概率论:Week 11

1 分布

性质 (测度的性质)**.** 设 $(\Omega, \mathcal{F}, \nu)$ 是一个测度空间.

1. *(单调性). 如果 $A \subset B$, 那么 $\nu(A) \leqslant \nu(B)$.*
2. *(次可加性). 对于任意序列 $A_1, A_2, \ldots$,*
$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} \nu\left(A_i\right).$$
3. *(连续性). 如果 $A_1 \subset A_2 \subset A_3 \subset \cdots$ (或者 $A_1 \supset A_2 \supset A_3 \supset \cdots$ 且 $\nu\left(A_1\right) < \infty$), 那么*
$$\nu\left(\lim_{n \to \infty} A_n\right) = \lim_{n \to \infty} \nu\left(A_n\right),$$
   其中
$$\lim_{n \to \infty} A_n = \bigcup_{i=1}^{\infty} A_i \quad \left(\text{或 } = \bigcap_{i=1}^{\infty} A_i\right).$$

设P是一个概率测度, P的累积分布函数定义为:

$$F(x) = P((-\infty, x]), \quad x \in \mathbb{R}$$

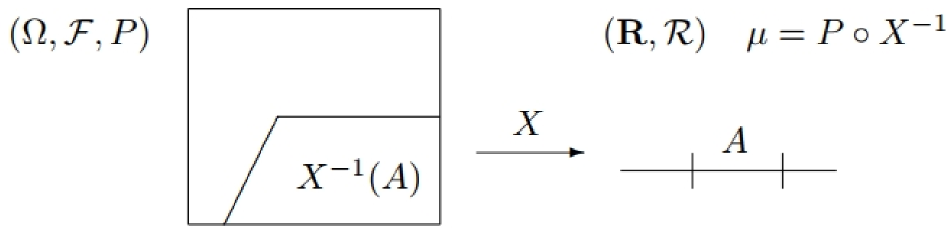
1. 设 F 是 $\mathbb{R}$ 上的一个 c.d.f., 那么
   (a) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$;
   (b) $F(\infty) = \lim_{x \to \infty} F(x) = 1$;
   (c) F 是非降的, 即若 $x \leqslant y$ 则 $F(x) \leqslant F(y)$;
   (d) F 是右连续的, 即 $\lim_{y \to x, y > x} F(y) = F(x)$.
2. 假设 $\mathbb{R}$ 上的实值函数 F 满足上述的四个条件, 那么 F 是 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上唯一的概率测度的 c.d.f.

注记 (对应关系)**.** 概率的基本性质和分布函数的基本性质对应.

定义 (随机变量)**.** 从可测空间$(\Omega, \mathscr{F})$到可测空间$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$上的可测映射$X$被称为随机变量. 即$X^{-1}(B) \in \mathscr{F}, \forall B \in \mathcal{B}(\mathbb{R})$, 或记为$X^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathscr{F}$. $X^{-1}(\mathcal{B}(\mathbb{R}))$是使随机变量$X$可测的最小$\sigma$代数.

之前, 我们说明了随机变量怎么产生一个分布函数(distribution function). 并说明了分布函数满足的一些条件: 单调非增, 规范性, 右连左极(书上是左连右极).

注记 (c.a.d.l.a.g.)**.** *a càdlàg (French: "continue à droite, limite à gauche"), RCLL ("right continuous with left limits"), or corlol ("continuous on (the) right, limit on (the) left") function is a function defined on the real numbers (or a subset of them) that is everywhere right-continuous and has left limits everywhere.*



例 1 (实轴上的测度)**.** *$(\mathbb{R}, \mathscr{B}(\mathbb{R}))$上的测度可以由一个Stirltjes测度函数$F$定义, Stirltjes测度函数$F$是指*

1. *F是单调非增的*
2. *F是右连续的.*

我们不加证明的给出以下的结论:

存在唯一的一个$(\mathbb{R}, \mathscr{B}(\mathbb{R}))$上的测度$\mu$满足$\mu((a,b]) = F(b) - F(a)$.

当$F(x) = x$时, 得到的测度μ是Lebesgue测度.

现在我们反过来说: 满足上述三个条件, 一定可以定义一个分布函数.

