目　录

第二章　随机变量及其分布

§2.7　随机变量的函数的概率分布

最简单的情形，是由一维随机变量X的概率分布去求其一给定函数$Y=g(X)$的分布。较常见的，是由$(X_1,X_2,\cdots,X_n)$的分布去求$Y=g(X_1,X_2,\cdots,X_n)$的分布。更一般地，由$(X_1,X_2,\cdots,X_n)$的分布去求$(Y_1,Y_2,\cdots,Y_m)$的分布，其中$Y_i=g_i(X_1,X_2,\cdots,X_n)$，$i=1,2,\cdots,m$。

这一部分内容，与数理统计中求统计量的分布有密切的联系。

1. 离散型随机变量的情形

设X的分布律为

$$P(X=x_i)=p_i,\quad i=1,2,\cdots$$

$g:R\to R$，令$Y=g(X)$，则Y的分布律为

$$P(Y=y_j)=P(g(X)=y_j)=\sum_{x_i:g(x_i)=y_j}P(X=x_i)=\sum_{i:g(x_i)=y_j}p_i$$

例 2.7.1. *设X的概率函数为*

X	-1	0	1	2
P	1/4	1/2	1/8	1/8

试求$Y=X^2$，$Z=X^3+1$的分布律。

解: 容易求得Y的分布律为:

Y	0	1	4
P	1/2	3/8	1/8

Z的分布律

Z	0	1	2	9
P	1/4	1/2	1/8	1/8

上述结论可以推广到多维随机变量的情形:

设随机向量X的分布律为$P(X=x)$，则X的函数$Y=g(X)$的分布律为

$$P(Y=y)=P(g(X)=y)=\sum_{x:g(x)=y}P(X=x)$$

特别当ξ,η是相互独立的非负整值随机变量，各有分布律$\{a_k\}$与$\{b_k\}$. 那么$\xi+\eta$有分布律

$$P(\xi+\eta=n)=\sum_{k=0}^{n}a_kb_{n-k}$$

称此公式为**离散卷积公式**

例 2.7.2. 设$X\sim B(n,p)$，$Y\sim B(m,p)$且X和Y相互独立，则$X+Y\sim B(n+m,p)$。这种性质称为再生性。可推广至多项和：设$X_i\sim B(n_i,p),(i=1,2,\cdots,m)$，且$X_1,X_2,\cdots,X_m$独立，则有:$\sum\limits_{i=1}^{m}X_i\sim B(\sum\limits_{i=1}^{m}n_i,p)$。特别，若$X_1,X_2,\cdots,X_n$为独立同分布，且$X_i\sim B(1,p),\ \ i=1,\cdots,n.$ 则有:$\sum\limits_{i=1}^{n}X_i\sim B(n,p)$。此结论揭示了二项分布与$0-1$分布之间的密切关系。

例 2.7.3. 设$X\sim P(\lambda)$，$Y\sim P(\mu)$，且X和Y独立，则有$X+Y\sim P(\lambda+\mu)$。即$Poisson$分布亦具有再生性，且可推广。

2. 连续型随机变量的情形

定理 2.7.1. *[密度变换公式] 设随机变量X有概率密度函数 $f(x)$, $x\in(a,b)(a,b$可以为$\infty)$, 而$y=g(x)$在$x\in(a,b)$上是严格单调的连续函数，存在唯一的反函数$x=h(y),y\in(\alpha,\beta)$并且$h'(y)$存在且连续，那么$Y=g(X)$也是连续型随机变量且有概率密度函数*

$$p(y)=f(h(y))|h'(y)|,\quad y\in(\alpha,\beta).$$

例 2.7.4. 设随机变量$X\sim U(-\frac{\pi}{2},\frac{\pi}{2})$, 求$Y=tgX$的概率密度函数。

由密度变换公式知Y的概率密度函数为

$$f(y)=\frac{1}{\pi}arctg'(y)=\frac{1}{\pi(1+y^2)},\quad -\infty<y<\infty$$

此分布称为Cauchy分布。本题我们也可以用一般的方法求解，即先求出分布函数，然后对分布函数求导数得到。

$$\begin{aligned} F(y) &= P(Y \le y) = P(tg(X) \le y) \\ &= P(X \le arctg(y)) = \int_{-\frac{\pi}{2}}^{arctg(y)} \frac{1}{\pi} dy = \frac{1}{\pi} arctg(y) + \frac{1}{2}. \end{aligned}$$

