概率论:Week 9

1 随机变量及其分布

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.

2 独立性

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

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

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 $F\left(x_1,\cdots,x_n\right)$, 则随机变量的独立性等价于对一切 $x_1,\cdots,x_n$ 成立

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

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 $f_1\left(\xi_1\right),\cdots,f_n\left(\xi_n\right)$ 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

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

□

3 随机变量函数的分布

3.1 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.

3.2 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

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

3.3 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 $y_i = g_i(x_1, \cdots, x_n), i = 1, 2, \cdots, n$, 存在唯一的反函数 $x_i(y_1, \cdots, y_n) = x_i (i = 1, \cdots, n)$, 而且 $(\eta_1, \cdots, \eta_n)$ 的密度函数为 $q(y_1, \cdots, y_n)$ ，那么

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

可知

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

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.

3.4 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

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