定理. *如果函数(右连左极)F满足单调非增, 规范性, 右连左极, 则函数F是某个随机变量X的分布函数.*

证明. 令$\Omega = (0,1)$, $\mathscr{F}$是上面的Borel集, 即$\mathscr{F} = \mathscr{B}(\mathbb{R}) \cap (0,1)$. 令$P$是一个Lebesgue测度(直线上的长度). 若$\omega \in (0,1)$, 定义

$$X(\omega) = \sup\{y : F(y) < \omega\}$$

只需证

$$\{\omega : X(\omega) \le x\} = \{\omega : \omega \le F(x)\}$$

因为$P(\omega : \omega \le F(x)) = F(x)$.

事实上, 一方面, 若$\omega \in \{\omega : \omega \le F(x)\}$, 即$\omega \le F(x)$, 则有$X(\omega) \le x$. 因为

$$\begin{aligned} X(\omega) &= \sup\{y : F(y) < \omega\} \\ &\le \sup\{y : F(y) < F(x)\} \\ &= x \end{aligned}$$

另一方面, 由于$X(\omega) \le x$, 即$\sup\{y : F(y) < \omega\} \le x = \sup\{y : F(y) < F(x)\}$, 根据$F$是右连左极的, 因此$\omega \le F(x)$.

$\square$

注记 (分布函数的逆)**.** 书上直接定义

$$F^{-1}(y) = \inf\{x : F(x) > y\}$$

在$0 < y < 1$时,

$$F^{-1}(F(y_0)) = \inf\{x : F(x) > F(y_0) = F(y_0 - 0)\}$$

若$F(y_0) = F(y_0+)$, 则$F^{-1}(F(y_0)) = y_0$; 若$F(y_0) < F(y_0+)$, 则$F^{-1}(F(y_0)) \geq y_0$. 上述逆也被称为"右连续逆"(right-continuous(RC) inverse), 也被记为$F^{\rightarrow}$.

若是右连左极的, 只需要把严格不等号改为不严格的.

$$F^{-1}(y) = \inf\{x : y \leq F(x)\}.$$

上述逆也被称为"左连续逆"(left-continuous(LC) inverse), 也被记为$F^{\leftarrow}$.

注记 (两种证明方法的等价性)**.** 定义$\sup\{\varnothing\} = 0$, 则

$$\inf\{x : F(x) > s\} = \sup\{x : F(x) \leq s\} \quad \inf\{x : F(x) \geq s\} = \sup\{x : F(x) < s\}$$

记$A = \inf\{x : F(x) > s\}$, 则当$F(x) > s$, $x \geq A$, $\forall \epsilon, \exists x_\epsilon, A < x + \epsilon$, 把$\epsilon$移项即由定义可得.

如果X和Y能推出相同的分布μ在$(\mathbb{R}, \mathscr{B}(\mathbb{R}))$上, 则称$X$和$Y$是"依分布相等"(equal in distribution)的. 根据测度的生成, 依分布相等等价于$P(X \leq x) = P(Y \leq x), \forall x$. 记之为

$$X =_d Y$$

例 2 (正态分布尾概率界的估计)**.** 对于任意的$x > 0$, 有

$$(x^{-1} - x^{-3}) \exp(-x^2/2) \leq \int_x^\infty \exp(-y^2/2) dy \leq x^{-1} \exp(-x^2/2)$$

证明. 对于上界, 写$y = x + z$, 有

$$\begin{aligned}
\int_x^\infty \exp(-y^2/2) dy &= \int_x^\infty \exp(-(x+z)^2/2) dz \\
&= \int_x^\infty \exp(-x^2/2 - xz - z^2/2) dz \\
&= \exp(-x^2/2) \int_0^\infty \exp(-xz) dz = x^{-1} \exp(-x^2/2)
\end{aligned}$$

