§3.4 随机变量函数的分布

华东师范大学统计学院

華東師範大學·统计
STATISTICS, EAST CHINA NORMAL UNIVERSITY

目标

- 已知X的分布，求$Y = f(X)$的分布
- 已知(X, Y)的分布，求$Z = g(X, Y)$的分布
- 已知$\boldsymbol{X} = (X_1, \cdots, X_n)$的分布，

  求$\boldsymbol{Y} = (h_1(X_1, \cdots, X_n), \cdots, h_m(X_1, \cdots, X_n))$的分布

主要内容

1. 单个随机变量函数的分布

主要内容

1. 单个随机变量函数的分布
2. 随机向量函数的分布
   - 离散情形
   - 连续情形
   - 最值函数的分布

主要内容

1. 单个随机变量函数的分布
2. 随机向量函数的分布
   - 离散情形
   - 连续情形
   - 最值函数的分布
3. 随机向量变换的分布

单个随机变量函数的分布

目的: 已知随机变量X的分布, 求$Y = f(X)$ 的分布, 其中f为$\mathbb{R}$到$\mathbb{R}$上的可测函数.

对任意的$y \in \mathbb{R}$,

$$\{Y \le y\} = \{f(X) \le y\} = \{X \in D_y\},$$

其中$D_y = \{x : f(x) \le y\}$.

对任意的$y \in \mathbb{R}$,

$$\{Y \le y\} = \{f(X) \le y\} = \{X \in D_y\},$$

其中$D_y = \{x : f(x) \le y\}$.

假设f严格单调增, 则

$$D_y = \{x : f(x) \le y\} = \{x : x \le f^{-1}(y)\},$$

对任意的$y \in \mathbb{R}$,

$$\{Y \le y\} = \{f(X) \le y\} = \{X \in D_y\},$$

其中$D_y = \{x : f(x) \le y\}$.

假设f严格单调增, 则

$$D_y = \{x : f(x) \le y\} = \{x : x \le f^{-1}(y)\},$$

于是Y的分布函数为

$$F_Y(y) = P(Y \le y) = P(X \in D_y) = P(X \le f^{-1}(y)) = F_X(f^{-1}(y)).$$

定理1

已知随机变量X具有分布列$\{p_k, k = 1, 2, \ldots\}$(离散情形)或概率密度函数$p_X(x)$(连续情形),

则$Y = f(X)$的分布函数为

$$F_Y(y) = \begin{cases} \sum_{x_k \in D_y} p_k, & \text{离散情形}, \\ \int_{D_y} p_X(x)\mathrm{d}x, & \text{连续情形}. \end{cases}$$

例1

设随机变量X具有概率密度函数$p_X(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & 其他. \end{cases}$

求随机变量$Y = 3X^2 + 5$的分布函数.

解: 对任意的$y \in \mathbb{R}$, 记$D_y = \{x \in \mathbb{R} : 3x^2 + 5 \le y\}$, 则

$$D_y \cap (0,1) = \begin{cases} \emptyset, & y < 5, \\ \left(0, \sqrt{\dfrac{y-5}{3}}\right], & 5 \le y < 8, \\ (0,1), & y \ge 8. \end{cases}$$

于是, $Y = 3X^2 + 5$的分布函数为

$$F_Y(y) = P(Y \le y) = P(3X^2 + 5 \le y) = P(X \in D_y) = \int_{D_y} p_X(x)\mathrm{d}x$$

$$= \int_{D_y \cap (0,1)} 2x\mathrm{d}x = \begin{cases} 0, & y < 5, \\ \displaystyle\int_0^{\sqrt{\frac{y-5}{3}}} 2x\mathrm{d}x, & 5 \le y < 8, \\ 1, & y \ge 8 \end{cases} = \begin{cases} 0, & y < 5, \\ \dfrac{y-5}{3}, & 5 \le y < 8, \\ 1, & y \ge 8. \end{cases}$$

例2

设随机变量X服从标准正态分布$N(0,1)$, 求随机变量$Y=X^2$ 的分布.

解: 设$Y = X^2$的分布函数为$F_Y(y)$, 则

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = \begin{cases} P(|X| \le \sqrt{y}), & y \ge 0, \\ 0, & y < 0 \end{cases}$$

$$= \begin{cases} \Phi(\sqrt{y}) - \Phi(-\sqrt{y}), & y \ge 0, \\ 0, & y < 0 \end{cases} = \begin{cases} 2\Phi(\sqrt{y}) - 1, & y \ge 0, \\ 0, & y < 0. \end{cases}$$

其中Φ是标准正态分布的分布函数.

进一步地, 我们可以求出Y的概率密度函数

$$p_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}y}\left(2\Phi(\sqrt{y}) - 1\right), & y > 0, \\ 0, & y \le 0 \end{cases}$$

$$= \begin{cases} \varphi(\sqrt{y})/\sqrt{y}, & y > 0, \\ 0, & y \le 0 \end{cases} = \begin{cases} \dfrac{(1/2)^{1/2}}{\Gamma(1/2)} y^{1/2-1} e^{-y/2}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

其中$\varphi(t) = \dfrac{1}{\sqrt{2\pi}} e^{-t^2/2}$ 是标准正态分布的概率密度函数.

进一步地, 我们可以求出Y的概率密度函数

$$p_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}y}\left(2\Phi(\sqrt{y}) - 1\right), & y > 0, \\ 0, & y \le 0 \end{cases}$$

$$= \begin{cases} \varphi(\sqrt{y})/\sqrt{y}, & y > 0, \\ 0, & y \le 0 \end{cases} = \begin{cases} \dfrac{(1/2)^{1/2}}{\Gamma(1/2)} y^{1/2-1} e^{-y/2}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

其中$\varphi(t) = \dfrac{1}{\sqrt{2\pi}} e^{-t^2/2}$ 是标准正态分布的概率密度函数.

