

2021 大二下学期概率论期末复习

温尊

目录

1	基本公式和定理	3
1.1	概率公式计算汇总	3
1.2	基础公式	3
1.3	概率论定理	6
2	离散分布及其性质总结	10
2.1	Bernoulli 分布	10
2.2	二项分布	10
2.3	超几何分布	10
2.4	Poisson 分布	11
2.5	几何分布	11
2.6	帕斯卡分布	11
2.7	多项分布	11
2.8	多元超几何分布	12
3	连续分布及其性质总结	13
3.1	均匀分布	13
3.2	正态分布	13
3.3	指数分布	14
3.4	Erlang 分布	14
3.5	Γ 分布	14
3.6	χ^2 分布	14
3.7	多元均匀分布	15
3.8	多元正态分布	15

4	其他零散分布	18
4.1	Cauchy 分布	18
4.2	F 分布	18
4.3	Rayleigh 分布	18
5	常见反例	19

1 基本公式和定理

1.1 概率公式计算汇总

定理 1.1 (加法公式). 对概率空间 (Ω, \mathcal{F}, P) , 若 $A_1, \dots, A_n \in \mathcal{F}$, 则

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n \sum_{i_1 < \dots < i_k} (-1)^{k-1} P(A_{i_1} \cdots A_{i_k}).$$

定理 1.2 (减法公式). 对概率空间 (Ω, \mathcal{F}, P) , 若 $A, B \in \mathcal{F}$, 则

$$P(A - B) = P(A) - P(AB).$$

定理 1.3 (乘法公式). 对概率空间 (Ω, \mathcal{F}, P) , 若 $A_1, \dots, A_n \in \mathcal{F}$, 则

$$P(A_1 \cdots A_n) = \prod_{i=1}^n P(A_i | A_1 \cdots A_{i-1}) = \sum_{k=1}^n \sum_{i_1 < \dots < i_k} (-1)^{k-1} P(A_{i_1} \cup \dots \cup A_{i_k}).$$

定理 1.4 (全概率公式). 对概率空间 (Ω, \mathcal{F}, P) , 若 $B, A_1, \dots, A_n, \dots \in \mathcal{F}$, 其中 A_i 是 Ω 的分割, 则

$$P(B) = \sum_j P(A_j)P(B|A_j).$$

定理 1.5 (Bayes 公式). 对概率空间 (Ω, \mathcal{F}, P) , 若 $B, A_1, \dots, A_n, \dots \in \mathcal{F}$, 其中 A_i 是 Ω 的分割, 则

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_j P(A_j)P(B|A_j)}.$$

1.2 基础公式

定理 1.6 (边际分布). (1)/离散/ 考虑二维随机变量 (ξ, η) , 其中 ξ 取值 x_1, x_2, \dots , 而 η 取值 y_1, y_2, \dots , 令 $P(\xi = x_i, \eta = y_j) = p(x_i, y_j)$, 令 $P(\xi = x_i) = p_1(x_i), P(\eta = y_j) = p_2(y_j)$, 则

$$\sum_j p(x_i, y_j) = p_1(x_i), \sum_i p(x_i, y_j) = p_2(y_j),$$

这就是边际分布;

(2)/一般/ 考虑二维随机变量 (ξ, η) , 其分布函数为 $F(x, y)$, 而 $F_1(x) = P(\xi < x) = F(x, +\infty)$ 且 $F_2(y) = P(\eta < y) = F(+\infty, y)$, 这就是 $F(x, y)$ 的边际分布函数.

(3)/连续/ 若分布函数为 $F(x, y)$ 是连续型, 有密度函数 $p(x, y)$, 则

$$F_1(x) = \int_{-\infty}^x du \int_{-\infty}^{\infty} p(u, y) dy, F_2(y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^y p(x, v) dv,$$

则 $F_1(x), F_2(y)$ 分别的密度函数为

$$p_1(x) = \int_{\mathbb{R}} p(x, y) dy, p_2(y) = \int_{\mathbb{R}} p(x, y) dx,$$

称为边际密度函数.

定理 1.7 (条件分布). (1)[一般] 定义为

$$P(\eta < y | \xi = x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x, y) - F(x, y)}{F(x + \Delta x, \infty) - F(x, \infty)};$$

(2)[离散] 已知 $\xi = x_i, p_1(x_i) > 0$, 则事件 $\{\eta = y_j\}$ 的条件概率为

$$P(\eta = y_j | \xi = x_i) = \frac{p(x_i, y_j)}{p_1(x_i)};$$

(3)[连续] 给定 $\xi = x$, 则 η 的分布密度函数为

$$p(y|x) = \frac{p(x, y)}{p_1(x)}.$$

定理 1.8 (离散卷积公式). 若 ξ, η 是相互独立的随机变量, 且取非负整数, 其概率分布为 $\{a_k\}, \{b_k\}$, 则 $\zeta = \xi + \eta$ 的概率分布为

$$c_r = P(\zeta = r) = a_0 b_r + \cdots + a_r b_0, r = 0, 1, \cdots.$$

定理 1.9. (1) 若 g 严格单调, 其反函数 $g^{-1}(x)$ 有连续导函数, 则 $\eta = g(\xi)$ 有密度函数

$$p(g^{-1}(y)) |(g^{-1}(y))'|;$$

(2) 若 $g(x)$ 在不相重叠的区间 I_1, \dots 上逐段单调, 反函数分别为 $h_1(y), \dots$, 且其导数连续, 则 $\eta = g(\xi)$ 密度函数为

$$\sum_i p(h_i(y)) |h'_i(y)|.$$

定理 1.10 (随机向量的函数分布律). 若 $\eta = g(\xi_1, \dots, \xi_n)$, 且 (ξ_1, \dots, ξ_n) 的密度函数为 $p(x_1, \dots, x_n)$, 则 η 的分布函数为

$$G(y) = P(\eta < y) = \int_{g < y} p(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