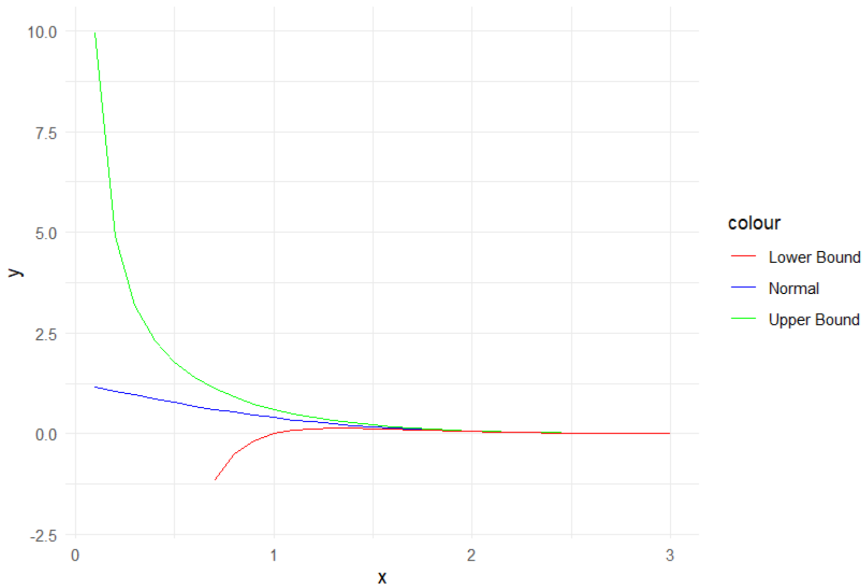
对于下界, 由于

$$\int_x^\infty (1 - 3y^{-4}) \exp(-y^2/2) dy = (x^{-1} - x^{-3}) \exp(-x^2/2)$$

即得. $\square$

例 3 (练习:正态分布的尾概率)**.** 用上述的上下界估计$P(N(0,1) > 4)$.

```
> x=4
> {(1-pnorm(x))*sqrt(2*pi)}
[1] 7.938803e-05
> {x^(-1) * exp(-(x^2)/2)}
[1] 8.386566e-05
> {(x^(-1) - x^(-3) ) * exp(-(x^2)/2)}
[1] 7.862405e-05
```



2 随机变量

在上一节, 我们介绍了一些具体的随机变量. 在这一节, 我们将梳理一些随机变量的具体性质.

定理. 若$\{\omega : X(\omega) \in A\} \in \mathscr{F}, \forall A \in \mathcal{A}$, 且$\mathcal{S} = \sigma(\mathcal{A})$. 则$X$是可测的.

证明. 简写$\{\omega : X(\omega) \in A\}$为$\{X \in B\}$, 则有

$$\{X \in \cup_i B_i\} = \cup_i \{X \in B_i\}$$
$$\{X \in B^c\} = \{X \in B\}^c$$

于是集合系$\mathcal{B} = \{B : \{X \in B\} \in \mathcal{F}\}$是一个$\sigma$域, $\mathcal{B} \supset \mathcal{A}, \mathcal{S} = \sigma(\mathcal{A})$, 从而$\mathcal{B} \supset \mathcal{S}$ □

上述结论表明了, 如果$\mathcal{S}$是一个σ域, 则$\{\{X \in B\} : B \in \mathcal{S}\}$也是一个$\Omega$上的$\sigma$域, 它是$\omega$上最小的保证$X$可测的结构, 也称它是"由随机变量$X$生成的$\sigma$域".

$$\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$$

例 4 (随机变量的证明)**.** 假设X和Y是定义在概率空间$(\Omega, \mathscr{F}, P)$, 当$A \in \mathscr{F}$上的随机变量. 定义$Z(\omega)$, 对于$\omega \in A$时, $Z(\omega) = X(\omega)$, 对于$\omega \in A^c$时, $Z(\omega) = Y(\omega)$, 证明Z是一个随机变量.

证明.
$$\begin{aligned} Z^{-1}(B) &= [Z^{-1}(B) \cap A] \cup [Z^{-1}(B) \cap A^c] \\ &= [X^{-1}(B) \cap A] \cup [Y^{-1}(B) \cap A^c] \\ &\in \mathcal{B}(\mathbb{R}) \end{aligned}$$

□

例 5. 证明: 分布函数只有至多可列的间断点.

证明. 设A是分布函数F的不连续点构成的集合, 则$\forall x < y \in A$, $F(x-) < F(x), F(y-) < F(y)$,特别的, 我们取$y-$充分靠近$y$, 即$x < y-$. 根据单调非降, 有

$$F(x-) < F(x) \leq F(y-) < F(y)$$

再根据有理数的稠密性, $\exists q_y \in (F(y-), F(y))$. 对每一个$A$中的元素重复上述操作. 定义$\phi : A \to \mathbb{Q}, x \mapsto \phi(x) = q_x$, 则$\phi$是一个到有理数上的单射, 因此$|A| \leq \aleph_0$, 从而间断点至多可列.

