2.5: 随机变量的函数的分布

张伟平

课件 http://staff.ustc.edu.cn/~zwp/

论坛 http://fisher.stat.ustc.edu.cn

第二章随机变量及其分布

2.5 随机变量的函数的概率分布 1
2.5.1 离散型随机变量的情形 1
2.5.2 连续型随机变量的情形 5
2.5.3 极小值和极大值的分布 28

2.5 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。

2.5.1 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

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

↑Example

设 X 的概率函数为

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

试求 $Y=X^2$, $\quad Z=X^3+1$ 的分布律。

↓Example

解:

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

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

称此公式为**离散卷积公式**

设 $X \sim B(n,p)$，$Y \sim B(m,p)$ 且 X 和 Y 相互独立，则 $X+Y \sim B(n+m,p)$。

↑Example ↓Example

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。

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。

↑Example ↓Example

2.5.2 连续型随机变量的情形

定理 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 ∞), 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

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

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

↑Example
↓Example

解:

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:

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

设 $X \sim N(0,1)$，求 $Y = X^2$ 的概率密度。

↑Example

↓Example

解:

定理 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 (ξ_1,ξ_2) 与 (ζ_1,ζ_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 (ζ_1,ζ_2) 亦为连续型随机向量, 且其联合概率密度为

$$q(y_1,y_2)=\begin{cases} p\left(h_1(y_1,y_2),h_n(y_1,y_2)\right)|J|, & (y_1,y_2)\in\mathbb{D}, \\ 0, & (y_1,y_2)\notin\mathbb{D}, \end{cases} \tag{2.1}$$

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 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

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

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

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

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.

↑Example

↓Example

解:

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 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:

证二: 2.1

↑Example

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 $P(X \le Y)$。

↓Example

解一:

解二:

设 $X_1, \ldots, X_n i.i.d \sim N(0,1)$, 试求 $Y_n = \sum_{i=1}^n X_i^2$ 的分布.

解:

由归纳法得 Y_n 的密度函数. 称 Y_n 的分布为自由度 n 的卡方分布, 记为 $Y_n \sim \chi^2_n$.

- χ^2_n 具有再生性
- $X \sim \chi^2_n$, 则 $EX = n, Var(X) = 2n$.
- http://en.wikipedia.org/wiki/Chi-squared_distribution

一些连续型随机变量，也有再生性性质。

↑Example

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 $X = \sum\limits_{i=1}^{n} a_i X_i + b$, 则有: $X \sim N(\mu, \sigma^2)$, 其中 $\mu = \sum\limits_{i=1}^{n} a_i \mu_i + b$, $\sigma^2 = \sum\limits_{i=1}^{n} a_i^2 \sigma_i^2$.

↓Example

↑Example

设 $X_1 \sim \chi_n^2, X_2 \sim \chi_m^2$, 且 X_1 和 X_2 相互独立, 则 $X_1 + X_2 \sim \chi_{n+m}^2$.

↓Example

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)
- χ^2 分布具有再生性

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 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

$$\begin{aligned} p_{\frac{\xi}{\eta}}(x) &= \int_{-\infty}^{\infty} |t| f(xt,t)dt, \quad \forall\ x \in \mathbb{R}. \\ \text{而 } p_{\frac{\eta}{\xi}}(x) &= \int_{-\infty}^{\infty} |u| f(u,xu)du, \quad \forall\ x \in \mathbb{R}. \end{aligned} \tag{2.4}$$

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.

↑Example ↓Example

解:

设 $X_1 \sim N(0,1)$，$X_2 \sim \chi^2_n$，且 X_1 与 X_2 独立。求 $Y = \frac{X_1}{\sqrt{X_2/n}}$ 的概率密度函数. (记 $Y \sim t_n$, 称为自由度为 n 的 t 分布)。

$\overline{\uparrow}$Example

$\underline{\downarrow}$Example

解:

- t_n 的分布关于原点对称
- $lim_{n \to \infty} f_Y(y) = \phi(y)$.
- http://en.wikipedia.org/wiki/Student's_t-distribution

↑Example

设 $X_1 \sim \chi_n^2$，$X_2 \sim \chi_m^2$，且 X_1 与 X_2 独立，求 $Y = \frac{X_1/n}{X_2/m}$ 的概率密度函数. (记 $Y \sim F_{n,m}$, 称为自由度为 n, m 的 F 分布)。

↓Example

解:

- http://en.wikipedia.org/wiki/F-distribution

2.5.3 极小值和极大值的分布

对于 n 个随机变量 $X_1, ..., X_n$, 我们可以考察它们的最大值和最小值:

$$X_{(n)} = \max\{X_1, ..., X_n\},$$

$$X_{(1)} = \min\{X_1, ..., X_n\}.$$

如此定义的 $X_{(n)}$ 与 $X_{(1)}$ 也是随机变量.

当 $X_1, ..., X_n$ 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 $X_{(n)}$ 与 $X_{(1)}$ 的分布函数 $F_{X_{(n)}}(x)$ 和 $F_{X_{(1)}}(x)$.

事实上,

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) = P(\max\{X_1, \cdots, X_n\} \leq x) = P\left(\bigcap_{k=1}^{n}(X_k \leq x)\right) \\ &= \prod_{k=1}^{n} P(X_k \leq x) = \prod_{k=1}^{n} F_k(x); \end{aligned} \tag{2.5}$$

而利用关系式

$$(X_{(1)} > x) = (X_1 > x,\ \ldots, X_n > x) = \bigcap_{k=1}^{n} (X_k > x)$$

可得

$$\begin{aligned} F_{X_{(1)}}(x) &= P(X_{(1)} \le x) = 1 - P(X_{(1)} > x) = 1 - P\left(\bigcap_{k=1}^{n} (X_k > x)\right) \\ &= 1 - \prod_{k=1}^{n} P(X_k > x) = 1 - \prod_{k=1}^{n} (1 - F_k(x)). \end{aligned} \tag{2.6}$$

在 $X_1, \ldots, X_n$ 还是同分布时候, (2.5) 和 (2.6) 还可以简化.

设 $X_1, \ldots, X_n i.i.d \sim U(0,\theta), \theta > 0$, 求 $X_{(n)} = max_{1\le i\le n} X_i$ 的密度函数.

↑Example

↓Example

解:

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;