1 基础概念

- 互不相容 (互斥) $A \cap B = \varnothing$
- 逆事件 (对立事件) $A \cup B = S \wedge A \cap B = \varnothing$
- 德摩根律 $\overline{A \cup B} = \bar{A} \cap \bar{B}$
- 概率
  - $0 \leqslant P(A) \leqslant 1$
  - 规范性 $P(S) = 1$ 性质扩展 $P(A) \leqslant P(S) = 1$
  - 可列可加性质, 有限可加性质
  - $A_1, A_2, \ldots$ 两两互不相容 $\Rightarrow f_n(A_1 \cup A_2 \cup \ldots) = f_n(A_1) + f_n(A_2) + \ldots$
  - $P(\varnothing) = 0$
  - $A \subset B \Rightarrow P(B - A) = P(B) - P(A), P(B) > P(A)$
  - $P(\bar{A}) = 1 - P(A)$
  - $\underline{\text{加法公式 } P(A \cup B) = P(A) + P(B) - P(AB)}$

组合数 $r \leqslant a, C_r^a = \binom{a}{r} = \dfrac{a(a-1)\ldots(a-r+1)}{r!}$

1.1 条件概率

条件概率 $P(B|A) = \dfrac{P(AB)}{P(A)}$

乘法定理 $P(AB) = P(B \mid A)P(A)$, $P(ABC) = P(C \mid AB)P(B \mid A)P(A)$

全概率公式 $P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n)$

贝叶斯公式 $P(B_i \mid A) = \frac{P(B_i A)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}$

1.2 独立性

$$P(AB) = P(A)P(B) \Leftrightarrow A, B\text{相互独立}$$

$$A, B\text{独立} \Leftrightarrow A, \bar{B}\text{独立}$$

多事件独立

$$\left.\begin{array}{l} P(AB)=P(A)P(B), \\ P(BC)=P(B)P(C), \\ P(AC)=P(A)P(C), \\ P(ABC)=P(A)P(B)P(C). \end{array}\right\} \Leftrightarrow A,B,C\text{相互独立}$$

2 随机变量及其分布

分布函数 F 设 X 是一个随机变量, x 是任意实数, 函数 $F(x) = P\{X \leqslant x\}, -\infty < x < \infty$

2.1 离散型随机变量

1. $(0-1)$ 分布

   随机变量只可能取 0 和 1 两个值,

   分布律 $P\{X=k\} = p^k(1-p)^{1-k}, k=0,1(0<p<1)$

   分布律表格

X	0	1
p_k	$1-p$	p

2. 伯努利试验, 二项分布: $X \sim b(n,p)$

   伯努利试验 只有两个可能结果的实验: A 及 $\bar{A}$

   设 $P(A) = p\ (0<p<1)$, 则 $P(\bar{A}) = 1-p$

   n 重伯努利试验 将**伯努力实验**重复 n 次

   二项分布 随机变量 X 表示 n 重伯努利试验中 A 发生的次数, X 服从参数为 n, p 的二项分布

   分布律 $P\{X=k\} = \begin{pmatrix} n \\ k \end{pmatrix} p^k q^{n-k}$

   数学期望 $E(X) = p$

   方差 $D(X) = p(1-p)$

3. 泊松分布: $X \sim \pi(\lambda)$

   随机变量可能取的值为 $0,1,2,\ldots$

   分布律 $P\{X=k\} = \frac{\lambda^k e^{-\lambda}}{k!}, k=0,1,2,\ldots$

   数学期望 $E(X) = \lambda$

   方差 $D(X) = E(X^2) - [E(X)]^2 = [E(X(X-1)) + E(X)] - \lambda^2 = \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k e^{-\lambda}}{k!} + \lambda - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

其中 $\lambda > 0$ 是常数, 则称 X 服从参数为 λ 的泊松分布

定理 2.1.1. 泊松定理

$$\lim_{n->\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k e^{-\lambda}}{k!},\lambda=np_n$$

由此可知, n 很大, p 很小时，二项分布可以由此近似

$$\binom{n}{k}p_n^k(1-p_n)^{n-k}\approx\frac{\lambda^k e^{-\lambda}}{k!},\lambda=np_n$$