□

例 6 (正变换法)**.** *之前我们介绍分布函数定义随机变量的方法, 链接起了均匀分布和一般分布函数之间的关系, 现在我们要说明: 随机变量有统一的方法变换为(0,1)上的均匀分布. 即取$Y = F(X)$. 我们对F是连续函数的情况加以证明. 根据分布函数的连续性, 我们有*

$$\{F^{-1}(F(X)) \leq F^{-1}(y)\} = \{X \leq F^{-1}(y)\}$$

因此

$$\begin{aligned} P(F(X) \leq y) &= P(F^{-1}(F(X)) \leq F^{-1}(y)) \\ &= P(X \leq F^{-1}(y)) \\ &= F(F^{-1}(y)) \\ &= y \end{aligned}$$

对于$y = 0, P(F(X) \leq 0) = \lim_n P(F(X) \leq \frac{1}{n}) = \lim_n \frac{1}{n} = 0$. 对于$y = 1, P(F(X) \leq 1) = \lim_n P(F(X) \leq 1 - \frac{1}{n}) = \lim_n 1 - \frac{1}{n} = 1$.

定理 (随机变量的可测变换还是随机变量)**.** *若$X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$和$f : (S, \mathcal{S}) \to (T, \mathcal{T})$是两个可测映射, 那么$f(X)$是$(\Omega, \mathcal{F}) \overset{\prime}{to} (T, \mathcal{T})(\Omega, \mathcal{F})$到$(T, \mathcal{T})$的可测映射.*

证明. 取$B \in \mathcal{T}.\{\omega : f(X(\omega)) \in B\} = \{\omega : X(\omega) \overset{-1}{\inf}(B)\} \in \mathcal{F}$,根据$f$的可测性:$f^{-1}(B) \in \mathcal{S}$. □

3 随机变量独立性

注记 (A history of the symbol)**.** *Graham, Knuth, and Patashnik proposed using $\perp$ for relatively prime numbers in their book Concrete Mathematics, at least by the second edition (1994).*

Philip Dawid proposed a similar symbol $\perp\!\!\!\perp$ for (conditionally) independent random variables in 1979.

目前教材中共提到过四种独立性, 在此加以梳理

1. 事件的独立性
2. 随机变量的独立性
3. 随机向量的独立性
4. 试验的独立性

定义 (事件的独立性)**.** *若两事件满足 $P(AB) = P(A)P(B)$ 则称两事件A, B独立.*

定义 (随机变量的独立性)**.** 两个随机变量 X, Y 相互独立是指对于直线上任意两个$Borel$点集 A, B , $P(X \in A, X \in B) = P(X \in A)P(X \in B)$. 即事件$\{X \in A\}$和$\{X \in B\}$是独立的事件.

注记. 由于A, B是任意两个$Borel$点集, 因此可以这样通俗理解: 随机变量的独立性是指由他们生成的所有的事件都独立.

注记. $P(A) = P(I_A = 1)$, 事件独立也可以写成这种随机变量的形式, 显然I_A, I_B独立则事件A, B独立, 反过来由事件独立性能推对立事件的独立性, 从而I_A, I_B独立.

书上直接给出多元的情况,

定义 (随机变量的独立性*)**.** 设 $\xi_1, \cdots, \xi_n$ 为 n 个随机变量, 若对于任意的 $x_1, \cdots, x_n$ 成立

$$P\{\xi_1 < x_1, \cdots, \xi_n < x_n\} = P\{\xi_1 < x_1\} \cdots P\{\xi_n < x_n\}$$

则称 $\xi_1, \cdots, \xi_n$ 是相互独立的. 若 ξ_i 的分布函数为 $F_i(x), i = 1, 2, \cdots, n$, 它们的联合分布函数为 $F(x_1, \cdots, x_n)$, 则随机变量的独立性等价于对一切 $x_1, \cdots, x_n$ 成立

$$F(x_1, \cdots, x_n) = F_1(x_1) \cdots F_n(x_n)$$

在离散和连续的情况下又分别等价于概率质量的乘积,概率密度的乘积.