所以Y的概率密度为

$$f(y) = F'(y) = \frac{1}{\pi(1+y^2)}.$$

这种方法更具有一般性。

注: 当g不是在全区间上单调而是逐段单调时，密度变换公式为下面的形式:
设随机变量ξ的密度函数为$p_\xi(x)$, $a < x < b$. 如果可以把(a,b)分割为一些(有限个或可列个)互不重叠的子区间的和$(a,b) = \bigcup_j I_j$, 使得函数$u = g(t)$, $t \in (a,b)$在每个子区间上有唯一的反函数$h_j(u)$, 并且$h_j^{'}(u)$存在连续, 则$\eta = g(\xi)$是连续型随机变量, 其密度函数为:

$$p_\eta(x) = \sum_j p_\xi(h_j(x))|h_j^{'}(x)| \ . \tag{2.7.1}$$

例 2.7.5. 设$X \sim N(0,1)$，求$Y = X^2$的概率密度。

解: 由于函数$y = x^2$在$(-\infty, 0)$和$[0,\infty)$上严格单调，因此由上述定理知Y的概率密度为

$$\begin{aligned} f(y) &= \phi(-\sqrt{y})| - \sqrt{y}'|I_{\{y>0\}} + \phi(\sqrt{y})|\sqrt{y}'|I_{\{y>0\}} \\ &= \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}} I_{\{y>0\}} \end{aligned}$$

定理 2.7.2. *设(ξ_1, ξ_2)是2维连续型随机向量, 具有联合密度函数$p(x_1, x_2)$, 设$\zeta_j = f_j(\xi_1, \xi_2), j = 1, 2$. 若$(\xi_1, \xi_2)$与$(\zeta_1, \zeta_2)$一一对应, 逆映射$\xi_j = h_j(\zeta_1, \zeta_2), j = 1, 2$.假定每个$h_j(y_1, y_2)$都有一阶连续偏导数. 则$(\zeta_1, \zeta_2)$亦为连续型随机向量, 且其联合概率密度为*

$$q(y_1, y_2) = \begin{cases} p\left(h_1(y_1, y_2), h_n(y_1, y_2)\right)|J|, & (y_1, y_2) \in \mathbb{D}, \\ 0, & (y_1, y_2) \notin \mathbb{D}, \end{cases} \tag{2.7.2}$$

其中$\mathbb{D}$是随机向量(ζ_1, ζ_2)的所有可能值的集合, J是变换的Jaccobi行列式, 即

$$J = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix}$$

在多元随机变量场合，更一般地有

定理 2.7.3. 如果$(\xi_1,\cdots,\xi_n)$是n维连续型随机向量, 具有联合密度函数$p(x_1,\cdots,x_n)$. 假设存在n个n元函数

$$y_j = f_j(x_1,\cdots,x_n), \quad j=1,\cdots,n,$$

使得

$$\zeta_j = f_j(\xi_1,\cdots,\xi_n), \quad j=1,\cdots,n,$$

若$(\xi_1,\cdots,\xi_n)$与$(\zeta_1,\cdots,\zeta_n)$之间一一对应, 逆映射为$\xi_j = h_j(\zeta_1,\cdots,\zeta_n)$, $j=1,\cdots,n$. 其中每个$h_j(y_1,\cdots,y_n)$都有一阶连续偏导数, 那么随机向量$(\zeta_1,\cdots,\zeta_n)$是连续型的, 且具有联合密度函数

$$q(y_1,\cdots,y_n) = \begin{cases} p\left(h_1(y_1,\cdots,y_n),\cdots,h_n(y_1,\cdots,y_n)\right)|J|, & (y_1,\cdots,y_n)\in\mathbb{D}, \\ 0, & (y_1,\cdots,y_n)\notin\mathbb{D}, \end{cases} \tag{2.7.3}$$

其中$\mathbb{D}$是随机向量$(\zeta_1,\cdots,\zeta_n)$的所有可能值的集合, J是变换的$Jaccobi$行列式，即

$$J = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \cdots & \frac{\partial h_n}{\partial y_n} \end{vmatrix}$$

例 2.7.6. 在直角坐标平面上随机选取一点, 分别以随机变量ξ与η表示其横坐标和纵坐标, 可以认为ξ与η相互独立. 如果ξ与η都服从正态分布$N(0,1)$, 试求其极坐标(ρ,θ)的分布.

解: 易知

$$\begin{cases} x = r\cos t \\ y = r\sin t \end{cases}$$

是$(0,\infty)\times[0,2\pi)$与$\mathbb{R}^2$(原点除外)之间的一一变换, 变换的Jaccobi行列式

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial t} \end{vmatrix} = \begin{vmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{vmatrix} = r.$$

由于(ξ,η)的联合密度为

$$p(x,y) = \frac{1}{2\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\},$$

所以由(2.7.3)式得知, (ρ,θ)的联合密度为

$$q(r,t)=\frac{1}{2\pi}r\exp\left\{-\frac{r^2}{2}\right\}=q_1(r)q_2(t),\quad r>0,\quad t\in[0,2\pi). \tag{2.7.4}$$

其中$q_1(r)=r\exp\left\{-\frac{r^2}{2}\right\},\quad r>0;\quad q_2(t)=\frac{1}{2\pi},t\in[0,2\pi)$.

