

概率论与数理统计

胡喜平

2019 年 12 月 28 日

1 随机事件与概率

1.1 条件概率

有事件 A 、 B ， $P(B) \neq 0$ ，则 A 在 B 发生条件下的概率

$$P(A|B) = \frac{P(AB)}{P(B)}$$

事件 B 是发生的前提条件，也有

$$P(AB) = P(A|B)P(B) = P(B|A)P(A)$$

全概率公式 (p9)

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i) \quad (1)$$

贝叶斯公式

$$P(A_i|B)P(B) = P(B|A_i)P(A_i)$$

将公式 1 代入

$$P(A_i|B) \sum_{i=1}^n P(B|A_i)P(A_i) = P(B|A_i)P(A_i)$$

最终得到

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

1.2 随机变量的分布

1.2.1 离散型

X	x_1	x_2	\cdots	x_i	\cdots
P	p_1	p_2	\cdots	p_i	\cdots

一维分布列 其中 $\sum_i p_i = 1$ ，分布函数 $F(x) = P(X \leq x)$

X \ Y	Y			
	y_1	y_2	y_3	
x_1	p_{11}	p_{12}	p_{13}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	p_{23}	$p_{2\cdot}$
	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	

二维分布列 其中 $\sum_{i,j} p_{ij} = 1$

二项分布 $X \sim B(n, p)$ 共实验 n 次, 成功的概率是 p , 分布列为

$$P(X = k) = C_n^k p^k (1-p)^{n-k}$$

分布列先增后减, 中心项 $P(X = [(n+1)p])$ 最大

当样本数量足够多的时候, 二项分布可以近似认为是和它方差和均值相等的正态分布。

泊松分布 $X \sim P(\lambda)$ 当二项分布的 n 很大, p 很小时, 可近似为泊松分布

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

先增后减, 中心项为 $P(X = [\lambda])$

超几何分布 (略) p19

几何分布 分布列为

$$P(X = k) = pq^{k-1}$$

计算分布函数要用到等比数列的求和公式

1.2.2 连续型随机变量

一维随机变量概率密度与分布函数

$$F(x) = \int_{-\infty}^x f(t)dt = P(X < x)$$

$$F(+\infty) = \int_{-\infty}^{+\infty} f(x)dt = 1$$

二维随机变量概率密度与分布函数

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v)dudv$$

$$F(+\infty, +\infty) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(u, v)dudv = 1$$

边缘概率密度

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

二维随机变量求分布函数需要画图分区域讨论，详见：p39 例 3.4

均匀分布 $X \sim U(a, b)$

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

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}$$

指数分布 $X \sim E(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

一维正态分布 $N \sim (\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\left(\frac{x-\mu}{\sigma}\right)^2}{2}} d\left(\frac{x-\mu}{\sigma}\right) \quad (2)$$

定义标准正态分布 $N \sim (0, 1)$ 的分布函数为

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

则方程 2 可化为

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = P(X < x)$$

二维正态分布 $N \sim (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \frac{1}{\sqrt{2\pi}\sigma_2} \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{2\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{2\sigma_2^2} \right]}$$

其中 $X \sim N(\mu_1, \sigma_1)$ ，即边缘密度

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

详见 p41

1.2.3 随机变量函数的分布

分布函数法 (详见 p28) 对于离散型随机变量, 根据分布列求概率就可以。对于连续型, 先根据概率求分布函数, 再对分布函数求导, 得到概率密度。

多维连续情形 (详见 p45)

- 和的分布
- 商的分布
- 最大值最小值分布

条件分布 根据公式 1.1, 相似的, 定义

$$f(y|x) = \frac{f(x, y)}{f_X(x)}$$

为 $X = x$ 时 Y 的条件密度

1.2.4 随机变量的独立性

离散型

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)P(X_2 = x_2)$$

连续型

$$f(x, y) = f_X(x)f_Y(y)$$

2 数字特征

2.1 数学期望与方差

2.1.1 离散型

$$EX = \sum_{i=1}^{\infty} x_i p_i$$
$$DX = E(X - EX)^2 = EX^2 - E(X^2)$$

2.1.2 连续型

$$EX = \int_{-\infty}^{+\infty} x f(x) dx$$
$$DX = \int_{-\infty}^{+\infty} (x - EX)^2 f(x) dx$$