对照Gamma分布的概率密度函数的形式知, Y服从Gamma分布$Ga(1/2, 1/2)$, 即自由度为1的卡方分布$\chi^2(1)$.

定理2

设随机变量X的分布函数$F(x)$是严格单调增的连续函数, 则随机变量$F(X)$服从均匀分布$U(0,1)$.

证明.

设$Y = F(X)$的分布函数为$F_Y(y)$, 则

$$F_Y(y) = P(Y \le y) = P(F(X) \le y) = \begin{cases} 0, & y < 0, \\ P(X \le F^{-1}(y)), & 0 \le y < 1, \\ 1, & y \ge 1 \end{cases}$$

$$= \begin{cases} 0, & y < 0, \\ F(F^{-1}(y)), & 0 \le y < 1, \\ 1, & y \ge 1 \end{cases} = \begin{cases} 0, & y < 0, \\ y, & 0 \le y < 1, \\ 1, & y \ge 1. \end{cases}$$

即Y服从均匀分布$U(0,1)$. □

产生已知分布的随机数

注记1

任何一个连续随机变量可以通过分布函数与均匀分布建立联系, 这在随机模拟中大有用处. 例如欲利用计算机产生一批数据来模拟成人的身高. 设成人身高服从已知分布$F(x)$, 由此定理我们可以先产生均匀分布的随机数$y_1, \ldots, y_n$, 通过解方程$F(x_i) = y_i$ 即可得到服从已知分布$F(x)$ 的随机数$x_1, \ldots, x_n$, 而均匀分布的随机数在几乎所有的统计软件中都能产生.

离散型

当随机变量X是离散型, 且分布列已知时, 我们知道$Y=f(X)$也是离散型的, 这时候求出Y的分布列并不困难.

例如, 设X的分布列为

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

则$Y=f(X)$的分布列为

$f(X)$	$f(x_1)$	$f(x_2)$	$\cdots$	$f(x_n)$	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

.

这里可能某些$f(x_i)$取值相同, 需要将相同项进行合并.

例3

设X的分布列为

X	-1	0	1	2
P	$1/4$	$1/4$	$1/4$	$1/4$

. 求$Y=|X|$的分布列.

解.

显然Y的可能取值为$0,1,2$, 且其分布列为

$$P(Y=0)=P(X=0)=\frac{1}{4},$$

$$P(Y=1)=P(|X|=1)=P(X=-1)+P(X=1)=\frac{1}{4}+\frac{1}{4}=\frac{1}{2},$$

$$P(Y=2)=P(X=2)=\frac{1}{4}.$$

例3

设X的分布列为

X	-1	0	1	2
P	$1/4$	$1/4$	$1/4$	$1/4$

. 求$Y=|X|$的分布列.

解.

显然Y的可能取值为$0,1,2$, 且其分布列为

$$P(Y=0)=P(X=0)=\frac{1}{4},$$

$$P(Y=1)=P(|X|=1)=P(X=-1)+P(X=1)=\frac{1}{4}+\frac{1}{4}=\frac{1}{2},$$

$$P(Y=2)=P(X=2)=\frac{1}{4}.$$ 即

Y	0	1	2
P	$1/4$	$1/2$	$1/4$

. □

连续型

定理3

设随机变量X的概率密度函数为$p_X(x)$, $y = f(x)$是严格单调函数, 且其反函数$x = h(y)$ 有连续的导函数, 则随机变量$Y = f(X)$的概率密度函数为

$$p_Y(y) = \begin{cases} p_X(h(y))|h'(y)|, & y \in E, \\ 0, & \text{其他}. \end{cases}$$

其中E是使得f的值域.

证明: 只需注意到f严格单调时,

$$D_y = \{x : f(x) \le y\} = \{x : x \le h(y)\} \text{ 或}$$

$$D_y = \{x : f(x) \le y\} = \{x : x \ge h(y)\}.$$

证明: 只需注意到f严格单调时,

$$D_y = \{x : f(x) \le y\} = \{x : x \le h(y)\} \text{ 或}$$

$$D_y = \{x : f(x) \le y\} = \{x : x \ge h(y)\}.$$

于是

$$F_Y(y) = P(Y \le y) = P(f(X) \le y) = P(X \in D_y)$$

$$= \begin{cases} P(X \le h(y)), & f\text{单调增}, \\ P(X \ge h(y)), & f\text{单调降}. \end{cases} = \begin{cases} F_X(h(y)), & f\text{单调增}, \\ 1 - F_X(h(y)), & f\text{单调降}. \end{cases}$$

求导即得概率密度函数.

例4

设X服从柯西分布, 即其概率密度函数为$p_X(x) = \dfrac{1}{\pi(1+x^2)}$, 求$Y = e^X$的分布.

解.

显然函数$y = e^x$的值域为$(0, \infty)$, 于是$Y = e^X$的概率密度函数为

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

□

定理4

正态分布的线性不变性

设随机变量X服从正态分布$N(\mu, \sigma^2)$, 则当$a \neq 0$时, 随机变量$Y = aX + b$服从正态分布$N(a\mu + b, a^2\sigma^2)$.

定理4

正态分布的线性不变性

设随机变量X服从正态分布$N(\mu,\sigma^2)$, 则当$a \neq 0$时, 随机变量$Y = aX + b$服从正态分布$N(a\mu + b, a^2\sigma^2)$.

特别地, 令$a = \dfrac{1}{\sigma}, b = -\dfrac{\mu}{\sigma}$, 我们有

推论5

若$X \sim N(\mu,\sigma^2)$, 则$\dfrac{X-\mu}{\sigma} \sim N(0,1)$.