这一结果表明: θ与ρ相互独立, 其中θ服从$[0,2\pi)$ 上的均匀分布; 而ρ则服从Weibull分布(参数$\lambda=1/2,\alpha=2$).

在计算两个随机变量之和时，我们还经常用到如下定理

定理 2.7.4. 设X,Y的联合概率密度为$f(x,y)$，则$X+Y$的概率密度$p(z)$为

$$p(z)=\int_{-\infty}^{\infty}f(x,z-x)dx=\int_{-\infty}^{\infty}f(z-y,y)dy$$

证一: 先求$X+Y$的分布函数$F(z)$. 我们有

$$\begin{aligned}F(z)=P(X+Y\le z)&=\int\int_{x+y\le z}f(x,y)dxdy=\int_{-\infty}^{\infty}dx\int_{-\infty}^{z-x}f(x,y)dy\\&=\int_{-\infty}^{\infty}du\int_{-\infty}^{z}f(x,t-x)dt=\int_{-\infty}^{z}\left\{\int_{-\infty}^{\infty}f(x,t-x)dx\right\}dt.\end{aligned}$$

这就说明,$X+Y$的分布函数$F(z)$是其中的花括弧中的函数在区间$(-\infty,z)$上的积分,所以$X+Y$是连续型随机变量,其密度函数如定理所述。

证二: 令$X=Z_1,X+Y=Z_2$, 利用单调映射的密度变换公式(2.7.2)可求得(Z_1,Z_2)的联合概率密度函数为$g(z_1,z_2)=f(z_1,z_2-z_1)$. 再对$g(z_1,z_2)$关于$z_1$在$R$上积分, 便求得$Z_2=X+Y$ 的密度为

$$\int_{-\infty}^{\infty}g(z_1,z_2)dz_1=\int_{-\infty}^{\infty}f(z_1,z_2-z_1)dz_1,$$

故得所证.

特别，当X与Y独立时，分别记X和Y的概率密度为$f_1(x)$和$f_2(y)$，则$X+Y$的概率密度为

$$p(z)=\int_{-\infty}^{\infty}f_1(x)f_2(z-x)dx=\int_{-\infty}^{\infty}f_1(z-y)f_2(y)dy\quad\triangleq\quad f_1*f_2(z)=f_2*f_1(z)$$

称此公式为**卷积公式**。

例 2.7.7. 设X服从期望为2的指数分布，$Y\sim U(0,1)$，且X和Y相互独立。求$X-Y$的概率密度和$P(X\le Y)$。

解一: 由题设知$-Y\sim U(-1,0)$，并记X和$-Y$的密度分别为f_1和f_2，从而由卷积公式有

$$\begin{aligned} f_{X-Y}(z) &= \int_{-\infty}^{\infty} f_1(x)f_2(z-x)dx \\ &= \begin{cases} e^{-\frac{z}{2}}(1-e^{-\frac{1}{2}}), & z\geq 0 \\ 1-e^{-\frac{z+1}{2}}, & -1<z<0 \\ 0, & z\leq -1 \end{cases} \end{aligned}$$

所以$P(X\leq Y)=P(X-Y\leq 0)=2e^{-\frac{1}{2}}-1$。

解二: 由于

$$\begin{aligned} P(X-Y\leq z) &= \int P(X\leq z+y|Y=y)f(y)dy \\ &= \begin{cases} \int_0^1 P(X\leq z+y)dy & z\geq 0 \\ \int_{-z}^1 P(X\leq z+y)dy & -1<z<0 \\ 0 & z\leq -1 \end{cases} \\ &= \begin{cases} 1-2e^{-z/2}(1-e^{-1/2}), & z\geq 0 \\ z+2e^{-(z+1)/2}-1, & -1<z<0 \\ 0, & z\leq -1 \end{cases} \end{aligned}$$

再对分布函数求导数即得所求.

一些连续型随机变量，也有再生性性质。

例 2.7.8. 设$X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$且$X$与$Y$相互独立，则*:*

$$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2).$$

更一般地, 设$X_i\sim N(\mu_i,\sigma_i^2), i=1,\cdots,n$, $X_1,\cdots,X_n$相互独立. $a_1,\cdots,a_n,b$为任意$n+1$个实数, 其中$a_1,\cdots,a_n$不全为零. 令$X=\sum\limits_{i=1}^n a_iX_i+b$, 则有*:* $X\sim N(\mu,\sigma^2)$, 其中$\mu=\sum\limits_{i=1}^n a_i\mu_i+b$, $\sigma^2=\sum\limits_{i=1}^n a_i^2\sigma_i^2$.