2.2 不同分布的数学期望与方差

2.2.1 离散型

X	EX	DX
$B(n, p)$	np	$np(1-p)$
$P(\lambda)$	λ	λ

2.2.2 连续型

X	EX	DX
$N(\mu, \sigma^2)$	μ	σ^2
$E(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$U(a, b)$	$\frac{b-a}{2}$	$\frac{(b-a)^2}{12}$

2.3 切比雪夫不等式

$$P(|X - EX| \geq \varepsilon) \leq \frac{DX}{\varepsilon^2} \quad (3)$$

2.4 协方差

$$Cov(X, Y) = E[(X - EX)(Y - EY)] = E(XY) - EXEY$$

协方差的性质有

$$Cov(aX, bY) = abCov(X, Y)$$

$$Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$$

特别地

$$DX = Cov(X, X)$$

2.5 相关系数

$$\rho_{xy} = \frac{Cov(X, Y)}{\sqrt{DX}\sqrt{DY}}$$

2.6 矩

k 阶原点矩 $\nu_k = EX^k$

k 阶原点绝对矩 $\alpha_k = E|X|^k$

k 阶中心矩 $\mu_k = E(X - EX)^k$

k 阶原点绝对矩 $\beta_k = E|X - EX|^k$

3 统计量的分布

3.1 统计量

均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

k 阶原点矩 $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$

k 阶中心矩 $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k$

3.2 分布

χ^2 分布 $X_n \sim N(0, 1)$, $X_1 + \cdots + X_n$ 满足自由度为 n 的 χ^2 分布, 该分布具有可加性。

t 分布 $X_n \sim N(0, 1)$, $Y \sim \chi^2(n)$, 则 $T = \frac{X}{\sqrt{Y/n}} \sim t(n)$

F 分布 $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 则 $F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$

3.2.1 正态总体样本均值与方差满足的分布

均值 $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

推论 1 $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$

推论 2 $\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$

方差 $\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} = (n-1) \frac{S^2}{\sigma^2} \sim \chi^2(n-1)$

4 参数估计

4.1 矩估计

- 求 EX 与 EX^2
- $A_1 = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$
- $A_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$
- $A_2 - A_1^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \tilde{S}^2$
- 用上面的式子解出分布中的参数, 消去 A_1 和 A_2

例子详见 p101

4.2 极大似然估计

- 根据概率密度用连乘的形式写出似然方程
- 对似然方程取对数, 求导判断参数取什么值的时候, 似然函数最大
- 如果参数取多个值的时候似然函数都能最大, 则取最大值作为极大似然估计

详见 p103

4.3 无偏估计

$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 是 σ^2 的无偏估计, $\tilde{S}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 是 σ^2 的有偏估计

无偏估计 $E(\hat{\theta}) = \theta$

一致估计 $\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| \leq \varepsilon) = 1$ 注意运用切比雪夫不等式 3

4.4 区间估计 (置信水平 $1 - \alpha$)

4.4.1 均值 μ 的区间估计

已知 σ^2 , 由章节 3.2.1, 推论 1: $\left(\bar{X} - \frac{\sigma}{\sqrt{n}} u_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha/2} \right)$

未知 σ^2 , 由章节 3.2.1, 推论 2: $\left(\bar{X} - \frac{S}{\sqrt{n}} u_{\alpha/2}, \bar{X} + \frac{S}{\sqrt{n}} u_{\alpha/2} \right)$

4.4.2 方差 σ 的区间估计

未知 μ , 由章节 3.2.1, 方差满足的分布: $\left(\frac{\sqrt{n-1}S}{\sqrt{\chi_{\alpha/2}^2(n-1)}}, \frac{\sqrt{n-1}S}{\sqrt{\chi_{1-\alpha/2}^2(n-1)}} \right)$

4.4.3 单侧置信区间

单侧置信下限 $\underline{\theta}$ $P(\theta > \underline{\theta}) = 1 - \alpha$

单侧置信上限 $\bar{\theta}$ $P(\theta < \bar{\theta}) = 1 - \alpha$