最后教材给出随机变量(向量)函数的独立性,

定理. 若 $\xi_1, \cdots, \xi_n$ 是相互独立的随机变量, 则 $f_1(\xi_1), \cdots, f_n(\xi_n)$ 也是相互独立的,这里 $f_i(i = 1, \cdots, n)$ 是任意的一元$Borel$函数.

证明. 对任意的一维Borel点集 $A_1, \cdots, A_n$ 有

$$\begin{aligned}
&P\{f_1(\xi_1) \in A_1, \cdots, f_n(\xi_n) \in A_n\} \\
=&P\left\{\xi_1 \in f_1^{-1}(A_1), \cdots, \xi_n \in f_n^{-1}(A_n)\right\} \\
=&P\left\{\xi_1 \in f_1^{-1}(A_1)\right\} \cdots P\left\{\xi_n \in f_n^{-1}(A_n)\right\} \\
=&P\{f_1(\xi_1) \in A_1\} \cdots P\{f_n(\xi_n) \in A_n\}
\end{aligned}$$

□

4 随机变量函数的分布

4.1 卷积公式

首先说明离散卷积公式若$Z = X + Y$ 可取 $0, 1, 2, \cdots$ 所有非负整数. 而事件 $\{Z = k\}$ 是如下不相容事件

$$\{X = i, Y = k - i\}, \quad i = 0, 1, \cdots, k$$

的并, 再考虑到独立性, 则对任意非负整数 k, 有

$$P(Z = k) = \sum_{i=0}^{k} P(X = i)P(Y = k - i).$$

这个概率等式被称为离散卷积公式.

在连续场合下, 若ξ_1, ξ_2相互独立, 则$\eta = \xi_1 + \xi_2$的密度函数为

$$q(y) = \int_{-\infty}^{\infty} p_1(u) p_2(y-u) du = \int_{-\infty}^{\infty} p_1(y-u) p_2(u) du$$

离散型应用: p相同情况下的二项分布, 不同参数的Poisson分布的可加性. 即:

$$Poi(\lambda_1) * Poi(\lambda_2) = Poi(\lambda_1 + \lambda_2)$$

$$b(n_1, p) * \cdots * b(n_k, p) = b(\sum_{i=1}^{k} n_i, p)$$

注记. 两*Poisson*随机变量差$K = N_1 - N_2$的分布不再是*Poisson*分布, 而是*Skellam*分布, 其形式为

$$p(k; \mu_1, \mu_2) = \Pr\{K = k\} = e^{-(\mu_1+\mu_2)} \left(\frac{\mu_1}{\mu_2}\right)^{k/2} I_k(2\sqrt{\mu_1 \mu_2})$$

其中μ_1, μ_2为*Poisson*分布的参数.

连续型应用: 正态分布的可加性,

$$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Γ分布的可加性,

$$X_1 \sim \Gamma(\alpha_1, \beta), X_2 \sim \Gamma(\alpha_2, \beta) \Rightarrow X_1 + X_2 \sim \Gamma(\alpha_1 + \alpha_2, \beta)$$

特别的, χ^2分布是Γ分布, 也具有可加性, 参数相同的指数分布加起来是Γ分布.

注记 (Γ分布的两种形式: shape form, scale form)**.** 根据第二个参数的不同写法有以下的两种形式*:*

- 尺度形式*:*

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(a)\beta^a} x^{a-1} \mathrm{e}^{-x/\beta}, & 0 < x < \infty \\ 0, & \text{其他} \end{cases}$$

- 速率形式*:*

$$f(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & 0 < x < \infty \\ 0, & \text{其他} \end{cases}$$

写成这两种情况的原因我们可以通过一个参数为λ的*Poisson*过程的例子说明*:* 我们假定拥有泊松过程的假设特别地，设随机变量W 表示所需获得的k 个事件*(*可能死亡*)*发生的时间，其中k是固定的正整数. 于是，W 的 cdf 为

$$G(w) = P(W \leqslant w) = 1 - P(W > w)$$

根据*Poisson*分布, 我们可以写出

$$P(W > w) = \sum_{x=0}^{k-1} P(X = x) = \sum_{x=0}^{k-1} \frac{(\lambda w)^x \mathrm{e}^{-\lambda w}}{x!}$$

于是λ是到达的速率, $\dfrac{1}{\lambda}$是等待时间.

