{-\infty}^{x_{0.5}} p(x)\mathrm{d}x = \int_{-\infty}^{x_{0.5}-c} p(c+y)\mathrm{d}y\\ &= \int_{-\infty}^{x_{0.5}-c} p(c-y)\mathrm{d}y = \int_{2c-x_{0.5}}^{\infty} p(t)\mathrm{d}t = \int_{x_{0.5}}^{\infty} p(x)\mathrm{d}x,\end{aligned}$$
所以 $2c - x_{0.5} = x_{0.5}$,由此得 $x_{0.5} = c$.

(2) 当 $c=0$ 时,
$$p = \int_{-\infty}^{x_p} p(x)\mathrm{d}x = \int_{-x_p}^{\infty} p(-y)\mathrm{d}y = \int_{-x_p}^{\infty} p(y)\mathrm{d}y = 1 - F(-x_p),$$
又由
$$F(-x_p) = 1-p, \quad 即 \quad -x_p = x_{1-p},$$
由此得结论.

10. 试证随机变量 X 的偏度系数与峰度系数对位移和改变比例尺是不变的,即对任意的实数 $a,b(b\neq 0)$,$Y=a+bX$ 与 X 有相同的偏度系数与峰度系数.

解 因为 $E(Y) = E[a+bX] = a+bE(X)$,所以
$$\frac{E[Y-E(Y)]^3}{\{E[Y-E(Y)]^2\}^{3/2}} = \frac{E[a+bX-a-bE(X)]^3}{\{E[a+bX-a-bE(X)]^2\}^{3/2}} = \frac{E[X-E(X)]^3}{\{E[X-E(X)]^2\}^{3/2}},$$

即 Y 与 X 有相同的偏度系数.又因为

$$\frac{E[Y-E(Y)]^4}{\{E[Y-E(Y)]^2\}^2}=\frac{E[a+bX-a-bE(X)]^4}{\{E[a+bX-a-bE(X)]^2\}^2}=\frac{E[X-E(X)]^4}{\{E[X-E(X)]^2\}^2},$$

所以 Y 与 X 有相同的峰度系数.

11. 设某项维修时间 T(单位:分) 服从对数正态分布 $LN(\mu,\sigma^2)$.

(1) 求 p 分位数 t_p;

(2) 若 $\mu=4.127\,1$,求该分布的中位数;

(3) 若 $\mu=4.127\,1$,$\sigma=1.036\,4$,求完成 95% 维修任务的时间.

解　因为 $T\sim LN(\mu,\sigma^2)$,所以 $X=\ln T\sim N(\mu,\sigma^2)$.记 x_p 为 $N(\mu,\sigma^2)$ 的 p 分位数,u_p 为 $N(0,1)$ 的 p 分位数,则由

$$p=P(X\leqslant x_p)=\Phi\left(\frac{x_p-\mu}{\sigma}\right)=\Phi(u_p)$$

知

$$x_p=\mu+\sigma u_p.$$

(1) 因为 $p=P(X\leqslant x_p)=P(\ln T\leqslant x_p)=P(T\leqslant \mathrm{e}^{x_p})$,

所以

$$t_p=\mathrm{e}^{x_p}=\exp\{\mu+\sigma u_p\}.$$

(2) 由本节习题 3(3) 知 $t_{0.5}=\mathrm{e}^{4.127\,1}=62$.

(3) 因为 $u_{0.95}=1.645$,所以当 $\mu=4.127\,1$,$\sigma=1.036\,4$ 时,完成 95% 的维修任务的时间 $t_{0.95}$ 为

$$t_{0.95}=\exp\{4.127\,1+1.036\,4\times1.645\}=341.$$

12. 某种绝缘材料的使用寿命 T(单位:小时) 服从对数正态分布 $LN(\mu,\sigma^2)$.若已知分位数 $t_{0.2}=5\,000$ 小时,$t_{0.8}=65\,000$ 小时,求 μ 和 σ.

解　由上一题知对数正态分布 $LN(\mu,\sigma^2)$ 的 p 分位数为

$$t_p=\exp\{\mu+\sigma u_p\},$$

其中 u_p 为标准正态分布 $N(0,1)$ 的 p 分位数,所以根据题意有

$$5\,000=t_{0.2}=\exp\{\mu+\sigma u_{0.2}\},$$
$$65\,000=t_{0.8}=\exp\{\mu+\sigma u_{0.8}\}.$$

将 $u_{0.2}=-0.845$,$u_{0.8}=0.845$ 代入上面两式,可解得

$$\mu=9.799\,7,\quad \sigma=1.517\,8.$$

13. 某厂决定按过去生产状况对月生产额最高的 5% 的工人发放高产奖.已知过去每人每月生产额 X(单位:千克) 服从正态分布 $N(4\,000,60^2)$,试问高产奖发放标准应把生产额定为多少?