定理6

设随机变量X服从$Gamma$分布$Ga(\alpha,\lambda)$, $c>0$, 则随机变量cX服从$Gamma$分布$Ga\left(\alpha,\dfrac{\lambda}{c}\right)$.

推论7

设随机变量X服从正态分布$N(0,\sigma^2)$, 则X^2服从$Gamma$分布$Ga\left(\dfrac{1}{2},\dfrac{1}{2\sigma^2}\right)$.

证明.

显然$X/\sigma\sim N(0,1)$. 于是$X^2/\sigma^2\sim\chi^2(1)$, 亦即$\dfrac{X^2}{\sigma^2}\sim Ga\left(\dfrac{1}{2},\dfrac{1}{2}\right)$.故$X^2=\sigma^2\cdot\dfrac{X^2}{\sigma^2}\sim Ga\left(\dfrac{1}{2},\dfrac{1}{2\sigma^2}\right)$. □

随机向量函数的分布

目的: 已知随机向量$\boldsymbol{X} = (X_1, \cdots, X_n)$的联合分布, 求$Z = g(X_1, \cdots, X_n)$ 的分布, 其中g为$\mathbb{R}^n$到$\mathbb{R}$上的可测函数.

随机向量函数的分布

目的: 已知随机向量$\boldsymbol{X} = (X_1, \cdots, X_n)$的联合分布, 求$Z = g(X_1, \cdots, X_n)$ 的分布, 其中g为$\mathbb{R}^n$到$\mathbb{R}$上的可测函数.

着重考虑$n = 2$的情形.

定理8

设随机变量(X,Y)具有分布列$\{p_{ij}, i,j=1,2,\ldots\}$(离散情形)或概率密度函数$p(x,y)$(连续情形), 则随机变量$Z=g(X,Y)$的分布函数为

$$F_Z(z)=P((X,Y)\in D_z)=\begin{cases}\displaystyle\sum_{(i,j):(x_i,y_j)\in D_z} p_{ij}, & \text{离散情形},\\ \displaystyle\iint_{D_z} p(x,y)\mathrm{d}x\mathrm{d}y, & \text{连续情形}.\end{cases}$$

其中$D_z=\{(x,y):g(x,y)\le z\}$.

定理8

设随机变量(X,Y)具有分布列$\{p_{ij}, i,j=1,2,\ldots\}$(离散情形)或概率密度函数$p(x,y)$(连续情形), 则随机变量$Z=g(X,Y)$的分布函数为

$$F_Z(z)=P((X,Y)\in D_z)=\begin{cases}\displaystyle\sum_{(i,j):(x_i,y_j)\in D_z} p_{ij}, & \text{离散情形},\\ \displaystyle\iint_{D_z} p(x,y)\mathrm{d}x\mathrm{d}y, & \text{连续情形}.\end{cases}$$

其中$D_z=\{(x,y): g(x,y)\le z\}$.

当$g(x,y)=x$时, 我们即可得到随机变量X的边际分布函数.

例5

设随机变量(X, Y)的联合分布列为

$X \backslash Y$	0	1
0	1/10	1/5
1	3/10	2/5

求随机变量X的分布函数$F_X(x)$.

解: X的分布函数为

$$F_X(x) = P(X \le x) = P((X,Y) \in (-\infty, x] \times \mathbb{R})$$

$$= \sum_{(i,j):(x_i,y_j)\in(-\infty,x]\times\mathbb{R}} p_{ij} = \begin{cases} 0, & x < 0, \\ \frac{1}{10} + \frac{1}{5}, & 0 \le x < 1, \\ \frac{1}{10} + \frac{1}{5} + \frac{3}{10} + \frac{2}{5}, & x \ge 1 \end{cases}$$

$$= \begin{cases} 0, & x < 0, \\ \frac{3}{10}, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

若(X,Y)是二维离散型随机变量, 则随机变量$Z=g(X,Y)$是一维离散型随机变量, 进而可以用分布列直接刻画其分布. 首先列出$Z=g(X,Y)$ 所有可能的取值$\{z_k, k=1,2,\ldots\}$, 接着求出每个概率$P(Z=z_k)$即可.

推论9

若二维离散型随机变量(X,Y)具有分布列$\{p_{ij}, i,j=1,2,\ldots\}$, 随机变量$Z=g(X,Y)$所有可能的取值为$\{z_k, k=1,2,\ldots\}$, 则$Z$的分布列为

$$P(Z=z_k)=\sum_{(i,j):g(x_i,y_j)=z_k} p_{ij}.$$

例6

若二维离散型随机变量(X, Y)有联合分布列如下:

$X \backslash Y$	0	1	2
0	1/2	1/8	1/4
1	1/16	1/16	0

求随机变量$Z = 2X - Y$的分布列.

解: 首先易知Z可能取值于$\{-2,-1,0,1,2\}$. 于是,

$$P(Z=-2)=P(X=0,Y=2)=\frac{1}{4},$$
$$P(Z=-1)=P(X=0,Y=1)=\frac{1}{8},$$
$$P(Z=1)=P(X=1,Y=1)=\frac{1}{16},$$
$$P(Z=2)=P(X=1,Y=0)=\frac{1}{16},$$

进而由正则性

$$P(Z=0)=1-\frac{1}{4}-\frac{1}{8}-\frac{1}{16}-\frac{1}{16}=\frac{1}{2}.$$

列表如下:

Z	-2	-1	0	1	2
P	$1/4$	$1/8$	$1/2$	$1/16$	$1/16$

例7

设随机变量X与Y相互独立, 且X, Y分别等可能地取值0和1.

求随机变量$Z = \max(X, Y)$ 的分布列.