例 2.1.1 证明泊松定理

解 由 $p_n=\frac{\lambda}{n}$ 可得:

$$\begin{aligned}\binom{n}{k}p_n^k(1-p_n)^{n-k}&=\frac{n(n-1)\dots(n-k+1)}{k!}(\frac{\lambda}{n})^k(1-\frac{\lambda}{n})^{n-k}\\&=\frac{\lambda^k}{k!}[1\cdot\frac{n-1}{n}\dots\frac{n-k+1}{n}](1-\frac{\lambda}{n})^{n-k}\end{aligned}$$

当 $n\to\infty$ 时: $[1\cdot\frac{n-1}{n}\dots\frac{n-k+1}{n}]\to 1$, $(1-\frac{\lambda}{n})^{n-k}\to e^{-\lambda}$

即 $\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k e^{-\lambda}}{k!}$

2.2 连续型随机变量

对与随机变量 X 分布函数 $F(X)$, 存在非负函数 $f(x)$, 使

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

成了, X 为**连续性随机变量**, $f(x)$ 为**概率密度函数**, 简称**概率密度**, 概率密度函数可以大于 1

1. 均匀分布: $X\sim U(a,b)$

   概率密度 $f(x)=\begin{cases}\dfrac{1}{b-a} & a<x<b\\ 0 & 其他\end{cases}$

分布函数 $F(x)=\begin{cases}0 & x<a\\ \dfrac{x-a}{b-a} & a\leqslant x<b\\ 1 & x\geqslant b\end{cases}$

数学期望 $E(X)=\frac{a+b}{2}$

方差 $D(X)=E(X^2)-[E(X)]^2=\int_a^b x^2\frac{1}{b-a}dx-(\frac{a+b}{2})^2=\frac{(b-a)^2}{12}$

2. 指数分布

概率密度 $f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & \text{其他}\end{cases}$

分布函数 $F(x)=\begin{cases}1-e^{-\frac{x}{\theta}} & x>0\\ 0 & \text{其他}\end{cases}$

无记忆性 $P\{X>s+t\mid X>s\}=P\{X>t\}$

3. 正态分布: $X\sim N(\mu,\sigma^2)$

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

分布函数 $F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dt$

数学期望 μ

方差 σ^2

4. 标准正态分布: $X\sim N(0,1^2)$

概率密度 $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}$

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

$1-\Phi(x)=\Phi(-x)$

转化为标准正态分布 若 $X\sim N(\mu,\sigma^2)$, $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$

例 2.2.1 $X\sim N(90,0.5^2)$, 求 $P\{X<89\}$

解

$$\begin{aligned}P\{X<89\}&=P\{\frac{X-90}{0.5}<\frac{89-90}{05}\}\\&=\Phi(\frac{89-90}{0.5})\\&=0.0228\end{aligned}$$

上 α 分位点 $P\{X>z_\alpha\}=\alpha, 0<\alpha<1$

2.3 随机变量的函数的分布

例 2.3.1 随机变量 X 具有概率密度 $f_X(x), -\infty < x < \infty$, 求 $Y = X^2$ 的概率密度

解

$$\begin{aligned} F_Y(y) &= P\{Y \leqslant y\} = P\{X^2 \leqslant y\} \\ &= P\{-\sqrt{y} \leqslant X \leqslant \sqrt{Y}\} \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})], & y > 0 \\ 0, & y \leqslant 0 \end{cases}$$

3 多维随机变量

联合分布函数 $F(x,y)$
边缘分布函数 $F_X(x,y)=F_X(x,\infty)$
边缘分布概率 $f_X(x)=\int_{\infty}^{\infty}f(x,y)dy$
边缘分布律 $p_{i\cdot}=\sum_{j=0}^{\infty}p_{ij}$
条件分布律 $P\{X=x_i\mid Y=y_i\}=\frac{p_{ij}}{p_{\cdot j}}$
条件概率密度 $P_{X|Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)}$

