

考研概率论

枫聆

2021 年 11 月 21 日

目录

1	概率运算	3
1.1	翻译事件要准确	3
1.2	贝叶斯的应用	3
2	常用分布	4
2.1	参数确定	4
3	随机变量函数	4
3.1	连续性判定	4
4	正态分布	5
4.1	线性运算	5
5	期望和方差	6
5.1	复杂随机变量函数	6
5.2	随机变量乘积	6
5.3	拟合	6
5.4	线性相关	7
6	参数估计	7
6.1	均匀分布的参数估计	7
7	三大分布	8
7.1	三大分布重要性质的应用	8

8	假设检验	8
8.1	犯错的概率	8

概率运算

翻译事件要准确

Example 1.1. 某种产品由自动生产线进行生成, 一旦出现不合格品就立即对其进行调整, 经过调整后生产出的产品为不合格的概率为 0.1, 求两次调整之间至少产生 3 件产品的概率.

hints 设 $A_i = \{\text{一次调整之后生产的第 } i \text{ 件为次品}\}$, $B = \{\text{两次调整之间至少产生 3 件产品}\}$, 那么

$$P(B) = 1 - P(A_1) - P(A_2) = 1 - 0.1 - 0.1 * 0.9 = 0.81.$$

贝叶斯的应用

Example 1.2. 假设有两箱同种零件: 第一箱内装有 50 件, 其中 10 件一等品; 第二箱内装有 30 件, 其中 18 件一等品. 现从两箱中随意挑选一箱, 然后从箱中随机取两个零件, 试求在第一次取出的零件是一等品的条件下, 第二次取出一等品的概率.

hints 设事件 A 为选择第一个箱子, 事件 B_1 为第一次取出一等品, 事件 B_2 为第二次取出一等品. 这里要求的是一个条件概率 $P(B_2|B_1)$, 首先我们用贝叶斯公式分别计算 $P(A|B_1)$ 和 $P(\bar{A}|B_1)$, 即

$$P(A|B_1) = \frac{P(A)P(B_1|A)}{P(A)P(B_1|A) + P(\bar{A})P(B_1|\bar{A})} = \frac{\frac{10}{50}}{\frac{10}{50} + \frac{18}{30}} = \frac{1}{4},$$

因此 $P(\bar{A}|B_1) = \frac{3}{4}$. 于是

$$P(B_2|B_1) = P(B_2|AB_1)P(A|B_1) + P(B_2|\bar{A}B_1)P(\bar{A}|B_1) = \frac{9}{49} \times \frac{1}{4} + \frac{17}{29} \times \frac{3}{4}$$

常用分布

参数确定

Example 2.1. 设随机变量 $X \sim U(a, b)$, 已知 $P\{-2 < X < 0\} = \frac{1}{4}$ 和 $P\{1 < X < 3\} = \frac{1}{2}$, 求 a, b .

hints 由已知条件, 我们知道 $[a, b]$ 和区间 $[-2, 0]$ 及 $[1, 3]$ 都是有重叠部分的, 因此

$$\begin{cases} a < 0 \\ b > 1 \end{cases}$$

由此我们知道 $[0, 1]$ 完全躺在 $[a, b]$ 里面的. (1 考虑由于 $P\{1 < X < 3\} = \frac{1}{2}$, 那么

$$P\{-2 < X < 1\} \leq \frac{1}{2} \Rightarrow P\{0 < X < 1\} \leq \frac{1}{4}.$$

考虑 $|[0, 1]| = \frac{1}{2}|[1, 2]|$, 那么

$$P\{0 < X < 1\} \geq \frac{1}{4}.$$

综上 $P\{0 < X < 1\} = \frac{1}{4}$. 因此 $a = -1, b = 3$.