解: 由题意知, X与Y独立同分布,

且$P(X=0)=P(X=1)=1/2$. Z的可能取值有0和1, 且

$$\begin{aligned}P(Z=0)=&P(\max(X,Y)=0)=P(X=0,Y=0)\\=&P(X=0)\cdot P(Y=0)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\\P(Z=1)=&1-P(Z=0)=1-\frac{1}{4}=\frac{3}{4}.\end{aligned}$$

解: 由题意知, X与Y独立同分布,

且$P(X=0)=P(X=1)=1/2$. Z的可能取值有0和1, 且

$$\begin{aligned}P(Z=0)=&P(\max(X,Y)=0)=P(X=0,Y=0)\\=&P(X=0)\cdot P(Y=0)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\\P(Z=1)=&1-P(Z=0)=1-\frac{1}{4}=\frac{3}{4}.\end{aligned}$$

故Z的分布列为

Z	0	1
P	1/4	3/4

.

卷积公式(离散情形)

定理10

设随机变量(X, Y)的联合分布列为

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j = 1, 2, \ldots,$$

则随机变量$Z = X + Y$的分布列为

$$P(Z = z_k) = \sum_i P(X = x_i, Y = z_k - x_i) = \sum_j P(X = z_k - y_j, Y = y_j)$$

其中$\{z_k, k = 1, 2, \ldots\}$是$Z$的可能取值集合.

卷积公式(离散情形)

定理10

设随机变量(X,Y)的联合分布列为

$$p_{ij} = P(X = x_i, Y = y_j), \quad i,j = 1,2,\ldots,$$

则随机变量$Z = X + Y$的分布列为

$$P(Z = z_k) = \sum_i P(X = x_i, Y = z_k - x_i) = \sum_j P(X = z_k - y_j, Y = y_j)$$

其中$\{z_k, k = 1,2,\ldots\}$是Z的可能取值集合.

特别地, 若X与Y相互独立, 则随机变量$Z = X + Y$的分布列为

$$\begin{aligned} P(Z = z_k) &= \sum_i P(X = x_i)P(Y = z_k - x_i) \\ &= \sum_j P(X = z_k - y_j)P(Y = y_j). \end{aligned}$$

证明.

按照X的取值, 将样本空间Ω写为$\Omega = \sum_i \{X = x_i\}$, 故由

$$\{Z = z_k\} = \sum_i \{Z = z_k, X = x_i\} = \sum_i \{X = x_i, Y = z_k - x_i\}$$

和概率的可列可加性即得

$$P(Z = z_k) = \sum_i P(X = x_i, Y = z_k - x_i).$$

剩下的同理可得. □

分布的可加性

定义1

若某类概率分布的独立随机变量和的分布仍是此类分布, 则称此类分布具有**可加性**.

分布的可加性

定义1

若某类概率分布的独立随机变量和的分布仍是此类分布, 则称此类分布具有**可加性**.

定理11

二项分布的可加性

若随机变量X服从二项分布$b(n,p)$, 随机变量Y服从二项分布$b(m,p)$, 且相互独立, 则随机变量$Z=X+Y$ 服从二项分布$b(n+m,p)$.

证明: 易知随机变量$Z = X + Y$ 的可能取值有$0, 1, \ldots, n + m$.

由卷积公式,

$$
\begin{aligned}
P(Z = k) &= \sum_i P(X = i)P(Y = k - i) \\
&= \sum_i C_n^i p^i (1-p)^{n-i} C_m^{k-i} p^{k-i} (1-p)^{m-(k-i)} \\
&= p^k (1-p)^{n+m-k} \sum_i C_n^i C_m^{k-i} \\
&= C_{n+m}^k p^k (1-p)^{n+m-k}, \qquad k = 0, 1, \ldots, n + m.
\end{aligned}
$$

证明: 易知随机变量$Z = X + Y$ 的可能取值有$0, 1, \ldots, n + m$.

由卷积公式,

$$
\begin{aligned}
P(Z = k) &= \sum_i P(X = i)P(Y = k - i) \\
&= \sum_i C_n^i p^i (1-p)^{n-i} C_m^{k-i} p^{k-i} (1-p)^{m-(k-i)} \\
&= p^k (1-p)^{n+m-k} \sum_i C_n^i C_m^{k-i} \\
&= C_{n+m}^k p^k (1-p)^{n+m-k}, \qquad k = 0, 1, \ldots, n + m.
\end{aligned}
$$

注记2

二项分布可以视为多个相互独立的两点分布的和.

定理12

若随机变量X服从泊松分布$P(\lambda_1)$, 随机变量Y服从泊松分布$P(\lambda_2)$, 且相互独立, 则随机变量$Z = X + Y$服从泊松分布$P(\lambda_1 + \lambda_2)$.

证明: 易知随机变量$Z = X + Y$ 的可能取值有$0, 1, 2, \ldots$.

$$P(Z = k) = \sum_i P(X = i)P(Y = k - i) = \sum_{i=0}^{k} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{k-i}}{(k-i)!} e^{-\lambda_2}$$

$$= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{k} C_k^i \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^i \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{k-i}$$

$$= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1+\lambda_2)}, \qquad k = 0, 1, \ldots$$

证明: 易知随机变量$Z = X + Y$ 的可能取值有$0, 1, 2, \ldots$.

$$P(Z = k) = \sum_i P(X = i)P(Y = k - i) = \sum_{i=0}^{k} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{k-i}}{(k-i)!} e^{-\lambda_2}$$

$$= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{k} C_k^i \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^i \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{k-i}$$

$$= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1+\lambda_2)}, \qquad k = 0, 1, \ldots$$

其中最后一个等号利用了二项分布$b\left(k, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$的正则性.