(1)[和的分布] 若 $\eta = \xi_1 + \xi_2$, 且 (ξ_1, ξ_2) 密度函数为 $p(x_1, x_2)$, 则其分布函数为

$$G(y) = \int_{\mathbb{R}} \int_{-\infty}^{y-x_1} p(x_1, x_2) dx_1 dx_2,$$

若 ξ_1, ξ_2 独立, 设 $p = p_1(x_1)p_2(x_2)$, 则 η 的密度函数为

$$q(y) = \int_{\mathbb{R}} p_1(u)p_2(y-u)du = \int_{\mathbb{R}} p_1(y-u)p_2(u)du;$$

(2)[商的分布] 若 $\eta = \frac{\xi_1}{\xi_2}$, 且 (ξ_1, ξ_2) 密度函数为 $p(x_1, x_2)$, 则其密度函数为

$$q(X) = \int_{\mathbb{R}} |z|p(zx, z)dz;$$

(3) [顺序统计量] 设相互独立且具有相同分布函数和密度函数的 (ξ_1, \dots, ξ_n) , 按顺序排位 $\xi_1^* \leq \dots \leq \xi_n^*$, 其中 $\xi_1^* = \min(\xi_i), \xi_n^* = \max(\xi_i)$.

连续性的话, 设密度为 $p(x)$, 则 ξ_k^* 的密度函数为

$$\frac{n!}{(k-1)!(n-k)!} (F(x))^{k-1} (1-F(x))^{n-k} p(x).$$

定理 1.11 (随机向量的变换). 若 (ξ_1, \dots, ξ_n) 的密度函数为 $p(x_1, \dots, x_n)$, 考虑 $\eta_k = g_k(\xi_1, \dots, \xi_n), k = 1, \dots, m$ 的分布律, 这时有

$$G(y_1, \dots, y_m) = \int_{g_1 < y_1, \dots, g_m < y_m} p(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

当 $m = n$, 且 $y_i = g_i(x_1, \dots, x_n)$ 有唯一反函数 $x_i = x_i(y_1, \dots, y_n)$ 时, 设 (η_1, \dots, η_n) 的密度函数为

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

其中 $J = \left(\frac{\partial x_i}{\partial y_j} \right)_{i,j}$ 为 *Jacobian* 矩阵.

定理 1.12 (数学期望的性质). (i) 若 $a \leq \xi \leq b$, 则 $a \leq E\xi \leq b$;

(ii) 对任意常数 $c_i, i = 1, \dots, n$ 及 b , 有

$$E \left(\sum_{i=1}^n c_i \xi_i + b \right) = \sum_{i=1}^n c_i E\xi_i + b.$$

注 1.1. 不难得知 $D\xi = E\xi^2 - (E\xi)^2$.

定理 1.13 (方差的性质). (i) 常数方差为 0;

(ii) 对常数 b, c , 有 $D(b\xi + c) = b^2 D\xi$;

(iii) 若 $c \neq E\xi$, 则 $D\xi < E(\xi - c)^2$.

注 1.2. 标准化随机变量 $\xi^* = \frac{\xi - E\xi}{\sqrt{D\xi}}$, 则 $E\xi^* = 0, D\xi^* = 1$.

定理 1.14. 对 n 维随机变量 $\xi = (\xi_1, \dots, \xi_n)^T$, 设其期望为 μ , 协方差矩阵为 Σ .

(i) 对线性组合 $\zeta = l^T \xi$, 有 $E\zeta = l^T \mu, D\zeta = l^T \Sigma l$;

(ii) 对 $\eta = C\xi$, 有 $E\eta = C\mu, D\eta = C\Sigma C^T$.

定理 1.15. 对 (ξ_1, \dots, ξ_n) . 定义协方差为 $\sigma_{ij} = \text{cov}(\xi_i, \xi_j) = E((\xi_i - E\xi_i)(\xi_j - E\xi_j))$, 则 $\text{cov}(\xi_i, \xi_j) = E\xi_i \xi_j - E\xi_i E\xi_j$ 和 $D(\sum_{i=1}^n \xi_i) = \sum_{i=1}^n D\xi_i + 2 \sum_{i < j} \text{cov}(\xi_i, \xi_j)$.

相关系数为 $\rho_{ij} = \frac{\text{cov}(\xi_i, \xi_j)}{\sqrt{D\xi_i D\xi_j}}$. 补充定义常数和任何随机变量的相关系数为 0.

$\text{cov}(a\xi + b, c\eta + d) = a \text{cov}(\xi, \eta)$, 且当 $ac > 0$ 时 $\rho_{a\xi+b, c\eta+d} = \rho_{\xi\eta}$; 当 $ac < 0$ 时 $\rho_{a\xi+b, c\eta+d} = -\rho_{\xi\eta}$.