解　根据题意知,求满足 $P(X>k)=0.05$ 的 k,即 $k=x_{0.95}$,其中 $x_{0.95}$ 为分布 $N(4\,000,60^2)$ 的 95% 分位数.又记 u_p 为标准正态分布 $N(0,1)$ 的 p 分位数,则由 $x_p=\mu+\sigma u_p$,及 $u_{0.95}=1.645$ 可得

$$x_{0.95}=4\,000+60\times1.645=4\,098.7$$

因此可将高产奖发放标准定在生产额为 4 099 千克.

补充习题及解答

14. 设 $X \sim N(\mu,\sigma^2)$,求 $E(|X-\mu|)^k$.

解

$$E(|X-\mu|)^k=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}|x-\mu|^k e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx.$$

若令 $y=(x-\mu)/\sigma$,$\sigma dy=dx$,可得

$$E(|X-\mu|)^k=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sigma^k|y|^k e^{-y^2/2}dy=\sqrt{\frac{2}{\pi}}\int_0^{\infty}\sigma^k y^k e^{-y^2/2}dy.$$

再令 $y^2/2=t$,$y=(2t)^{1/2}$,$dy=(2t)^{-1/2}dt$,可得

$$E(|X-\mu|)^k=\frac{\sigma^k}{\sqrt{\pi}}2^{k/2}\int_0^{\infty}t^{k/2-1/2}e^{-t}dt=\frac{\sigma^k}{\sqrt{\pi}}2^{k/2}\Gamma\left(\frac{k+1}{2}\right).$$

当 k 为偶数时,$E(|X-\mu|^k)=(k-1)!!\ \sigma^k$;

当 k 为奇数时,$E(|X-\mu|^k)=(k-1)!!\ \sigma^k\cdot\sqrt{2/\pi}$,其中 $n!!$ 表示不超过 n 且与 n 有相同奇偶性的所有正整数的乘积.

15. 设随机变量 X 服从参数为 λ 的泊松分布,试求 X 的前四阶原点矩、中心矩、偏度与峰度.

解 分几步进行.

(1) 先求 k 阶原点矩的递推公式.按定义

$$\mu_k=\sum_{x=0}^{\infty}x^k\frac{\lambda^x}{x!}e^{-\lambda},\qquad \lambda>0.$$

显然 $\mu_0=1$,而当 $k\geqslant 1$ 时有

$$\begin{aligned}\mu_k&=\lambda\sum_{x=1}^{\infty}[(x-1)+1]^{k-1}\frac{\lambda^{x-1}}{(x-1)!}e^{-\lambda}\\&=\lambda\sum_{x=1}^{\infty}\sum_{i=0}^{k-1}\binom{k-1}{i}(x-1)^i\frac{\lambda^{x-1}}{(x-1)!}e^{-\lambda}\\&=\lambda\sum_{i=0}^{k-1}\binom{k-1}{i}\sum_{x=1}^{\infty}(x-1)^i\frac{\lambda^{x-1}}{(x-1)!}e^{-\lambda}=\lambda\sum_{i=0}^{k-1}\binom{k-1}{i}\mu_i.\end{aligned}$$

(2) 由此递推公式可导出前四阶原点矩.

$\mu_1=\lambda\mu_0=\lambda$.

$\mu_2=\lambda(\mu_0+\mu_1)=\lambda(1+\lambda)$.

$\mu_3=\lambda(\mu_0+2\mu_1+\mu_2)=\lambda(1+3\lambda+\lambda^2)$.

$$\begin{aligned}\mu_4&=\lambda(\mu_0+3\mu_1+3\mu_2+\mu_3)=\lambda[1+3\lambda+3\lambda(1+\lambda)+\lambda(\lambda^2+3\lambda+1)]\\&=\lambda(1+7\lambda+6\lambda^2+\lambda^3).\end{aligned}$$

(3) 再计算前四阶中心矩;

$\nu_1=0$.

$\nu_2=\lambda$.

$\nu_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \lambda(1 + 3\lambda + \lambda^2) - 3\lambda^2(1 + \lambda) + 2\lambda^3 = \lambda.$

$\nu_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4$

$\quad = \lambda(1 + 7\lambda + 6\lambda^2 + \lambda^3) - 4\lambda^2(1 + 3\lambda + \lambda^2) + 6\lambda^3(1 + \lambda) - 3\lambda^4$

$\quad = \lambda(1 + 3\lambda).$

（4）最后计算偏度 β_S 与峰度 β_k.

$$\beta_S = \frac{\nu_3}{\nu_2^{3/2}} = \frac{\lambda}{\lambda^{3/2}} = \frac{1}{\sqrt{\lambda}} > 0.$$

所以泊松分布是正偏分布，λ 愈小偏度愈大.

$$\beta_k = \frac{\nu_4}{\nu_2^2} - 3 = \frac{\lambda(1 + 3\lambda)}{\lambda^2} - 3 = \frac{1}{\lambda} > 0.$$

所以泊松分布比标准正态分布更尖峭一些，λ 愈小分布愈尖峭.