定理13

设随机变量(X, Y)的联合概率密度函数为$p(x, y)$, 则对任意的$t \in \mathbb{R}$, 随机变量$Z = tX + Y$的概率密度函数为

$$p_Z(z) = \int_{-\infty}^{\infty} p(x, z - tx)\mathrm{d}x.$$

证明: 设随机变量$Z = tX + Y$的分布函数为$F_Z(z)$, 令

$$D_z = \{(x, y) : tx+y \le z\} = \{(x, y) : -\infty < x < \infty, -\infty < y \le z-tx\}.$$

于是

$$\begin{aligned} F_Z(z) =& P(Z \le z) = P(tX + Y \le z) = P((X, Y) \in D_z) = \iint_{D_z} p(x, y)\mathrm{d}x\mathrm{d}y \\ =& \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{z-tx} p(x, y)\mathrm{d}y = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{z} p(x, u - tx)\mathrm{d}u \\ =& \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} p(x, u - tx)\mathrm{d}x \right) \mathrm{d}u, \end{aligned}$$

于是$Z = tX + Y$的概率密度函数为

$$
\begin{aligned}
p_Z(z) &= \frac{\mathrm{d}F_Z(z)}{\mathrm{d}z} = \frac{\mathrm{d}}{\mathrm{d}z}\left[\int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} p(x, u - tx)\mathrm{d}x\right)\mathrm{d}u\right] \\
&= \int_{-\infty}^{\infty} p(x, z - tx)\mathrm{d}x.
\end{aligned}
$$

定理14

设随机变量(X, Y)的联合概率密度函数为$p(x, y)$, 则对任意的$t \in \mathbb{R}$, 随机变量$Z = X + tY$的概率密度函数为

$$p_Z(z) = \int_{-\infty}^{\infty} p(z - ty, y)\mathrm{d}y.$$

卷积公式(连续情形)

推论15

设随机变量(X, Y)的联合概率密度函数为$p(x, y)$, 则随机变量$Z = X + Y$的概率密度函数为

$$p_Z(z) = \int_{-\infty}^{\infty} p(x, z - x)\mathrm{d}x = \int_{-\infty}^{\infty} p(z - y, y)\mathrm{d}y.$$

若X与Y相互独立, 其概率密度函数分别为$p_X(x)$和$p_Y(y)$, 则随机变量$Z = X + Y$ 的概率密度函数为

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(x)p_Y(z - x)\mathrm{d}x = \int_{-\infty}^{\infty} p_X(z - y)p_Y(y)\mathrm{d}y.$$

$X-Y$的分布

推论16

设随机变量(X,Y)的联合概率密度函数为$p(x,y)$, 则随机变量$Z=X-Y$的概率密度函数为

$$p_Z(z)=\int_{-\infty}^{\infty}p(x,x-z)\mathrm{d}x=\int_{-\infty}^{\infty}p(z+y,y)\mathrm{d}y.$$

例8

设二维随机变量(X, Y)的联合概率密度函数为

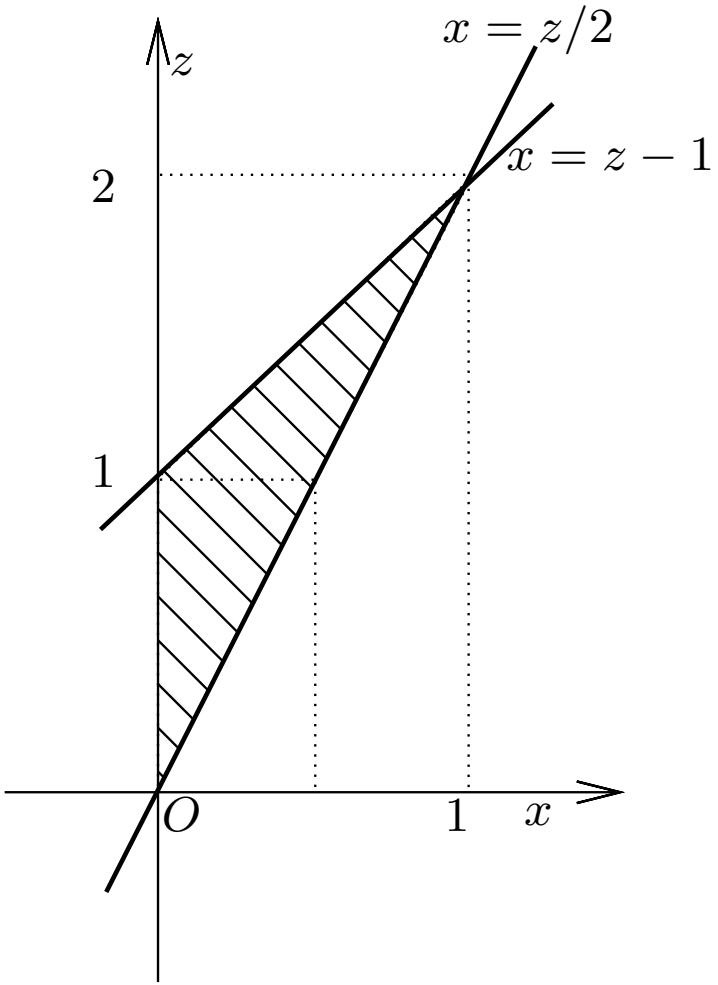
$$p(x, y) = \begin{cases} 2, & 若0 < x < y < 1, \\ 0, & 其他. \end{cases}$$

求随机变量$Z = X + Y$的概率密度函数$p_Z(z)$.