定理 1.16 (相关系数的性质). (i) 由柯西不等式得到 $|\rho| \leq 1$, 而当 $\rho = \pm 1$ 时, 二者为线性关系 (完全正相关, 完全负相关), 当 $\rho = 0$, 称为不相关;

(ii) 对随机变量 ξ, η , 二者不相关当且仅当 $\text{cov}(\xi, \eta) = 0$ 当且仅当 $E\xi\eta = E\xi E\eta$ 当且仅当 $D(\xi + \eta) = D\xi + D\eta$;

(iii) 若 ξ, η 独立, 则不相关, 反之不对; (不相关也可以有函数关系, 见李贤平 P211)

(iv) 二元正态分布和两个二值随机变量的情况, 独立和不相关等价;

定理 1.17 (特征函数的性质 (多元类似)). (1) $f(0) = 1, |f(t)| \leq f(0), f(-t) = \overline{f(t)}$;

(2) 特征函数在 \mathbb{R} 上一致连续;

(3) $\forall n \in \mathbb{N}, t_i \in \mathbb{R}, \lambda_i \in \mathbb{C}$, 则 $\sum_k \sum_j f(t_k - t_j) \lambda_k \bar{\lambda}_j \geq 0$;

(4) 相互独立随机变量之和的特征函数为其特征函数之积;

(5) 随机变量 ξ 的 n 阶矩存在, 则其特征函数 n 次可微, 且当 $k \leq n$ 时 $f^{(k)}(0) = i^k E\xi^k$;

(6) 设 $\eta = a\xi + b$, 则 $f_\eta(t) = e^{ibt} f_\xi(at)$;

(7) 分布函数唯一被特征函数决定.

定理 1.18 (多元特征函数的性质). (1) 若 (ξ_1, \dots, ξ_n) 特征函数 $f(t_1, \dots, t_n)$, 而 ξ_i 的特征函数 $f_{\xi_i}(t)$, 则他们相互独立当且仅当 $f(t_1, \dots, t_n) = f_{\xi_1}(t) \cdots f_{\xi_n}(t)$;

(2) 特征函数列 $\{f_k(x_1, \dots, x_n)\}$ 收敛于连续函数 $f(x_1, \dots, x_n)$, 则其为某特征函数.

定义 1.1. 设 $\xi_1, \dots, \xi_n, \dots$ 是独立随机变量序列, 令 $\eta_n = \frac{\sum_{i=1}^n \xi_i}{n}$, 如果存在常数列 a_i 使得对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P(|\eta_n - a_n| < \varepsilon) = 1,$$

则称 $\{\xi_n\}$ 服从大数定律.

假定 $E\xi_i, D\xi_i$ 存在, 令 $\zeta_n = \frac{\sum_{i=1}^n \xi_i - \sum_{i=1}^n E\xi_i}{\sqrt{\sum_{i=1}^n D\xi_i}}$, 如果

$$\lim_{n \rightarrow \infty} P(\zeta_n < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2} dt,$$

则称 $\{\xi_n\}$ 服从中心极限定理.

1.3 概率论定理

定理 1.19 (二项分布的 Poisson 逼近). 独立实验中, 设 p_n 为事件 A 出现的概率, 它和实验总数有关, 若 $np_n \rightarrow \lambda$, 则当 $n \rightarrow \infty$ 时

$$b(k; n, p_n) \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}.$$

定理 1.20 (随机变量函数独立性). 若 ξ_1, \dots, ξ_n 相互独立, 则对任意的 Borel 函数 f_i , 变量 $f_k(\xi_k)$ 也相互独立.

定理 1.21 (佚名统计学家公式). 若 $g(x)$ 是一元 Borel 函数, 而 $\eta = g(\xi)$, 则

$$E\eta = \int_{\mathbb{R}} y dF_{\eta}(y) = \int_{\mathbb{R}} g(x) dF_{\xi}(x).$$

若 ξ 具有密度函数 $p(x)$, 则 $E\eta = \int_{\mathbb{R}} g(x)p(x)dx$.

若 (ξ_1, \dots, ξ_n) 分布函数为 $F(x_1, \dots, x_n)$, 而 $g(x_1, \dots, x_n)$ 是 n 元 Borel 函数, 则

$$Eg(\xi_1, \dots, \xi_n) = \int_{\mathbb{R}} g(x_1, \dots, x_n) dF(x_1, \dots, x_n),$$

特别的,

$$E\xi_1 = \int_{\mathbb{R}} x_1 dF(x_1, \dots, x_n) = \int_{\mathbb{R}} x_1 dF_1(x_1),$$

其中 F_1 是 ξ_1 分布函数.

复佚名统计学家公式:

$$Ee^{itg(\xi)} = \int_{\mathbb{R}} e^{itg(x)} dF_{\xi}(x).$$

定理 1.22 (切比雪夫不等式). 对任何具有有限方差的随机变量 ξ , 有

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2},$$

其中 ε 是任意正数.

证明. 注意到 $\int_{\mathbb{R}} (x - E\xi)^2 dF(x) \geq \int_{|x - E\xi| \geq \varepsilon} (x - E\xi)^2 dF(x)$ 即可. □

注 1.3. (i) 有时将其变形为 $P(|\xi - E\xi| < \varepsilon) \geq 1 - \frac{D\xi}{\varepsilon^2}$ 或 $P\left(\left|\frac{\xi - E\xi}{\sqrt{D\xi}}\right| \geq \delta\right) \leq \frac{1}{\delta^2}$;

(ii) 可以证明方差为 0 的随机变量必为常数.

定理 1.23 (柯西不等式). 对任何随机变量 ξ, η , 有

$$|E\xi\eta|^2 \leq E\xi^2 E\eta^2,$$

等号成立当且仅当存在 t_0 使得 $P(\eta = t_0\xi) = 1$.