例 3.0.1 设数 X 在区间 $(0,1)$ 上随机地取值, 当观察到 $X=x(0<x<1)$ 时, 数 Y 在区间 $(x,1)$ 上随机地取值, 求 Y 的概率密度函数 $f_Y(y)$
解

$$f_X(x)=\begin{cases}1, & 0<x<1\\ 0, & \text{其他}\end{cases}$$

$$f_{Y|X}(y\mid x)=\begin{cases}\frac{1}{1-x}, & x<y<1\\ 0, & \text{其他}\end{cases}$$

$$f(x,y)=f_{Y|X}(y\mid x)f_X(x)=\begin{cases}\frac{1}{1-x}, & x<y<1\\ 0, & \text{其他}\end{cases}$$

$$\begin{aligned}f_Y(y)&=\int_{-\infty}^{\infty}f(x,y)dx\\&=\begin{cases}\int_0^y\frac{1}{1-x}\frac{1}{1-x}dx=-ln(1-y), & 0<y<1\\ 0, & \text{其他}\end{cases}\end{aligned}$$

3.1 两个随机变量相互独立

$$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}P\{Y\leqslant y\}$$
$$F(x,y)=F_X(x)F_Y(y)$$

3.2 两个随机变量的函数的分布

1. $Z = X + Y$

$$
\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f(z-y, y)dy \\
&= \int_{-\infty}^{\infty} f(x, z-x)dx
\end{aligned}
$$

2. $Z = \frac{Y}{X}$

$$
f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f(x, xz)dx
$$

3. $Z = XY$

$$
f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x, \frac{z}{x})dx
$$

4. $max\{X, Y\}, min\{X, Y\}$

   当 X, Y 独立时:

$$
F_{max}(z) = F_X(z)F_Y(z)
$$

$$
F_{min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)]
$$

<u>补充? 例题 P87</u>

4 随机变量的数字特征

切比雪夫不等式 $P\{|X-E(X)| \geqslant \varepsilon\} \leqslant \frac{D(X)}{\varepsilon^2}$, $P\{|X-E(X)| < \varepsilon\} \geqslant 1-\frac{D(X)}{\varepsilon^2}$

4.1 数学期望

数学期望 $E(X)=\sum_{k=1}^{\infty} x_k p_k$, $E(X)=\int_{-\infty}^{\infty} x f(x) dx$

$E(Y)=E(g(X))=\sum_{k=1}^{\infty} g(x_k) pk$, $E(Y)=E(g(X))=\int_{-\infty}^{\infty} g(x_k) pk$

4.1.1 数学期望性质

1. $E(C)=C$
2. $E(CX)=CE(X)$
3. $E(X+Y)=E(X)+E(Y)$
4. $E(XY)=E(X)E(Y)$ (X, Y 相互独立)

4.2 方差

方差 $D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2 p_k$, $D(X)=\int_{-\infty}^{\infty}[x-E(X)]^2 f(x) dx$, $D(X)=E(X^2)-E^2(X)$

4.2.1 方差性质

1. $D(C)=0$
2. $D(CX)=C^2 D(X)$
3. $D(X+Y)=D(X)+D(Y)+Cov(X,Y)$
4. $D(X)=0 \Leftrightarrow P\{X=E(X)\}=1$

4.3 协方差

协方差 $Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$

1. $Cov(X,Y) = Cov(Y,X)$
2. $Cov(X,X) = D(X)$

X, Y 独立时 $Cov(X,Y) = 0$

相关系数 $\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$

5 大数定律和中心极限定理

5.1 大数定律

弱大数定理 (辛钦大数定理) 对于 $X_1, X_2, \ldots$ 相互独立且服从同一分布,

$\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu| < \varepsilon\} = 1, \forall\varepsilon > 0, E(X_k) = \mu$

例 5.1.1 证明弱大数定理

解 $E(\frac{1}{n}\sum_{k=1}^{n} X_k) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) = \frac{1}{n}n\mu = \mu$

$D(\frac{1}{n}\sum_{k=1}^{n} X_k) = \frac{1}{n}\sum_{k=1}^{n} D(X_k) = \frac{1}{n}n\sigma^2 = \sigma^2$