不加证明的给出下面的概率等式,

$$\int_{\lambda w}^{\infty}\frac{z^{k-1}\mathrm{e}^{-z}}{(k-1)!}\mathrm{d}z=\sum_{x=0}^{k-1}\frac{(\lambda x)^{x}\,\mathrm{e}^{-\lambda w}}{x!}$$

则

$$G(w)=1-\int_{\lambda w}^{\infty}\frac{z^{k-1}\mathrm{e}^{-z}}{\Gamma(k)}\mathrm{d}z=\int_{0}^{\lambda w}\frac{z^{k-1}\mathrm{e}^{-z}}{\Gamma(k)}\mathrm{d}z$$

这对于 $w\leqslant 0, G(w)=0$ 也成立.若我们通过令 z=λy 定义$G(w)$对积分中的积分变量进行变量变换, 则

$$G(w)=\int_{0}^{w}\frac{\lambda^{k}y^{k-1}\mathrm{e}^{-\lambda y}}{\Gamma(k)}\mathrm{d}y,\quad w>0$$

因此*pdf*为:

$$g(w)=G'(w)=\begin{cases}\dfrac{\lambda^{k}w^{k-1}\mathrm{e}^{-\lambda w}}{\Gamma(k)}, & 0<w<\infty\\ 0, & \text{其他}\end{cases}$$

因此尺度形式刻画的是等待间隔, 速率形式刻画的是发生强度.

定理 (卷积定理)**.** 假设X,Y是两个独立的随机变量, F,G为其对应的累计分布函数, 则

$$P(X+Y\leq z)=\int F(z-y)dG(y).$$

其中$dG(y)$表示“使用分布函数G的, 关于测度ν的积分”.

$$\begin{aligned}P(X+Y\leq z)&=\int\int 1_{(x+y\leq z)}\mu(dx)\nu(dy)\\&=\int F(z-y)\nu(dy)=\int F(z-y)dG(y)\end{aligned}$$

4.2 顺序统计量

极值的分布首先求极大值 ξ_n^* 的分布函数,

$$\begin{aligned}&P\left\{\xi_n^*<x\right\}=P\left\{\max\left(\xi_1,\xi_2,\cdots,\xi_n\right)<x;\right.\\&=P\left\{\xi_1<x,\xi_2<x,\cdots,\xi_n<x\right\}\\&=P\left\{\xi_1<x\right\}\cdot P\left\{\xi_2<x\right\}\cdots P\left\{\xi_n<x\right\}\\&=[F(x)]^n\end{aligned}$$

其次求极小值 ξ_1^* 的分布函数, 注意到

$$\begin{aligned}P\left\{\xi_1^*\geqslant x\right\}&=P\left\{\min\left(\xi_1,\xi_2,\cdots,\xi_n\right)\geqslant x\right\}\\&=P\left\{\xi_1\geqslant x,\xi_2\geqslant x,\cdots,\xi_n\geqslant x\right\}\\&=P\left\{\xi_1\geqslant x\right\}P\left\{\xi_2\geqslant x\right\}\cdots P\left\{\xi_n\geqslant x\right\}\\&=[1-F(x)]^n\end{aligned}$$

顺序统计量方法采用概率元的方法(习题46).

样本中观测值落于 $(-\infty, x]$ 内的概率为 $F(x)$, 落人区间 $(x, x+\Delta x]$ 内的概率为 $F(x, x+\Delta x)-F(x)$, 落人区间 $(x+\Delta x, \infty)$ 内的概率为 $1-F(x, x+\Delta x)$. 将 n 个观测值分成这样的 3 组, 则其总的分法共有 $\dfrac{n!}{(k-1)!1!(n-k)!}$ 种. 于是, 由多项分布可得:

$$F_k(x+\Delta x)-F_k(x) \approx \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1}(F(x+\Delta x)-F(x))(1-F(x+\Delta x))^{n-k}.$$

对两边除以 Δx, 并令 $\Delta x \to 0$, 即有:

$$f_k(x)=\lim_{\Delta x} \frac{F_k(x+\Delta x)-F_k(x)}{\Delta x}=\frac{n!}{(k-1)!(n-k)!} F(x)^{k-1} f(x)(1-F(x))^{n-k}.$$

联合

$$\left\{\begin{array}{l} V=x_{(j)}-x_{(i)} \\ z=x_{(i)} \end{array}\right. \Rightarrow \quad \begin{array}{l} x_{(i)}=z \\ x_{ij}=v+z \end{array} \quad \Rightarrow\left\{\begin{array}{l} x=z \\ x=v+z \end{array}\right.$$