定理 1.24 (重期望公式). 对任何随机变量 ξ, η , 有

$$E\eta = E(E(\eta|\xi)).$$

定理 1.25 (逆转公式). 分布函数 F 的特征函数 f , 设 x_1, x_2 是 F 的连续点, 则

$$F(x_2) - F(x_1) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-itx_1} - e^{-itx_2}}{it} f(t) dt.$$

推论 1. 若 $\int_{\mathbb{R}} |f| < \infty$, 则其分布函数 F 导数存在且连续, 且

$$F'(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} f(t) dt.$$

定理 1.26 (切比雪夫大数定律). 若 $\{\xi_i\}$ 两两不相关, 且都有有限方差, 且有公共上界 $D\xi_i \leq C$, 则对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E\xi_k \right| < \varepsilon \right) = 1.$$

定理 1.27 (马尔可夫大数定律). 若 $\{\xi_i\}$ 满足 $\frac{1}{n} D \sum_{k=1}^n \xi_k \rightarrow 0$ (无独立性假设), 则对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n E\xi_k \right| < \varepsilon \right) = 1.$$

推论 2 (伯努利大数定律). 设 μ_n 是 n 次伯努利实验中 A 出现的次数, 而概率是 p , 则任给 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{\mu_n}{n} - p \right| < \varepsilon \right) = 1.$$

推论 3 (泊松大数定律). 若独立试验序列中, 事件 A 在第 k 次实验的概率为 p_k , 设 μ_n 是前 n 次实验中 A 出现的次数, 则任给 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{\mu_n}{n} - \frac{p_1 + \dots + p_n}{n} \right| < \varepsilon \right) = 1.$$

定理 1.28 (棣莫弗-拉普拉斯极限定理). 若 μ_n 是 n 次伯努利实验中 A 出现的次数, 而概率是 p , 则对任一有限区间 $[a, b]$:

(1) [局部极限定理] 当 $a \leq x_k \equiv \frac{k-np}{\sqrt{npq}} \leq b$ 及 $n \rightarrow \infty$ 时, 一致的有

$$\frac{P(\mu_n = k)}{\left(\frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-1/2x_k^2} \right)} \rightarrow 1;$$

(2) [积分极限定理] 当 $n \rightarrow \infty$ 时, 一致的有

$$P \left(a \leq \frac{\mu_n - np}{\sqrt{npq}} < b \right) \rightarrow \int_a^b \phi(x) dx, \text{ where } \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

注 1.4. 局部极限定理更精细的有

$$P(\mu_n = k) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{npq}} e^{-x_k^2/2} \left(1 + \frac{(q-p)(x_k^3 - 3x_k)}{6\sqrt{npq}} + O\left(\frac{1}{n}\right) \right).$$

注 1.5 (一些应用). (a) 频率估计概率的计算: 注意到

$$P \left(\left| \frac{\mu_n}{n} - p \right| < \varepsilon \right) \approx 2\Phi \left(\varepsilon \sqrt{\frac{n}{pq}} \right) - 1;$$

(b) 概率的置信区间估计: 置信水平 β 下得到的概率 p 的置信区间为

$$\left(\frac{\mu_n}{n} - z_\beta \sqrt{\frac{\frac{\mu_n}{n}(1 - \frac{\mu_n}{n})}{n}}, \frac{\mu_n}{n} + z_\beta \sqrt{\frac{\frac{\mu_n}{n}(1 - \frac{\mu_n}{n})}{n}} \right),$$

其中 $2\Phi(z_\beta) - 1 = \beta$;

(c) 局部极限定理估计二项分布: 当 n 较大时有

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{npq}} \phi\left(\frac{k-np}{\sqrt{npq}}\right),$$

其中 $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$;

(d) 积分极限定理估计二项分布: 修正公式为

$$P(k_1 \leq \mu_n \leq k_2) \approx \Phi\left(\frac{k_2 - np + 0.5}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np - 0.5}{\sqrt{npq}}\right).$$

定理 1.29 (辛钦大数定律). 设 $\xi_1, \dots, \xi_n, \dots$ 是相互独立的随机变量序列, 服从相同分布, 且具有有限数学期望 $a = E\xi_n$, 则对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n \xi_i - a\right| < \varepsilon\right) = 1.$$

定理 1.30 (中心极限定理-林德伯格-莱维). 设 $\xi_1, \dots, \xi_n, \dots$ 是相互独立的随机变量序列, 服从相同分布, 且 $E\xi_k = \mu, D\xi_k = \sigma^2$, 考虑 $\zeta_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (\xi_k - \mu)$, 若 $0 < \sigma^2 \leq \infty$, 则

$$\lim_{n \rightarrow \infty} P(\zeta_n < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

定理 1.31 (多元中心极限定理). 若 p 维随机向量 $\xi_1, \dots, \xi_n, \dots$ 相互独立, 且具有相同分布, 数学期望为 μ , 协方差矩阵为 Σ , 则

$$\zeta_n = \frac{\sum_{k=1}^n (\xi_k - \mu)}{\sqrt{n}}$$

的极限分布为 $N(\mathbf{0}, \Sigma)$.

2 离散分布及其性质总结

2.1 Bernoulli 分布

定义 2.1. 只有两种结果, 事件域为 $\mathcal{F} = \{\emptyset, A, \bar{A}, \Omega\}$, 满足 $P(A) = p, P(\bar{A}) = q$, 且 $p, q \geq 0, p + q = 1$.

注 2.1. 分布为 $b_k = P(\beta = k) = p^k q^{1-k}, k = 0, 1$, 其中 β 为 A 出现次数, 仅取 $0, 1$.

注 2.2 (推广 Bernoulli 分布). n 次重复独立实验, 每次实验可能结果为 A_1, \dots, A_r , 满足 $P(A_i) = p_i, \sum_i p_i = 1$.