解: $Z = X + Y$的概率密度函数为$p_Z(z) = \int_{-\infty}^{\infty} p(x, z-x)\mathrm{d}x$. 此积分中被积函数$p(x, z-x)$的非零区域为

$$
\begin{aligned}
D =& \{(x,z) : 0 < x < z - x < 1\} \\
=& \{(x,z) : 0 < z \le 1, 0 < x < z/2\} \\
&+ \{(x,z) : 1 < z < 2, z - 1 < x < z/2\} \\
=& \{(x,z) : 0 < z < 2, \max(z-1, 0) < x < z/2\}.
\end{aligned}
$$

如图所示.


$x = z/2$
$x = z - 1$
z
2
1
O
1
x

于是, $Z = X + Y$的概率密度函数为

$$p_Z(z) = \int_{-\infty}^{\infty} p(x, z-x)\mathrm{d}x = \begin{cases} \int_{\max(z-1,0)}^{z/2} 2\mathrm{d}x, & 0 < z < 2, \\ 0, & z \le 0 \text{或} z \ge 2 \end{cases}$$

$$= \begin{cases} z - 2\max(z-1, 0), & 0 < z < 2, \\ 0, & z \le 0 \text{或} z \ge 2 \end{cases} = \begin{cases} z, & 0 < z \le 1, \\ 2 - z, & 1 < z < 2, \\ 0, & z \le 0 \text{或} z \ge \end{cases}$$

即随机变量$Z = X + Y$服从三角形分布.

Gamma分布的可加性

定理17

若随机变量X服从$Gamma$分布$Ga(\alpha_1, \lambda)$, Y服从$Gamma$分布$Ga(\alpha_2, \lambda)$, 且相互独立, 则随机变量$Z = X + Y$服从$Gamma$分布$Ga(\alpha_1 + \alpha_2, \lambda)$.

Gamma分布的可加性

定理17

若随机变量X服从$Gamma$分布$Ga(\alpha_1, \lambda)$, Y服从$Gamma$分布$Ga(\alpha_2, \lambda)$, 且相互独立, 则随机变量$Z = X + Y$服从$Gamma$分布$Ga(\alpha_1 + \alpha_2, \lambda)$.

注记3

若随机变量X服从卡方分布$\chi^2(n)$, Y服从卡方分布$\chi^2(m)$, 且相互独立, 则随机变量$Z = X + Y$服从卡方分布$\chi^2(n + m)$.

Gamma分布的可加性

定理17

若随机变量X服从$Gamma$分布$Ga(\alpha_1, \lambda)$, Y服从$Gamma$分布$Ga(\alpha_2, \lambda)$, 且相互独立, 则随机变量$Z = X + Y$服从$Gamma$分布$Ga(\alpha_1 + \alpha_2, \lambda)$.

注记3

若随机变量X服从卡方分布$\chi^2(n)$, Y服从卡方分布$\chi^2(m)$, 且相互独立, 则随机变量$Z = X + Y$服从卡方分布$\chi^2(n + m)$.

注记4

指数分布不具有可加性.

正态分布的可加性

例9

设X与Y相互独立, 且都服从标准正态分布$N(0,1)$,

求$Z = X + Y$的分布.

解:由卷积公式, $Z = X + Y$的概率密度函数为

$$p_Z(z) = \int_{-\infty}^{\infty} \varphi(x)\varphi(z-x)\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-(z-x)^2/2} \mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{2}} \exp\left(-\frac{z^2}{4}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\frac{1}{\sqrt{2}}} \exp\left[-\frac{\left(x-\frac{z}{2}\right)^2}{2\cdot\frac{1}{2}}\right] \mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{2}} \exp\left(-\frac{z^2}{2\cdot 2}\right),$$

故$Z = X + Y$服从正态分布$N(0, 2)$.

正态分布的可加性

定理18

设随机变量(X,Y)服从二维正态分布$N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$, 则随机变量$Z=X+Y$服从正态分布$N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)$.

正态分布的可加性

定理18

设随机变量(X, Y)服从二维正态分布$N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 则随机变量$Z = X + Y$服从正态分布$N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$.

证明: (1)记$X^* = \dfrac{X - \mu_1}{\sigma_1}$, $Y^* = \dfrac{Y - \mu_2}{\sigma_2}$, 则$(X^*, Y^*)$服从二维正态分布$N(0, 1; 0, 1; \rho)$.

正态分布的可加性

定理18

设随机变量(X,Y)服从二维正态分布$N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$, 则随机变量$Z=X+Y$服从正态分布$N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)$.

证明: (1)记$X^*=\dfrac{X-\mu_1}{\sigma_1}$, $Y^*=\dfrac{Y-\mu_2}{\sigma_2}$, 则$(X^*,Y^*)$服从二维正态分布$N(0,1;0,1;\rho)$.

(2)若(X^*,Y^*)服从二维正态分布$N(0,1;0,1;\rho)$, 则对任意的$t\in\mathbb{R}$, tX^*+Y^*服从正态分布$N(0,1+t^2+2t\rho)$.

正态分布的可加性

定理18

设随机变量(X,Y)服从二维正态分布$N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$, 则随机变量$Z=X+Y$服从正态分布$N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)$.

证明: (1)记$X^*=\dfrac{X-\mu_1}{\sigma_1}$, $Y^*=\dfrac{Y-\mu_2}{\sigma_2}$, 则$(X^*,Y^*)$服从二维正态分布$N(0,1;0,1;\rho)$.

(2)若(X^*,Y^*)服从二维正态分布$N(0,1;0,1;\rho)$, 则对任意的$t\in\mathbb{R}$, tX^*+Y^*服从正态分布$N(0,1+t^2+2t\rho)$.

(3)$X+Y=\sigma_2\left(\dfrac{\sigma_1}{\sigma_2}X^*+Y^*\right)+\mu_1+\mu_2$,正态分布的线性不变性.