随机变量函数

连续性判定

Example 3.1. 设随机变量 X 与 Y 相互独立, X 服从参数为 λ 的指数分布, Y 的分布律为 $P\{Y = -1\} = \frac{1}{2}, P\{Y = 1\} = \frac{1}{2}$. 判定 $Z = X + Y$ 的分布函数 $F_Z(z)$ 的连续性.

hints 求出 $F_Z(z)$ 来判断是下策! 这里要结合分布函数的性质来做就比较简单, 如果 $F_Z(z)$ 有间断点 a , 那么它是左间断的, 即 $P\{Z = a\} = F(a) - F(a-0) > 0$. 因此我们来求 $P\{Z = a\}$,

$$P\{Z = a\} = P\{X \leq a+1\}P\{Y = -1\} + P\{X \leq a-1\}P\{Y = 1\} = \frac{1}{2}[P\{X \leq a+1\} + P\{X \leq a-1\}] = 0$$

最后一个等式成立条件是因为 X 是连续的.

正态分布

线性运算

Example 4.1. 设 X_1, X_2 是两个独立的正态分布 (μ, σ^2) , 证明: $X_1 - X_2$ 和 $X_1 + X_2$ 也是独立的.

[hints](#)

$$\text{Cov}(X_1 - X_2, X_1 + X_2) = D(X_1) - D(X_2) = 0,$$

且 $X_1 - X_2 \sim (0, 2\sigma^2), X_1 + X_2 \sim (2\mu, 2\sigma^2)$.

期望和方差

复杂随机变量函数

Example 5.1. 相互独立的随机变量 X_1 和 X_2 均服从正态分布 $N(0, \frac{1}{2})$, 求 $D(|X_1 - X_2|)$.

hints 这里求期望不需要计算出 $|X_1 - X_2|$ 的概率分布, 只需要确定 $X_1 - X_2$ 概率分布即可, 设 $Z = X_1 - X_2$, 那么显然有 $Z \sim N(0, 1)$. 首先求 $E(|X_1 - X_2|)$

$$E(|X_1 - X_2|) = \int_{-\infty}^{+\infty} |z| f_z(z) dz = 2 \int_0^{+\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \frac{\sqrt{2}}{\sqrt{\pi}}.$$

再来求 $D(|X_1 - X_2|)$

$$D(|X_1 - X_2|) = D(|Z|) = E(Z^2) - E^2(|Z|) = 1 - \frac{2}{\pi}.$$

Example 5.2. 设随机变量 X 的分布函数为 $F(x) = 0.4\Phi(\frac{x-5}{2}) + 0.6\Phi(\frac{x+1}{3})$, 其中 $\Phi(x)$ 为标准正态分布的分布函数, 求 $E(X)$.

hints 常规思路是先求出 $f(x)$, 再积分

$$\int_{-\infty}^{+\infty} \frac{x}{5} f(\frac{x-5}{2}) dx + \int_{-\infty}^{+\infty} \frac{x}{5} f(\frac{x+1}{3}) dx = \int_{-\infty}^{+\infty} (\frac{4}{5}t + 2) f(t) dt + \int_{-\infty}^{+\infty} (\frac{9}{5}t - \frac{3}{5}) f(t) dt = \frac{7}{5}.$$

也可以这样思考 $\Phi(\frac{x-5}{2}) \sim N(5, 4)$, $\Phi(\frac{x+1}{3}) \sim N(-1, 9)$, 因此 $E(X) = \frac{2}{5} \cdot 5 - 1 \cdot \frac{3}{5} = \frac{7}{5}$.

随机变量乘积

Example 5.3. 设随机变量 X 服从标准正态分布 $N(0, 1)$, 求 $E[(X - 2)^2 e^{2X}]$.

hints 这里要用求随机变量函数期望的公式.

Example 5.4. 设随机变量 X, Y 不相关, 且 $E(X) = 2, E(Y) = 1, D(X) = 3$, 求 $E[X(X + Y - 2)]$.

hints

$$\text{Cov}(X, X + Y - 2) = E[X(X + Y - 2)] - E(X)E(X + Y - 2).$$

拟合

Example 5.5. 设随机变量 $X_1, X_2, \dots, X_{2n}, \dots$ 相互独立服从指数分布 $E(\lambda)$, 记 $Z_i = X_{2i} - X_{2i-1}, i = 1, 2, 3, \dots$, 则 $\sum_{i=1}^n Z_i$ 近似服从正态分布, 求其参数.

hints 不要去构造尝试构造 $\sum_{i=1}^n Z_i$, 题目已经告诉你是正态分布了, 那么直接求其期望和方差即可.