命题 2.1 (Bernoulli 分布的数字特征). (1) 数学期望 $E1_A = P(A)$;

(2) 方差 $D\xi = p(1-p)$;

2.2 二项分布

定义 2.2. 做 n 重 Bernoulli 实验, 设 μ 为成功次数, 其为随机变量, 取值为 $0, 1, \dots, n$, 概率分布 $\mu \sim B(n, p)$ 为

$$b(k; n, p) = P(\mu = k) = \binom{n}{k} p^k q^{n-k}, k = 0, 1, \dots, n.$$

命题 2.2 (二项分布的性质). (1) 对称性: $b(k; n, p) = b(n-k; n, 1-p)$;

(2) 单调性: $k \leq (n+1)p$ 时递增; $k > (n+1)p$ 时递减;

(3) 二项分布在 p 相当小的时候, 用 Poisson 逼近 (定理 1.19) 有

$$b(k; n, p) \approx \frac{(np)^k}{k!} e^{-np}.$$

命题 2.3 (二项分布的特征). (1) 数学期望 $E\xi = np$;

(ii) 方差 $D\xi = np(1-p)$;

(iii) 特征函数 $f(t) = (pe^{it} + q)^n$

2.3 超几何分布

定义 2.3. 某批 N 件产品有 M 件次品, 随机抽出 n 件, 出现次品数量 ν 为随机变量, 且

$$h_k = P(\nu = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, 0 \leq k \leq n \leq N, k \leq M.$$

注 2.3. 当 $N \gg n$ 时, 可以用二项分布近似.

2.4 Poisson 分布

定义 2.4. 设 $\lambda > 0$ 且 ξ 可取一切非负整数, 称 ξ 服从 Poisson 分布 $\xi \sim P(\lambda)$, 如果

$$P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k \in \mathbb{Z}_{\geq 0}.$$

命题 2.4 (Poisson 分布的特征). (1) 数学期望 $E = \lambda$;

(2) 方差 $D\xi = \lambda$;

(3) 特征函数 $f(t) = e^{\lambda(e^{it}-1)}$

2.5 几何分布

定义 2.5. 成功概率为 p 的 Bernoulli 实验中, 以 η 记成功首次出现时的试验次数, 则其为随机变量, 可能取值为 $1, 2, 3, \dots$, 概率分布为

$$g(k, p) = P(\eta = k) = q^{k-1}p, k = 1, 2, \dots.$$

命题 2.5 (性质). (1) 无记忆性: 若已知前 m 次实验没有成功, 设达到首次成功的等待时间为 η' , 则 $P(\eta' = k) = q^{k-1}p, k = 1, 2, \dots$ (离散分布里只有几何分布有此性质);

命题 2.6 (几何分布的数字特征). (1) 数学期望 $E = \frac{1}{p}$;

2.6 帕斯卡分布

定义 2.6. 成功概率为 p 的 Bernoulli 实验中, 以 ζ 记成功第 r 次出现时的试验次数, 取值为 $r, r+1, \dots$, 概率分布为

$$P(\zeta = k) = \binom{k-1}{r-1} p^r q^{k-r}, k = r, r+1, \dots.$$

注 2.4 (推广). 对任意实数 $r > 0$, 称

$$Nb(l; r, p) = \binom{-r}{l} p^r (-q)^l, l = 0, 1, \dots$$

为负二项分布.

2.7 多项分布

定义 2.7. 设实验可能结果为 A_1, \dots, A_r , 且 $P(A_i) = p_i$ 且 $p_1 + \dots + p_r = 1$, 重复 n 次, 假定相互独立, 若以 ξ_i 为 A_i 出现次数, 则

$$P(\xi_i = k_i) = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}.$$

命题 2.7 (多项分布的性质). (1) 协方差 $\text{cov}(\xi_i, \xi_j) = -np_i p_j$, 相关系数 $\rho_{ij} = -\sqrt{\frac{p_i p_j}{(1-p_i)(1-p_j)}}$;

2.8 多元超几何分布

定义 2.8. 袋子中有 N_i 个 i 号球, $i = 1, \dots, r$, 且 $N_1 + \dots + N_r = N$, 从中摸出 n 只, 若以 ξ_i 为 i 号球的出现次数, 则

$$P(\xi_i = n_i) = \frac{\binom{N_1}{n_1} \dots \binom{N_r}{n_r}}{\binom{N}{n}}.$$

3 连续分布及其性质总结

3.1 均匀分布

定义 3.1. 设 a, b 为有限数, 则下面密度函数定义的就是 $[a, b]$ 上的均匀分布:

$$p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & \text{其它.} \end{cases}$$

注 3.1. (1) 其分布函数为 $F(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b; \\ 1, & x > b. \end{cases}$;

(2) 若 θ 服从 $[0, 1]$ 均匀分布, 那么对任意分布函数 $F(x)$, 令 $\xi = F^{-1}(\theta)$, 则不难看出 ξ 是服从分布函数 F 的随机变量;

命题 3.1 (均匀分布的数字特征). (1) 数学期望 $E\xi = \frac{b+a}{2}$;

(1) 方差 $D\xi = \frac{(b-a)^2}{12}$;

3.2 正态分布

定义 3.2. 密度函数为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R},$$

其中 $\sigma > 0, \mu, \sigma$ 为常数, 其分布函数为

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy, x \in \mathbb{R},$$

记为 $N(\mu, \sigma^2)$.

注 3.2. 当 $\mu = 0, \sigma = 1$ 称为标准正态分布, 其密度函数和分布函数分别记为 ϕ, Φ .

命题 3.2 (性质). (1) 若 $\xi \sim N(\mu, \sigma^2)$, 则 $\eta = \frac{\xi-\mu}{\sigma} \sim N(0, 1)$;

(2) 若 $\xi \sim N(\mu, \sigma^2)$, 则 $F(x) = P\left(\frac{\xi-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, 且 $P(|\xi - \mu| < k\sigma) = 2\Phi(k) - 1$, 且 $P(a \leq \xi < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$;

(3) 相互独立分布相同的两个随机变量 ξ, η 如果满足密度函数不等于零且二阶可导, 且 $\xi + \eta, \xi - \eta$ 相互独立, 那么 $\xi, \eta, \xi + \eta, \xi - \eta$ 都服从正态分布;

命题 3.3 (正态分布的特征). (1) 数学期望 $E\xi = \mu$;

(2) 方差 $D\xi = \sigma^2$, 则 σ 为标准差;

(3) 特征函数 $f(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}$

3.3 指数分布

定义 3.3. 密度函数为 $p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0, \end{cases}$ 分布函数为 $F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0, \end{cases}$, 这里 $\lambda > 0$ 为参数, 称为指数分布, 记为 $\text{Exp}(\lambda)$.