推论19

设X服从正态分布$N(\mu_1, \sigma_1^2)$, Y 服从正态分布$N(\mu_2, \sigma_2^2)$, 且相互独立, 则$X \pm Y$服从正态分布$N(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$.

推论19

设X服从正态分布$N(\mu_1,\sigma_1^2)$, Y 服从正态分布$N(\mu_2,\sigma_2^2)$, 且相互独立, 则$X\pm Y$服从正态分布$N(\mu_1\pm\mu_2,\sigma_1^2+\sigma_2^2)$.

由正态分布的线性不变性和可加性, 我们可以得到

推论20

任意n个相互独立的正态随机变量的线性组合仍是正态随机变量. 即若X_i服从正态分布$N(\mu_i,\sigma_i^2)$, $i=1,2,\ldots,n$, 且相互独立, 若a_i是任意常数, $i=1,2,\ldots,n$, 则随机变量$\sum_{i=1}^{n}a_iX_i$ 服从正态分布$N(a_1\mu_1+\cdots+a_n\mu_n,a_1^2\sigma_1^2+\cdots+a_n^2\sigma_n^2)$.

定理21

设随机变量$X_1, X_2, \ldots, X_n$相互独立, 分布函数分别为$F_i(x)$, $i = 1, \ldots, n$. 记

$$Y = \max(X_1, X_2, \ldots, X_n), \qquad Z = \min(X_1, X_2, \ldots, X_n),$$

则Y和Z的分布函数分别为

$$F_Y(y) = \prod_{i=1}^{n} F_i(y), \qquad F_Z(z) = 1 - \prod_{i=1}^{n}(1 - F_i(z)).$$

证明: 由分布函数的定义和独立性, Y和Z的分布函数分别为

$$
\begin{aligned}
F_Y(y) =& P(Y \le y) = P(\max(X_1, X_2, \ldots, X_n) \le y) \\
=& P(X_1 \le y, X_2 \le y, \ldots, X_n \le y) \\
=& P(X_1 \le y)P(X_2 \le y)\ldots P(X_n \le y) = \prod_{i=1}^{n} F_i(y)
\end{aligned}
$$

$$
\begin{aligned}
F_Z(z) =& P(Z \le z) = 1 - P(Z > z) = 1 - P(\min(X_1, X_2, \ldots, X_n) > z) \\
=& 1 - P(X_1 > z, X_2 > z, \ldots, X_n > z) = 1 - \prod_{i=1}^{n} P(X_i > z) \\
=& 1 - \prod_{i=1}^{n}(1 - P(X_i \le z)) = 1 - \prod_{i=1}^{n}(1 - F_i(z)).
\end{aligned}
$$

推论22

设随机变量$X_1, X_2, \ldots, X_n$相互独立且同分布(简记为$i.i.d.$), 共同的分布函数为$F_X(x)$. 记

$$Y = \max(X_1, X_2, \ldots, X_n), \qquad Z = \min(X_1, X_2, \ldots, X_n),$$

则Y和Z的分布函数分别为

$$F_Y(y) = [F_X(y)]^n, \qquad F_Z(z) = 1 - [1 - F_X(z)]^n.$$

进一步地, 若$X_1, X_2, \ldots, X_n$是连续型随机变量, 具有概率密度函数$p_X(x)$, 则Y和Z的概率密度函数分别为

$$p_Y(y) = n\,[F_X(y)]^{n-1}\,p_X(y), \qquad p_Z(z) = n\,[1 - F_X(z)]^{n-1}\,p_X(z).$$

例10

在区间$(0,1)$上随机地取n个点, X和Y分别表示最右端和最左端两个点的坐标, 求X与Y的概率密度函数.

解: 设n个点的坐标分别为$X_1,\ldots,X_n$, 则由已知, $X_1,\ldots,X_n$独立同分布, 共同分布为均匀分布$U(0,1)$, $X=\max\{X_1,\ldots,X_n\}$, $Y=\min\{X_1,\ldots,X_n\}$.

已知均匀分布$U(0,1)$的概率密度函数和分布函数分别为

$$p(t)=\begin{cases}1, & 0<t<1,\\ 0, & \text{其他}.\end{cases}\qquad F(t)=\begin{cases}0, & t<0,\\ t, & 0\le t<1,\\ 1, & t\ge 1.\end{cases}$$

故由推论知, X与Y的概率密度函数分别为

$$p_X(x) = n(F(x))^{n-1}p(x) = \begin{cases} nx^{n-1}, & 0 < x < 1, \\ 0, & \text{其他}. \end{cases}$$

和

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

随机向量变换的分布

目的: 已知随机向量$\boldsymbol{X} = (X_1, \cdots, X_n)$的联合分布, 求

$$\boldsymbol{Y} = (h_1(X_1, \cdots, X_n), \cdots, h_m(X_1, \cdots, X_n))$$

的联合分布.

仅考虑连续型随机向量的变换的分布.

已知随机向量$\boldsymbol{X} = (X_1, \cdots, X_n)$的概率密度函数$p(x_1, \cdots, x_n)$, $Y_i = h_i(\boldsymbol{X})$, $i = 1, \cdots, m$, 我们要求$\boldsymbol{Y} = (Y_1, \cdots, Y_m)$的分布, 这里诸$h_i$是定义在$\mathbb{R}^n$ 到$\mathbb{R}$ 上的可测函数.

同一维情形类似, 求$\boldsymbol{Y}$的分布函数归结为计算多重积分, 即$\boldsymbol{Y}$的分布函数为

$$G(y_1, \cdots, y_m) = \int \cdots \int_D p(x_1, \cdots, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n,$$

其中$D = \{(x_1, \cdots, x_n) : h_i(x_1, \cdots, x_n) \leq y_i, i = 1, \cdots, m\}$.