线性相关

Example 5.6. 设随机变量 $X \sim N(0, 1), Y \sim N(1, 4)$, 且相关系数 $\rho_{XY} = 1$, 求 $Y = aX + b$.

hints 这个题非常经典, $\rho_{XY} = 1$ 表示 X 和 Y 线性相关, 来求它们之间的线性表达式.

方法 1. 直接祭关键表达式

$$E\{[t(X - E(X)) + (Y - E(Y))]^2\} = t^2 D(X) + 2t \text{Cov}(X, Y) + D(Y),$$

使得上式左边等于 0, 求出 t . 因此

$$t(X - E(X)) + (Y - E(Y)) = 0,$$

带入 $t = 2$, 即可求得 $Y = 2X + 1$.

方法 2. 由 $Y = aX + b$, 那么 $E(Y) = E(aX + b)$, 求出 $b = 1$. 再由

$$\text{Cov}(X, Y) = \text{Cov}(X, aX + b) = a \text{Cov}(X, X) = 2,$$

求出 $a = 2$.

参数估计

均匀分布的参数估计

Example 6.1. 设总体 X 的概率密度为

$$f(x; \theta) = \begin{cases} \frac{1}{1-\theta}, & \theta \leq x \leq 1, \\ 0, & \text{其他.} \end{cases}$$

hints 那么似然函数为

$$\begin{aligned} L(x_1, x_2, \dots, x_n; \theta) &= \begin{cases} \frac{1}{(1-\theta)^n}, & \theta \leq x_i \leq 1, i = 1, 2, \dots, n \\ 0, & \text{其他.} \end{cases} \\ &= \begin{cases} \frac{1}{(1-\theta)^n}, & \theta \leq \min\{x_1, x_2, \dots, x_n\} \\ 0, & \text{其他.} \end{cases} \end{aligned}$$

因此当 $\theta = \min\{x_1, x_2, \dots, x_n\}$ 时, $L(x_1, x_2, \dots, x_n; \theta)$ 取最大值. 一定理解这个表示形式, 因为拿到手你不太好表示这个似然函数的.

三大分布

三大分布重要性质的应用

Example 7.1. 设 X_1, X_2, \dots, X_n 来自正态总体 $N(\mu, \sigma^2)$ 的简单随机样本, 求 $E \left\{ \sum_{i=1}^n X_i \left[\sum_{j=1}^n (nX_j - \sum_{k=1}^n X_k)^2 \right] \right\}$.

hints \bar{X} 和 S^2 线性无关.

Example 7.2. 设 X_1, X_2, \dots, X_n 和 Y_1, Y_2, \dots, Y_n 分别来自正态总体 $N(\mu, \sigma^2)$ 两个相互独立简单随机样本, 设它们样本方差分别为 S_X^2 和 S_Y^2 , 求统计量 $T = (n-1)(S_X^2 + S_Y^2)$ 的方差 DT .

hints $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.

Example 7.3. 设随机变量 $X \sim F(n, n)$, $p_1 = P\{X \geq 1\}$, $p_2 = P\{X \leq 1\}$, 证明: $p_1 = p_2$.

hints $X \sim F(n_1, n_2) \Rightarrow \frac{1}{X} \sim F(n_2, n_1)$.

假设检验

犯错的概率

Example 8.1. 设 X_1, X_2, \dots, X_{16} 是来自总体 $N(\mu, 4)$ 的简单样本, 考虑假设检验问题 $H_0: \mu \leq 10, H_1: \mu > 10$. 若该检验问题的拒绝域为 $W = \{\bar{X} > 11\}$, 其中 $\bar{X} = \frac{1}{16} \sum_{i=1}^{16} X_i$, 则 $\mu = 11.5$ 时, 求该检验犯第二类错误的概率.

hints 给定 $\mu = 11.5$, 那么原假设 H_0 为假的, 因此按照定义

$$\beta = P\{\bar{X} \leq 11\} = P\left\{\frac{\bar{X} - 11.5}{\frac{1}{2}} \leq -1\right\} = \Phi(-1) = 1 - \Phi(1).$$