命题 3.4 (性质). (1) 无记忆性: 对任意的 $s, t > 0$, 有 $P(\xi \geq s + t | \xi \geq s) = P(\xi \geq t)$ (连续分布里只有指数分布有此性质);

命题 3.5 (正态分布的数字特征). (1) 数学期望 $E = \frac{1}{\lambda}$;

3.4 Erlang 分布

定义 3.4. 对任意的 $r \in \mathbb{Z}_{>0}, \lambda > 0$, 密度函数

$$p(x) = \frac{\lambda^r}{(r-1)!} x^{r-1} e^{-\lambda x}, x \geq 0,$$

称为 *Erlang* 分布.

3.5 Γ 分布

定义 3.5. 对任意的 $r, \lambda > 0$, 密度函数

$$f(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x > 0; \\ 0, & x \leq 0, \end{cases}$$

称为 Γ 分布.

3.6 χ^2 分布

定义 3.6. 密度函数为

$$p(x) = \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2}, x > 0$$

的分布为具有自由度 n 的 χ^2 分布, 记作 $\xi \sim \chi_n^2$.

注 3.3. 1. 不难得知 χ^2 分布是 Γ 分布的特例;

2. 若相互独立的 χ^2 分布 ξ, η 自由度分别为 m, n , 则 $\xi + \eta$ 服从 $m + n$ 的 χ^2 分布;

3. 对相互独立且服从 $N(0, 1)$ 的随机变量 ξ_1, \dots, ξ_n , 则 $\eta = \xi_1^2 + \dots + \xi_n^2$ 服从 χ^2 分布;

3.7 多元均匀分布

定义 3.7. 若 $G \subset \mathbb{R}^n$ 的正测有限区域, 则密度函数

$$p(x_1, \dots, x_n) = \begin{cases} \frac{1}{m(G)}, & (x_1, \dots, x_n) \in G; \\ 0, & (x_1, \dots, x_n) \notin G, \end{cases}$$

称为多元均匀分布.

3.8 多元正态分布

定义 3.8. 若 $\Sigma = (\sigma_{ij})$ 是 n 阶正定矩阵, 设逆矩阵 $\Sigma^{-1} = (\gamma_{ij})$. 设 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$ 为任意实值列向量, 则由密度函数

$$\begin{aligned} p(x_1, \dots, x_n) &= \frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2} \sum_{j,k=1}^n r_{jk}(x_j - \mu_j)(x_k - \mu_k)\right) \\ &= \frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \end{aligned}$$

定义的分量为 n 元正态分布, 记为 $N(\boldsymbol{\mu}, \Sigma)$.

命题 3.6 (性质). 1. (二元正态分布的典型分解) 二元正态分布密度函数 $p(x, y)$ 有两个分解:

$$\begin{aligned} p(x, y) &= \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{\left(y - \left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)\right)\right)^2}{2\sigma_2^2(1-\rho^2)}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(y - \mu_2)^2}{2\sigma_2^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp\left(-\frac{\left(x - \left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2)\right)\right)^2}{2\sigma_1^2(1-\rho^2)}\right). \end{aligned}$$

2. 二元正态分布的边际密度函数: 为 $p_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right)$, 同理 $p_2(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(y - \mu_2)^2}{2\sigma_2^2}\right)$, 这就得知二元正态分布的边际分布也是正态分布 (反之不对);

3. 二元正态分布的条件分布仍然是正态分布;

4. 二元正态分布独立当且仅当 $\rho = 0$, 且独立当且仅当不相关;

5. 二元正态分布中 μ_1, μ_2 为其两个边际分布的数学期望, 而 σ_1, σ_2, ρ 构成协方差矩阵

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix};$$

定理 3.1 (特征函数). 为 $f(\mathbf{t}) = \exp\left(i\boldsymbol{\mu}^T \mathbf{t} - \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right)$.

证明. 我们知道

$$\begin{aligned} f(\mathbf{t}) &= \int_{\mathbb{R}^n} e^{i\mathbf{t}^T \mathbf{x}} p(\mathbf{x}) d\mathbf{x} \\ &= \frac{1}{(2\pi)^{n/2} (\det \boldsymbol{\Sigma})^{1/2}} \int_{\mathbb{R}^n} e^{i\mathbf{t}^T \mathbf{x}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) d\mathbf{x}, \end{aligned}$$

设可逆矩阵 \mathbf{L} 满足 $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T$, 做变换 $\mathbf{y} = \mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu})$, 其逆变换为 $\mathbf{x} = \mathbf{L}\mathbf{y} + \boldsymbol{\mu}$, 设 $\mathbf{s} = \mathbf{L}^T \mathbf{t}$, 则 $i\mathbf{t}^T \mathbf{x} = i\mathbf{t}^T \boldsymbol{\mu} + i(\mathbf{L}^T \mathbf{t})^T \mathbf{y}$, 故

$$i\mathbf{t}^T \mathbf{x} - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = i\boldsymbol{\mu}^T \mathbf{t} - \frac{1}{2} \sum_{k=1}^n (y_k - i s_k)^2 - \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t},$$

则

$$\begin{aligned} f(\mathbf{t}) &= \frac{\exp\left(i\boldsymbol{\mu}^T \mathbf{t} - \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right)}{(2\pi)^{n/2} (\det \boldsymbol{\Sigma})^{1/2}} \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2} \sum_{k=1}^n (y_k - i s_k)^2\right) (\det \boldsymbol{\Sigma})^{1/2} d\mathbf{y} \\ &= \exp\left(i\boldsymbol{\mu}^T \mathbf{t} - \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right), \end{aligned}$$

得到结论. □

定理 3.2. $\boldsymbol{\xi}$ 的任一子向量 $(\xi_{k_1}, \dots, \xi_{k_m})^T$ 服从 $N(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}})$, 其中 $\tilde{\boldsymbol{\mu}} = (\mu_{k_1}, \dots, \mu_{k_m})$, 且 $\tilde{\boldsymbol{\Sigma}}$ 为 $\boldsymbol{\Sigma}$ 的 (k_1, \dots, k_m) 子式.