当$\boldsymbol{X}=(X_1,\cdots,X_n)$和$\boldsymbol{Y}=(Y_1,\cdots,Y_m)$有一一对应的变换关系时(此时必有$n=m$), 可以直接求出$\boldsymbol{Y}$的概率密度函数.

定理23

设随机向量$\boldsymbol{X}=(X_1,\cdots,X_n)$有概率密度函数$p(x_1,\cdots,x_n)$, $y_i=h_i(x_1,\cdots,x_n)$, $i=1,\cdots,n$是$\mathbb{R}^n$ 到$\mathbb{R}$ 上的可测函数. 若对每个$i=1,\cdots,n$, $y_i=h_i(x_1,\cdots,x_n)$的逆变换$x_i=k_i(y_1,\cdots,y_n)$存在且连续, 并有连续的偏导数, 则随机向量$\boldsymbol{Y}=(Y_1,\cdots,Y_m)$是连续型随机向量, 其概率密度函数为

$$p^*(y_1,\cdots,y_n)=\begin{cases}p(k_1(y_1,\cdots,y_n),\cdots,k_n(y_1,\cdots,y_n))|J|, & (y_1,\cdots,y_n)\\ 0, & \text{其他情形}.\end{cases}$$

这里E是$(h_1(x_1,\cdots,x_n),\cdots,h_n(x_1,\cdots,x_n))$的值域, J为坐标变换的雅可比行列式, 即

$$J=\frac{\partial(x_1,\cdots,x_n)}{\partial(y_1,\cdots,y_n)}=\begin{vmatrix}\frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n}\\ \vdots & \vdots & \vdots\\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n}\end{vmatrix}$$

定理24

设随机变量(X, Y)服从二维正态分布

$N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 记$X^* = \dfrac{X - \mu_1}{\sigma_1}$, $Y^* = \dfrac{Y - \mu_2}{\sigma_2}$,

则(X^*, Y^*)服从二维正态分布$N(0, 1; 0, 1; \rho)$.

(X, Y)的联合概率密度函数为

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2 c} \exp\left[-\frac{1}{2c^2}(a^2 + b^2 - 2\rho ab)\right]$$

其中$a = \dfrac{x - \mu_1}{\sigma_1}, b = \dfrac{y - \mu_2}{\sigma_2}, c = \sqrt{1 - \rho^2}.$

(X,Y)的联合概率密度函数为

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2 c}\exp\left[-\frac{1}{2c^2}(a^2+b^2-2\rho ab)\right]$$

其中$a=\dfrac{x-\mu_1}{\sigma_1}, b=\dfrac{y-\mu_2}{\sigma_2}, c=\sqrt{1-\rho^2}$.

作变换$\begin{cases} s=\dfrac{x-\mu_1}{\sigma_1}, \\ t=\dfrac{y-\mu_2}{\sigma_2} \end{cases}$，则逆变换为$\begin{cases} x=\sigma_1 s+\mu_1, \\ y=\sigma_2 t+\mu_2. \end{cases}$，雅可比行列式

$$J=\begin{vmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial x}{\partial t} \\ \dfrac{\partial y}{\partial s} & \dfrac{\partial y}{\partial t} \end{vmatrix}=\begin{vmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{vmatrix}=\sigma_1\sigma_2.$$

于是, (X^*, Y^*)的联合概率密度函数为

$$
\begin{aligned}
p^*(s,t) =& \sigma_1\sigma_2 p(\sigma_1 s+\mu_1, \sigma_2 t+\mu_2) \\
=& \frac{1}{2\pi c}\exp\left[-\frac{1}{2c^2}(s^2+t^2-2\rho st)\right].
\end{aligned}
$$

故(X^*, Y^*)服从二维正态分布$N(0,1;0,1;\rho)$.

例11

设X与Y独立同分布, 共同分布为指数分布$Exp(1)$, 记$U = X + Y, V = X/Y$, 求$U = X + Y$的概率密度函数和随机变量(U, V) 的联合概率密度函数.

解: 易知(X,Y)的联合概率密度函数为

$$p(x,y)=\begin{cases} e^{-(x+y)}, & x>0, y>0; \\ 0, & \text{其他情形}. \end{cases}$$

作变换$\begin{cases} u=x+y, \\ v=x/y. \end{cases}$ 当$x>0, y>0$时, 逆变换为$\begin{cases} x=\dfrac{uv}{1+v}, \\ y=\dfrac{u}{1+v}. \end{cases}$

此时$u>0, v>0$. 计算雅可比行列式

$$J=\begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}=\begin{vmatrix} \dfrac{v}{1+v} & \dfrac{u}{(1+v)^2} \\ \dfrac{1}{1+v} & -\dfrac{u}{(1+v)^2} \end{vmatrix}=-\frac{u}{(1+v)^2}.$$

于是, (U,V)的联合概率密度函数为

$$p^*(u,v)=\begin{cases} e^{-u}\dfrac{u}{(1+v)^2}, & u>0, v>0; \\ 0, & \text{其他情形}. \end{cases}$$

U的概率密度函数

$$p_U^*(u)=\int p^*(u,v)\mathrm{d}v=\begin{cases} \displaystyle\int_0^\infty e^{-u}\frac{u}{(1+v)^2}\mathrm{d}v, & u>0; \\ 0, & u\le 0. \end{cases}$$

$$=\begin{cases} ue^{-u}, & u>0; \\ 0, & u\le 0. \end{cases}$$

注记5

为求$U = g(X, Y)$的分布, 除了直接使用求随机向量函数的分布的方法之外, 可以先增补一个新的变量$V = h(X, Y)$(通常令$V = X$或Y), 先求(U, V) 的联合分布再求$U = X + Y$ 的边际分布. 这种方法称之为**增补变量法**.