4.3 变量变换方法

设二维随机变量 (X, Y) 的联合密度函数为 $p(x, y)$, 如果函数

$$\left\{\begin{array}{l} u=g_1(x, y), \\ v=g_2(x, y) \end{array}\right.$$

有连续偏导数, 且存在唯一的反函数

$$\left\{\begin{array}{l} x=x(u, v), \\ y=y(u, v), \end{array}\right.$$

由反函数存在定理, 其变换的雅可比行列式

$$J=\frac{\partial(x, y)}{\partial(u, v)}=\left|\begin{array}{cc} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{array}\right|=\left(\frac{\partial(u, v)}{\partial(x, y)}\right)^{-1}=\left|\begin{array}{cc} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{array}\right|^{-1} \neq 0.$$

若

$$\left\{\begin{array}{l} U=g_1(X, Y), \\ V=g_2(X, Y), \end{array}\right.$$

则 (U, V) 的联合密度函数为

$$p(u, v)=p(x(u, v), y(u, v))|J|$$

教材上给出高维的例子, 如果对 $y_i=g_i\left(x_1, \cdots, x_n\right), i=1,2, \cdots, n$, 存在唯一的反函数 $x_i\left(y_1, \cdots, y_n\right)=x_i(i=1, \cdots, n)$, 而且 $\left(\eta_1, \cdots, \eta_n\right)$ 的密度函数为 $q\left(y_1, \cdots, y_n\right)$, 那么

$$G\left(y_1, \cdots, y_n\right)=\int_{\substack{u_1<y_1 \\ \cdots \\ u_n<y_n}} \cdots \int q\left(u_1, \cdots, u_n\right) \mathrm{d} u_1 \cdots \mathrm{d} u_n$$

可知

$$\begin{aligned} & q\left(y_1, \cdots, y_n\right) \\ = & \left\{\begin{array}{cc} p\left(x_1\left(y_1, \cdots, y_n\right), \cdots, x_n\left(y_1, \cdots, y_n\right)\right)|J| & \text{若 } \left(y_1, \cdots, y_n\right) \text{ 属于 } g_1, \cdots, g_n \text{ 的值域} \\ 0, & \text{其他} \end{array}\right. \end{aligned}$$

其中 J 为坐标变换的雅可比行列式

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \vdots & & \vdots \\ \dfrac{\partial x_n}{\partial y_1} & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix}$$

这里, 我们假定上述偏导数存在而且连续.

4.4 增补变量法

增补变量法实质上是变量变换法的一种应用: 为了求出二维连续随机乏量 (X,Y) 的函数 $U = g(X,Y)$ 的密度函数, 增补一个新的随机变量 $V = h(X,Y)$, 一般令 $V = X$ 或 $V = Y$. 先用变量变换法求出 (U,V) 的联合密度函数 $p(u,v)$, 再对 $p(u,v)$ 关于 v 积分, 从而得出关于 U 的边际密度函数. 设随机变量 X 与 Y 相互独立, 其密度函数分别为 $p_X(x)$ 和 $p_Y(y)$. 则 $U = XY$ 的密度函数为

$$p_U(u) = \int_{-\infty}^{\infty} p_X\left(\frac{u}{v}\right) p_Y(v) \frac{1}{|v|} \mathrm{d}v.$$

证明记 $V = Y$, 则 $\begin{cases} u = xy, \\ v = y \end{cases}$ 的反函数为 $\begin{cases} x = \dfrac{u}{v}, \\ y = v, \end{cases}$ 雅可比行列式为

$$J = \begin{vmatrix} \dfrac{1}{v} & -\dfrac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v},$$

所以 (U,V) 的联合密度函数为

$$p(u,v) = p_X\left(\frac{u}{v}\right) \cdot p_Y(v)|J| = p_X\left(\frac{u}{v}\right) p_Y(v) \frac{1}{|v|}.$$

对 $p(u,v)$ 关于 v 积分, 就可得 $U = XY$ 的密度函数为

$$p_U(u) = \int_{-\infty}^{\infty} p_X\left(\frac{u}{v}\right) p_Y(v) \frac{1}{|v|} \mathrm{d}v.$$