Chapter 4

正态分布

4.1 设随机变量X服从标准正态分布$N(0,1)$，求下列概率：

(2) $P\{X > 2.5\}$;

(3) $P\{|X| < 1.68\}$;

解： 查表得：

(2) $P\{X > 2.5\} = 1 - P\{X \le 2.5\} = 1 - 0.9938 = 0.0062$；

(3) $P\{|X| < 1.68\} = 2\Phi(1.68) - 1 = 2 \times 0.9535 - 1 = 0.9070$；

□

4.2 设随机变量X服从正态分布$N(1, 2^2)$，求下列概率：

(2) $P\{-1.6 \le X < 5.8\}$;

解：

(2)

$$\begin{aligned} P\{-1.6 \le X < 5.8\} &= \Phi\left(\frac{5.8-1}{2}\right) - \Phi\left(\frac{-1.6-1}{2}\right) = \Phi(2.4) - \Phi(-1.3) \\ &= 0.9918 - (1 - 0.9032) = 0.8950. \end{aligned}$$

□

4.3 已知某种机械零件的直径（单位：mm）服从正态分布$N(100, 0.6^2)$，规定直径在范围(100 ± 1.2)mm内为合格品，求这种机械零件的不合格率。

解： 用X表示这种机械零件的直径，根据题意$X \sim N(100, 0.6^2)$，这种机械零件的不合格率为

$$P\{|x-100| \geq 1.2\} = 1 - P\{|x-100| < 1.2\} = 1 - P\left\{\left|\frac{x-100}{0.6}\right| < 2\right\}$$
$$= 1 - [2\Phi(2) - 1] = 2[1-\Phi(2)] = 2[1-0.9772] = 0.0456.$$

□

4.5 某次考试的成绩X近似服从正态分布，平均分为75分。已知95分以上的考生比例为2.3%，求这次考试的不及格率（60分及以上为及格）。

解： 设$X \sim N(75, \sigma^2)$。因为已知95分以上的考生比例为2.3%，即

$$P\{X > 95\} = 1 - P\{X \leq 95\} = 1 - P\left\{\frac{X-75}{\sigma} \leq \frac{95-75}{\sigma}\right\} = 1 - \Phi(\frac{20}{\sigma}) = 0.023.$$

所以$\Phi(\frac{20}{\sigma}) = 0.9770$，查表得$\frac{20}{\sigma} \approx 2$，即$\sigma \approx 10$。从而这次考试的不及格率为

$$P\{X < 60\} = P\left\{\frac{X-75}{10} < \frac{60-75}{10}\right\} = \Phi(-1.5) = 1-\Phi(1.5) = 1-0.9332 = 0.0668.$$

□

4.7 设随机变量$X \sim N(\mu, \sigma^2)$，求随机变量函数$Y = e^X$的概率密度。（所得的概率分布叫做对数正态分布。）

解： 已知$X \sim N(\mu, \sigma^2)$，则X的概率密度

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty.$$

考虑随机变量函数$Y = e^X$的分布函数

$$F_Y(y) = P\{Y \leq y\} = P\{e^X \leq y\}.$$

当$y\le 0$时，有$F_Y(y)=0$；当$y>0$时，有

$$F_Y(y)=P\{X\le \ln y\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\ln y}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx.$$

所以，Y的分布函数为

$$F_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma}\displaystyle\int_{-\infty}^{\ln y}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx, & y>0,\\ 0, & y\le 0.\end{cases}$$

对y求导得Y的概率密度

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sigma y}e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}, & y>0,\\ 0, & y\le 0.\end{cases}$$

□

4.8 设随机变量$X\sim N(0,\sigma^2)$，求随机变量函数$Y=|X|$的概率密度、数学期望与方差。

解： 因为随机变量$X\sim N(0,\sigma^2)$，则X的概率密度函数是

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}},\quad -\infty<x<+\infty.$$

记Y的概率分布函数、概率密度函数分别为$F_Y(y)$、$f_Y(y)$。注意到$Y=|X|$的取值范围为$[0,+\infty)$，当$y<0$时$F_Y(y)=0$。当$y>0$时，

$$\begin{aligned}F_Y(y)&=P(Y\le y)=P(|X|\le y)=P(-y\le X\le y)\\&=\int_{-y}^{y}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx=\int_0^y\frac{2}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx,\end{aligned}$$

所以，Y的分布函数为

$$F_Y(y)=\begin{cases}\displaystyle\int_0^y\frac{2}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx, & y>0,\\ 0, & y\le 0.\end{cases}$$