定理 3.3. 而 $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ 是 $\boldsymbol{\xi}$ 的期望和协方差矩阵.

证明. 第一个显然, 第二个我们知道由柯西不等式知协方差存在, 则

$$E\xi_j \xi_k = \frac{1}{i^2} \left. \frac{\partial^2 f(t_1, \dots, t_n)}{\partial t_j \partial t_k} \right|_{t_i=0} = \sigma_{jk} + \mu_j \mu_k,$$

则 $E(\xi_j - \mu_j)(\xi_k - \mu_k) = E\xi_j \xi_k - \mu_j \mu_k = \sigma_{jk}$, 成立. □

定理 3.4. ξ_1, \dots, ξ_n 相互独立当且仅当两两不相关.

证明. 若两两不相关, 则 $j \neq k$ 有 $\sigma_{jk} = E(\xi_j - E\xi_j)(\xi_k - E\xi_k) = 0$, 则特征函数

$$f(t_1, \dots, t_n) = \exp\left(i\boldsymbol{\mu}^T \mathbf{t} - \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right) = \prod_{k=1}^n f_{\xi_k}(t_k),$$

故命题成立. □

定理 3.5. 若 $\boldsymbol{\xi} = (\boldsymbol{\xi}_1^T, \boldsymbol{\xi}_2^T)^T$, 且 $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$ 为其协方差矩阵, 其中 $\boldsymbol{\Sigma}_{ii}$ 对应 $\boldsymbol{\xi}_i$. 则

$\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ 相互独立当且仅当 $\boldsymbol{\Sigma}_{12} = \mathbf{O}$.

证明. 若 $\boldsymbol{\Sigma}_{21}^T = \boldsymbol{\Sigma}_{12} = \mathbf{O}$, 设 $\mathbf{t} = (\mathbf{t}_1^T, \mathbf{t}_2^T)^T$, 则

$$f(\mathbf{t}) = \exp\left(i\boldsymbol{\mu}_1^T \mathbf{t}_1 - \frac{1}{2}\mathbf{t}_1^T \boldsymbol{\Sigma}_{11} \mathbf{t}_1\right) \exp\left(i\boldsymbol{\mu}_2^T \mathbf{t}_2 - \frac{1}{2}\mathbf{t}_2^T \boldsymbol{\Sigma}_{22} \mathbf{t}_2\right) = f_{\boldsymbol{\xi}_1}(\mathbf{t}_1) f_{\boldsymbol{\xi}_2}(\mathbf{t}_2),$$

故命题成立. □

定理 3.6. $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$ 服从 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 当且仅当任一个线性组合 $\zeta = \sum_{j=1}^n l_j \xi_j$ 服从 $N(\sum_{j=1}^n l_j \mu_j, \sum_{j,k} l_j l_k \sigma_{jk})$.

证明. 一方面, 取 $\mathbf{t} = u\mathbf{1}$, 代入特征函数得到结论; 另一方面, 特征函数与之前相同, 只需要让 $u = 1$. □

定理 3.7. $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$ 服从 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 取 mn 矩阵 \mathbf{C} , 则 $\boldsymbol{\eta} = \mathbf{C}\boldsymbol{\xi}$ 服从 $N(\mathbf{C}\boldsymbol{\mu}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^T)$.

证明. 考虑特征函数, 显然. □

命题 3.7. 若 $\boldsymbol{\xi} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 其中 $\boldsymbol{\Sigma}$ 是 n 阶正定矩阵, 则 $(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \sim \chi_n^2$.

证明. 存在可逆矩阵 \mathbf{L} 使得 $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^T$, 则

$$(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) = (\boldsymbol{\xi} - \boldsymbol{\mu})^T (\mathbf{L}\mathbf{L}^T)^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) = \boldsymbol{\eta}^T \boldsymbol{\eta},$$

其中 $\boldsymbol{\eta} = \mathbf{L}^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu})$ 为均值为 0 的 n 维正态分布, 且协方差矩阵为 $\mathbf{L}^{-1}\boldsymbol{\Sigma}(\mathbf{L}^T)^{-1} = \mathbf{I}$, 故其分量相互独立, 那么

$$(\boldsymbol{\xi} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) = \boldsymbol{\eta}^T \boldsymbol{\eta} \sim \chi_n^2,$$

得到结论. □

定理 3.8. 若 $\boldsymbol{\xi} = (\boldsymbol{\xi}_1^T, \boldsymbol{\xi}_2^T)^T$, 且 $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$ 为其协方差矩阵, 其中 $\boldsymbol{\Sigma}_{ii}$ 对应 $\boldsymbol{\xi}_i$,

且 $E\boldsymbol{\xi}_1 = \boldsymbol{\mu}_1, E\boldsymbol{\xi}_2 = \boldsymbol{\mu}_2$, 则在给定 $\boldsymbol{\xi}_1 = \mathbf{x}_1$ 下的 $\boldsymbol{\xi}_2$ 的条件分布 $N(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})$.