16. 设随机变量 X 服从二项分布 $b(n,p)$，试求 X 的前四阶原点矩、中心矩、偏度与峰度.

解　分几步进行.

（1）先求 k 阶原点矩的递推公式.记

$$J_k(n) = \sum_{x=0}^{n} x^k \binom{n}{x} p^x(1 - p)^{n-x}.$$

显然有 $J_0(n) = J_0(n - i) = 1, i = 1,2,\cdots,n$.而当 $k \geqslant 1$ 时有

$$\begin{aligned} J_k(n) &= np\sum_{x=1}^{n}[(x-1)+1]^{k-1}\binom{n-1}{x-1}p^{x-1}(1-p)^{n-x} \\ &= np\sum_{x=1}^{n}\sum_{i=0}^{k-1}\binom{k-1}{i}(x-1)^i\binom{n-1}{x-1}p^{x-1}(1-p)^{n-x} \\ &= np\sum_{i=0}^{k-1}\binom{k-1}{i}\sum_{x=1}^{n}(x-1)^i\binom{n-1}{x-1}p^{x-1}(1-p)^{(n-1)-(x-1)} \\ &= np\sum_{i=0}^{k-1}\binom{k-1}{i}J_i(n-1). \end{aligned}$$

（2）由此递推公式可导出前四阶原点矩.

$\mu_1 = J_1(n) = npJ_0(n - 1) = np.$

$\mu_2 = J_2(n) = np[J_0(n - 1) + J_1(n - 1)] = np[1 + (n - 1)p].$

$\mu_3 = J_3(n) = np[J_0(n - 1) + 2J_1(n - 1) + J_2(n - 1)]$

$\quad = np\{1 + 2(n - 1)p + (n - 1)p[1 + (n - 2)p]\}$

$\quad = np[1 + 3(n - 1)p + (n - 1)(n - 2)p^2].$

$\mu_4 = J_4(n) = np[J_0(n - 1) + 3J_1(n - 1) + 3J_2(n - 1) + J_3(n - 1)]$

$\quad = np[1 + 7(n - 1)p + 6(n - 1)(n - 2)p^2 + (n - 1)(n - 2)(n - 3)p^3].$

（3）再计算前四阶中心矩：

$\nu_1 = 0.$

$\nu_2 = np(1 - p).$

$\nu_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$

$$= np[1+3(n-1)p+(n-1)(n-2)p^2] - 3n^2p^2[1+(n-1)p] + 2n^3p^3$$
$$= np(1-p)(1-2p).$$

$$\nu_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4 = np(1-p)[1+3(n-2)p(1-p)].$$

（4）最后计算偏度 β_S 与峰度 β_k.

$$\beta_S = \frac{\nu_3}{\nu_2^{3/2}} = \frac{1-2p}{\sqrt{np(1-p)}} \begin{cases} >0, & p<\dfrac{1}{2}, \\ =0, & p=\dfrac{1}{2}, \\ <0, & p>\dfrac{1}{2}, \end{cases}$$

由此可见：二项分布在 $p = 1/2$ 时是对称分布；当 $p < 1/2$ 时，二项分布正偏；当 $p > 1/2$ 时，二项分布负偏.

$$\beta_k = \frac{\nu_4}{\nu_2^2} - 3 = \frac{np(1-p)[1+3(n-2)p(1-p)]}{n^2p^2(1-p)^2} - 3$$
$$= \frac{1}{np(1-p)} - \frac{6}{n} = \frac{1}{n}\left(\frac{1}{p(1-p)} - 6\right).$$

更细致地讨论会发现：(i) 当 p 在区间 $[0.5-\sqrt{3}/6, 0.5+\sqrt{3}/6] \approx [0.21, 0.79]$ 内，$\beta_k \leqslant 0$，此时二项分布比标准正态分布更平坦，譬如在 $p=0.5$ 时，$\beta_k = -2/n < 0$，此时二项分布是对称的，且比标准正态分布更平坦；(ii) 当 p 在区间 $[0.5-\sqrt{3}/6, 0.5+\sqrt{3}/6] \approx [0.21, 0.79]$ 外，$\beta_k > 0$，此时二项分布比标准正态分布更尖峭.