对y求导得Y的概率密度

$$f_Y(y)=\begin{cases}\dfrac{2}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}, & y>0,\\ 0, & y\le 0.\end{cases}$$

由概率密度求Y的期望

$$E(Y)=\int_{-\infty}^{+\infty}yf(y)dy=\frac{2}{\sqrt{2\pi}\sigma}\int_{0}^{+\infty}ye^{-\frac{y^2}{2\sigma^2}}dy$$
$$=\frac{2\sigma}{\sqrt{2\pi}}\cdot\left(-e^{-\frac{y^2}{2\sigma^2}}\Big|_0^{+\infty}\right)=\frac{2\sigma}{\sqrt{2\pi}}.$$

其次由Y与X的关系得

$$E(Y^2)=E(X^2)=D(X)+[E(X)]^2=\sigma^2.$$

因此Y的方差为

$$D(Y)=E(Y^2)-[E(Y)]^2=\sigma^2-\frac{2}{\pi}\sigma^2.$$

□

4.9 设随机变量X服从标准正态分布$N(0,1)$，求随机变量函数$Y=X^n$（n是正整数）的数学期望与方差。

解： 已知$X\sim N(0,1)$，则X的概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad -\infty<x<+\infty.$$

随机变量函数$Y=X^n$的数学期望

$$E(Y)=E(X^n)=\int_{-\infty}^{+\infty}x^n\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}x^ne^{-\frac{x^2}{2}}dx.$$

当n为奇数时，因为反常积分绝对收敛，且被积函数为奇函数，所以有$E(Y)=0$；当n为偶数时，因为被积函数为偶函数，所以有

$$E(Y)=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}x^ne^{-\frac{x^2}{2}}dx.$$

做变量替换$t=\frac{x^2}{2}=t$，则$x=\sqrt{2t}$，$dx=\frac{\sqrt{2}}{2}t^{-\frac{1}{2}}dt$，

$$E(Y)=\sqrt{\frac{2}{\pi}}\int_{0}^{+\infty}(2t)^{\frac{n}{2}}e^{-t}\cdot\frac{\sqrt{2}}{2}t^{-\frac{1}{2}}dt$$
$$=\frac{2^{\frac{n}{2}}}{\sqrt{\pi}}\int_{0}^{+\infty}t^{\frac{n-1}{2}}e^{-t}dt$$
$$=\frac{2^{\frac{n}{2}}}{\sqrt{\pi}}\Gamma\left(\frac{n+1}{2}\right)=\frac{2^{\frac{n}{2}}}{\sqrt{\pi}}\cdot\frac{n-1}{2}\cdot\frac{n-3}{2}\cdot\cdots\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)$$
$$=\frac{2^{\frac{n}{2}}}{\sqrt{\pi}}\cdot\frac{(n-1)!!}{2^{\frac{n}{2}}}\sqrt{\pi}=(n-1)!!.$$

所以，Y的数学期望

$$E(Y)=\begin{cases}0, & \text{当}n\text{为奇数},\\ (n-1)!!, & \text{当}n\text{为偶数}.\end{cases}$$

又因为$2n$为偶数，所以直接利用上式得

$$E(Y^2)=E(X^{2n})=(2n-1)!!.$$

于是，Y的方差为

$$D(Y)=(2n-1)!!-[E(Y)]^2=\begin{cases}(2n-1)!!, & \text{当}n\text{为奇数},\\ (2n-1)!!-[(n-1)!!]^2, & \text{当}n\text{为偶数}.\end{cases}$$

□

注：该题直接用课本公式(4.9)、(4.10)亦可。

4.10 设随机变量X与Y独立，$X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$，求：

(1) 随机变量函数$Z_1=aX+bY$的数学期望与方差，其中a及b为常数；

(2) 随机变量函数$Z_2=XY$的数学期望与方差。

解：

(1) 由期望的线性性质：

$$EZ_1=E(aX+bY)=aEX+bEY=a\mu_1+b\mu_2.$$

由X与Y的独立性及方差的性质，

$$DZ_1=D(aX+bY)=a^2DX+b^2DY=a^2\sigma_1^2+b^2\sigma_2^2.$$

(2) 首先由方差的公式$DX=E(X^2)-(EX)^2$得$E(X^2)=DX+(EX)^2=\sigma_1^2+\mu_1^2$，同理可得$E(Y^2)=DY+(EY)^2=\sigma_2^2+\mu_2^2$。由$X$与$Y$的独立性（该性质蕴含$X^2$与$Y^2$也是独立的)、期望的乘积性质得：