我们把具有再生性性质的分布总结一下为

- 二项分布(关于试验次数具有再生性)
- $Poisson$分布(关于参数λ具有再生性)

- $Pascal$分布(关于成功次数r具有再生性)
- 正态分布(关于两个参数都具有再生性)
- 具有再生性的连续型分布还有χ^2分布和Γ分布

有时我们还会碰到计算随机变量之商的概率密度. 我们有

定理 2.7.5. 如果(ξ,η)是二维连续型随机向量, 它们的联合密度为$f(x,y)$, 则它们的商ξ/η是连续型随机变量, 具有密度函数

$$\begin{aligned} p_{\frac{\xi}{\eta}}(x) &= \int_{-\infty}^{\infty} |t| f(xt,t)dt, \quad \forall\ x\in\mathbb{R}. \\ \text{而}\ p_{\frac{\eta}{\xi}}(x) &= \int_{-\infty}^{\infty} |u| f(u,xu)du, \quad \forall\ x\in\mathbb{R}. \end{aligned} \tag{2.7.5}$$

例 2.7.9. 设随机变量ξ与η相互独立,同服从参数$\lambda=1$的指数分布, 试求ξ/η的密度函数.

解: 我们利用(2.7.5)式求$p_{\frac{\xi}{\eta}}(x)$.由于(ξ,η) 的联合密度为

$$p(u,v)=e^{-u-v}, \quad u>0,\ v>0,$$

所以欲(2.7.5)式中的被积函数$|t|p(xt,t)\neq 0$, 当且仅当, $t>0$和$xt>0$, 从而知有

$$p_{\frac{\xi}{\eta}}(x)=\begin{cases} \int_0^\infty t\ e^{-xt-t}dt=\frac{1}{(1+x)^2} & x>0; \\ 0 & x\leq 0. \end{cases}$$

易见$p_{\frac{\eta}{\xi}}(x)$同上。

例 2.7.10. 设$X_1\sim N(0,1)$, $X_2\sim\chi_n^2$, 且X_1与X_2独立。命$Y=\frac{X_1}{\sqrt{X_2/n}}$, 则: $Y\sim t_n$(自由度为n的t分布)。

例 2.7.11. 设$X_1\sim\chi_m^2$, $X_2\sim\chi_n^2$, 且X_1与X_2独立, 则$Y=\frac{X_1/m}{X_2/n}\sim F_{m,n}$(自由度为$m,n$的$F$分布)。

例 2.7.12. 极小值和极大值的分布

对于n个随机变量$\xi_1, ..., \xi_n$,我们可以考察它们的最大值和最小值:

$$\eta_1 = \max\{\xi_1, ..., \xi_n\},$$

$$\eta_2 = \min\{\xi_1, ..., \xi_n\}.$$

如此定义的η_1与η_2也是随机变量.

当$\xi_1, ..., \xi_n$相互独立时, 我们不难利用它们的分布函数$F_1(x), ..., F_n(x)$求出η_1与η_2 的分布函数$F_{\eta_1}(x)$和$F_{\eta_2}(x)$.

事实上,

$$\begin{aligned} F_{\eta_1}(x) &= P(\eta_1 \le x) = P(\max\{\xi_1, \cdots, \xi_n\} \le x) = P\left(\bigcap_{k=1}^{n}(\xi_k \le x)\right) \\ &= \prod_{k=1}^{n} P(\xi_k \le x) = \prod_{k=1}^{n} F_k(x); \end{aligned} \tag{2.7.6}$$

而利用关系式

$$(\eta_2 > x) = (\xi_1 > x,\ \ldots, \xi_n > x) = \bigcap_{k=1}^{n}(\xi_k > x)$$

可得

$$\begin{aligned} F_{\eta_2}(x) &= P(\eta_2 \le x) = 1 - P(\eta_2 > x) = 1 - P\left(\bigcap_{k=1}^{n}(\xi_k > x)\right) \\ &= 1 - \prod_{k=1}^{n} P(\xi_k > x) = 1 - \prod_{k=1}^{n}(1 - F_k(x)). \end{aligned} \tag{2.7.7}$$

目前我们接触到的分布的关系为

- n个独立同分布$B(1,p)$的0-1分布随机变量之和为二项分布$B(n,p)$;
- 有限个独立二项随机变量(成功的概率相同)之和仍为二项分布;
- 有限个独立的$Poisson$分布随机变量之和服从$Poisson$分布，参数相加;
- r个独立同分布几何分布$G(p)$的随机变量之和服从参数为r和p的$Pascal$分布;
- 任意有限个独立的正态分布随机变量的线性组合仍然服从正态分布;