证明. 考虑变换 $\begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{pmatrix}$, 计算得到 $E\boldsymbol{\eta}_1 = \boldsymbol{\mu}_1, E\boldsymbol{\eta}_2 = \boldsymbol{\mu}_2 -$

$\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1$ 且 $D\boldsymbol{\eta}_1 = \boldsymbol{\Sigma}_{11}, D\boldsymbol{\eta}_2 = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$. 不难得知变换之后的二者协方差矩阵为零, 故 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ 独立, 且变换的 Jacobian 是零, 则 $P_{\boldsymbol{\xi}}(\mathbf{x}_1, \mathbf{x}_2) = P_{\boldsymbol{\eta}_1}(\mathbf{y}_1)P_{\boldsymbol{\eta}_2}(\mathbf{y}_2)$, 故 $p(\mathbf{x}_2|\boldsymbol{\xi}_1 = \mathbf{x}_1) = p_{\boldsymbol{\eta}_2}(\mathbf{y}_2) = p_{\boldsymbol{\eta}_2}(\mathbf{x}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{x}_1)$, 则得到结论. □

4 其他零散分布

4.1 Cauchy 分布

定义 4.1. 密度函数为 $q(y) = \frac{1}{\pi} \frac{1}{1+y^2}$, $y \in \mathbb{R}$ 的分布为 *Cauchy* 分布.

命题 4.1 (Cauchy 分布的数字特征). (1) 数学期望不存在;
(2) 一般的 *Cauchy* 分布 $\frac{1}{\pi} \frac{\lambda}{\lambda^2+(x-\mu)^2}$, $\lambda > 0$ 的特征函数为 $e^{i\mu t - \lambda|t|}$.

4.2 F 分布

定义 4.2. 相互独立的 χ^2 分布 ξ, η 自由度分别为 m, n , 则随机变量 $\beta = \frac{n\xi}{m\eta}$ 的密度函数为

$$f(x; m, n) = \begin{cases} \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \frac{(\frac{m}{n})^{m/2} x^{m/2-1}}{(1 + \frac{m}{n}x)^{(m+n)/2}}, & x > 0; \\ 0, & x \leq 0, \end{cases}$$

称为 *F* 分布.

4.3 Rayleigh 分布

定义 4.3. 密度函数为

$$R(r) = \begin{cases} re^{-r^2/2}, & r \geq 0; \\ 0, & r < 0, \end{cases}$$

称为 *Rayleigh* 分布.

5 常见反例

例 1 (相互独立和两两独立). (a) [两两独立不蕴含相互独立] 在正四面体上, 三个面分别写上 A, B, C , 在第四个面上同时写上 A, B, C , 判断投一次后底面颜色. 发现 $P(A) = P(B) = P(C) = \frac{1}{2}$, 且 $P(AB) = P(BC) = P(AC) = \frac{1}{4}$, 但 $P(ABC) = \frac{1}{4} \neq \frac{1}{8}$;

(b) 在正八面体上, 在第 $1, 2, 3, 4$ 面写上 A , 第 $1, 2, 3, 5$ 面写上 B , 第 $1, 6, 7, 8$ 面写上 C , 则 $P(A) = P(B) = P(C) = \frac{1}{2}$, 且 $P(ABC) = \frac{1}{8} = P(A)P(B)P(C)$, 但 $P(AB) = \frac{3}{8} \neq \frac{1}{4} = P(A)P(B)$.

例 2 (相同随机变量组成不同函数也独立). 若 ξ, η 相互独立, 且 $\xi, \eta \sim N(0, 1)$, 则 $\rho = \sqrt{\xi^2 + \eta^2}, \phi = \arctan \frac{\eta}{\xi}$ 相互独立.

这是因为考虑极坐标 $x = r \cos t, y = r \sin t$, 则 $r = \sqrt{x^2 + y^2}, t = \arctan \frac{y}{x}$, 则 (ρ, ϕ) 的密度函数 $q(r, t) = \frac{1}{2\pi} r e^{-r^2/2}, r \geq 0, 0 \leq t \leq 2\pi$, 则 ρ 服从 *Rayleigh* 分布而 ϕ 服从均匀分布. 独立.

例 3 (随机变量函数的独立性). 若 (ξ, η) 服从联合密度函数 $p(x, y) = \begin{cases} \frac{1+xy}{4}, & |x| < 1, |y| < 1, \\ 0, & \text{其他;} \end{cases}$,

则 ξ, η 不独立, 但 ξ^2, η^2 相互独立.

例 4 (不相关不蕴含独立). 设 θ 服从 $[0, 2\pi]$ 的均匀分布, 设 $\xi = \cos \theta, \eta = \cos(\theta + a)$. 若 $a = \frac{\pi}{2}$, 则不相关, 但 $\xi^2 + \eta^2 = 1$, 故不独立.

若 (ξ, η) 联合密度函数为 $\frac{1}{\pi} \chi_{x^2+y^2 \leq 1}$, 则不相关也不独立.

例 5 (和的特征函数是特征函数的积不蕴含独立). 设 ξ 服从 *Cauchy* 分布且 $\eta = \xi$, 则其特征函数满足 $f_{\xi+\eta}(t) = f_{\xi}(t)f_{\eta}(t)$, 但二者显然不独立.

例 6 (多元场合满足分布函数三条性质但不是分布函数). 考虑 $F(x, y) = \chi_{x+y>0}$, 则其满足非降性, 左连续性, 且 $\lim_{x(y) \rightarrow \infty} F = 0, \lim_{x \rightarrow +\infty, y \rightarrow +\infty} F = 1$, 但注意到 $P(0 \leq \xi \leq 1, 0 \leq \eta \leq 1) = F(1, 1) - F(1, 0) - F(0, 1) + F(0, 0) = -1$, 这不可能. 但在一元场合下, 满足三条性质一定是分布函数.