$$\begin{aligned}EZ_2&=E(XY)=EX\cdot EY=\mu_1\mu_2,\\ E(Z_2^2)&=E(X^2Y^2)=E(X^2)\cdot E(Y^2)=(\sigma_1^2+\mu_1^2)(\sigma_2^2+\mu_2^2),\\ D(Z_2^2)&=E(Z_2^2)-(EZ_2)^2=\sigma_1^2\sigma_2^2+\mu_1^2\sigma_2^2+\mu_2^2\sigma_1^2.\end{aligned}$$

□

4.11 设随机变量X服从标准正态分布，概率密度为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2},$$

随机变量$Y = X^n$（n是正整数），求X与Y的相关系数。

解： 已知随机变量$X \sim N(0,1)$，则X的数学期望与方差分别是

$$E(X) = 0, \quad D(X) = 1.$$

习题4.7已求得随机变量$Y = X^n$的数学期望与方差分别是

$$E(Y) = \begin{cases} 0, & \text{当}n\text{为奇数}, \\ (n-1)!!, & \text{当}n\text{为偶数}. \end{cases}$$

$$D(Y) = (2n-1)!! - [E(Y)]^2.$$

易知

$$E(XY) = E(X^{n+1}) = \begin{cases} n!!, & \text{当}n\text{为奇数}, \\ 0, & \text{当}n\text{为偶数}. \end{cases}$$

则X与Y的协方差

$$\mathrm{cov}(X,Y) = E(XY) - 0 \times E(Y) = \begin{cases} n!!, & \text{当}n\text{为奇数}, \\ 0, & \text{当}n\text{为偶数}. \end{cases}$$

X与Y的相关系数

$$R(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \begin{cases} \dfrac{n!!}{\sqrt{(2n-1)!!}}, & \text{当}n\text{为奇数}, \\ 0, & \text{当}n\text{为偶数}. \end{cases}$$

□

4.12 设二维随机变量(X,Y)服从二维正态分布，已知$E(X) = E(Y) = 0$, $D(X) = 16$, $D(Y) = 25$, $\mathrm{cov}(X,Y) = 12$，求(X,Y)的联合概率密度。

解: 由已知条件得X与Y的相关系数为

$$R(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}=0.6.$$

从而(X,Y)服从二维正态分布$N(0,0,16,25,0.6)$，其联合概率密度为

$$\begin{aligned}f(x,y)&=\frac{1}{32\pi}\exp\left\{-\frac{1}{1.28}\left[\frac{x^2}{16}-\frac{1.2xy}{20}+\frac{y^2}{25}\right]\right\}\\&=\frac{1}{32\pi}\exp\left\{-\frac{25}{32}\left[\frac{x^2}{16}-\frac{3xy}{50}+\frac{y^2}{25}\right]\right\}.\end{aligned}$$

□

4.15 设随机变量X与Y独立，$X\sim N(0,1)$，$Y\sim N(1,2^2)$，求随机变量函数$Z=2X-Y+3$的概率密度。

解: 由于X与Y独立，故$Z\sim N(2,8)$，其概率密度为

$$\frac{1}{4\sqrt{\pi}}e^{-\frac{(z-2)^2}{16}}.$$

□

4.16 设二维随机变量(X,Y)的联合概率密度为

$$f(x,y)=\frac{1}{\pi}e^{-x^2-y^2+2x-1},$$

求随机变量函数$Z=X-2Y$的概率密度。

解: 整理(X,Y)的联合概率密度

$$f(x,y)=\frac{1}{\pi}e^{-x^2-y^2+2x-1}=\frac{1}{\pi}e^{-[(x-1)^2+y^2]}$$

并与二元正态分布的密度函数

$$\begin{aligned}f(x,y)=&\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\\&\cdot\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}\end{aligned}$$

相比较，可知(X,Y)服从二元正态分布，且参数$\rho=0$。代入后继续比较函数各部分得$\mu_1=1$，$\mu_2=0$，$\sigma_1^2=\sigma_2^2=\frac{1}{2}$。故$(X,Y)\sim N(1,0,\frac{1}{2},\frac{1}{2},0)$，边缘分布为$X\in N(1,\frac{1}{2})$，$Y\in N(0,\frac{1}{2})$，$X$与$Y$相互独立。因为$Z=X-2Y$，$Z$也服从正态分布，且

$$E(Z)=E(X)-2E(Y)=1,\qquad D(Z)=D(X)+4D(Y)=\frac{5}{2},$$

其概率密度为

$$\frac{1}{\sqrt{5\pi}}e^{-\frac{(z-1)^2}{5}}.$$

□

4.20 已知一本300页的书中每页印刷错误的个数服从泊松分布$P(0.2)$，求这本书的印刷错误总数不多于70的概率。

解： 记X_i为第i页中的印刷错误的个数$(i=1,2,\cdots,300)$，则所有X_i独立同分布，有共同的期望0.2与方差0.2。再记X为整本书的印刷错误总数。则

$$X=\sum_{i=1}^{300}X_i=X_1+X_2+\cdots+X_{300},$$
$$EX=\sum_{i=1}^{300}EX_i=300\times0.2=60,\quad DX=\sum_{i=1}^{300}DX_i=60.$$

由中心极限定理，可以近似认为$X\sim N(60,60)$，故这本书的印刷错误总数不多于70的概率为

$$\begin{aligned}P\{X\le70\}&=P\left\{\frac{X-60}{\sqrt{60}}\le\frac{70-60}{\sqrt{60}}\right\}\approx P\left\{\frac{X-60}{\sqrt{60}}\le1.29\right\}\\&=\Phi(1.29)\approx0.9015.\end{aligned}$$

□

4.22 在习题3.30中，利用棣莫弗—拉普拉斯定理估计所求的概率。

解： 设随机变量X_n表示在n次重复独立试验中事件A发生的次数，则

$$X_n\sim B(n,p),$$

其中p是事件A在每次试验中发生的概率。事件A在n次试验中发生的频率$f_n(A)$与其概率p之差的绝对值小于0.01的概率

$$\begin{aligned} P\{|f_n(A)-p|<0.01\} &= P\left\{\left|\frac{X_n}{n}-p\right|<0.01\right\} \\ &= P\left\{\left|\frac{X_n-np}{\sqrt{np(1-p)}}\right|<\frac{0.01\sqrt{n}}{\sqrt{p(1-p)}}\right\}. \end{aligned}$$

因为$n=10000$充分大，所以由棣莫弗—拉普拉斯中心极限定理知：$\dfrac{X_n-np}{\sqrt{np(1-p)}}$近似服从标准正态分布$N(0,1)$，于是

$$P(|f_n(A)-p|<0.01)\approx 2\Phi\left(\frac{0.01\times 100}{\sqrt{p(1-p)}}\right)-1.$$

因为$p(1-p)\le\dfrac{1}{4}$，$\dfrac{1}{\sqrt{p(1-p)}}\ge 2$，所以有

$$\begin{aligned} P(|f_n(A)-p|<0.01) &\ge 2\Phi(2)-1 \\ &= 2\times 0.9772-1=0.9544. \end{aligned}$$

□

4.23 某单位设置一台电话总机，共有200个分机。设每个分机有5%的时间要使用外线通话，并且各个分机使用外线与否是互相独立的。该单位需要多少条外线才能保证每个分机要使用外线时可供使用的概率达到0.9？

解： 设随机变量X表示同时要使用外线通话的分机数，则

$$X\sim B(n,p),$$

其中$n=200$为分机总数，$p=0.05$为每个分机要使用外线通话的概率。假设该单位共有k条外线，则按题意应有

$$P\{X\le k\}\ge 0.9.$$

因为$n=200$充分大，所以由棣莫弗—拉普拉斯中心极限定理知：$\dfrac{X-np}{\sqrt{np(1-p)}}$近

似服从标准正态分布$N(0,1)$，于是

$$
\begin{aligned}
P\{X \le k\} &= P\left\{\frac{X-np}{\sqrt{np(1-p)}} \le \frac{k-np}{\sqrt{np(1-p)}}\right\} \\
&\approx \Phi\left(\frac{k-np}{\sqrt{np(1-p)}}\right) = \Phi\left(\frac{k-10}{\sqrt{9.5}}\right).
\end{aligned}
$$

由此得

$$\Phi\left(\frac{k-10}{\sqrt{9.5}}\right) \ge 0.9.$$

查表有$\Phi(1.28) \approx 0.9$，从而有

$$\frac{k-10}{\sqrt{9.5}} \ge 1.28,$$

解得$k \ge 13.945$，即至少需要14条外线才能保证每个分机要使用外线时可供使用的概率达到0.9。 □