$1 \geqslant P\{|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu| < \varepsilon\} \geqslant 1 - \frac{\sigma^2}{\varepsilon^2}$

$n \to \infty, P\{|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu| < \varepsilon\} \to 1$

伯努力大数定理 $\lim_{n\to\infty} P\{|\frac{f_A}{n} - p| < \varepsilon\} = 1$

5.2 中心极限定理

独立同分布的中心极限定理 $\lim_{n\to\infty} P\{\frac{\sum_{k=1}^{\infty} X_k - n\mu}{\sqrt{n}\sigma} \leqslant x\} = \Phi(x)$

李雅普诺夫定理 <u>补充? P122</u>

棣莫弗—拉普拉斯定理 设随机变量 $\eta_n \sim b(n, p)$

$\lim_{n\to\infty} P\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leqslant x\} = \Phi(x)$

6 样本及抽样分析

样本平均差 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

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

样本标准差 S

样本 k 阶 (原点) 矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

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

经验分布函数 $F_n(x)$

χ^2 **分布:** $\chi^2 \sim \chi^2(n)$ $X_1, X_2, \ldots$ 是来自总体 $N(0,1)$ 的样本，则称统计量 $\chi^2 = X_1^2 + X_2^2 + \ldots + X_n^2$, 服从自由度为 n 的 χ^2 分布

数学期望: $E(\chi^2) = n$

方差: $D(\chi^2) = n^2$

χ^2 分布的可加性: $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$

6. 设总体 $X \sim b(1,p)$，$X_1, X_2, \cdots, X_n$ 是来自 X 的样本.

(1) 求 $(X_1, X_2, \cdots, X_n)$ 的分布律.

(2) 求 $\sum\limits_{i=1}^{n} X_i$ 的分布律.

(3) 求 $E(\overline{X}), D(\overline{X}), E(S^2)$.

解 (1) 因 $X_1, X_2, \cdots, X_n$ 相互独立，且有 $X_i \sim b(1,p)$，$i=1,2,\cdots,n$，即 X_i 具有分布律 $P\{X_i = x_i\} = p^{x_i}(1-p)^{1-x_i}$，$x_i = 0,1$，因此 $(X_1, X_2, \cdots, X_n)$ 的分布律为

$$
\begin{aligned}
& P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} \\
&= \prod_{i=1}^{n} P\{X_i = x_i\} \\
&= \prod_{i=1}^{n} [p^{x_i}(1-p)^{1-x_i}] = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}.
\end{aligned}
$$

(2) 因 $X_1, X_2, \cdots, X_n$ 相互独立，且有 $X_i \sim b(1,p)$，$i=1,2,\cdots,n$，故 $\sum\limits_{i=1}^{n} X_i \sim b(n,p)$，其分布律为

$$
P\left\{\sum_{i=1}^{n} X_i = k\right\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0,1,2,\cdots,n.
$$

(3) 由于总体 $X \sim b(1,p)$，$E(X) = p$，$D(X) = p(1-p)$，故有

$$
E(\overline{X}) = p, D(\overline{X}) = \frac{p(1-p)}{n}, E(S^2) = D(X) = p(1-p).
$$

图 6.1: 例题

7 参数估计

补充？P149

8 假设检验

原假设和备择假设

1. 双边检验: $H_0 : \Theta = \Theta_0$, $H_1 : \Theta \neq \Theta_0$
2. 右边检验: $H_0 : \Theta \leqslant \Theta_0$, $H_1 : \Theta > \Theta_0$
3. 左边检验: $H_0 : \Theta \geqslant \Theta_0$, $H_1 : \Theta < \Theta_0$

例 8.0.1 右边检验 $\mu_0 = -0.545, \sigma = 0.008, \bar{x} = -0.535, n = 5, a = 0.05$

解 原假设和备择假设: $H_0 : \mu \leqslant \mu_0 = -0.545, H_1 : \mu > \mu_0$

拒绝域: $z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = 2.7951 \geqslant z_{0.05} = 1.645$

即在显著性水平 $a = 0.05$ 下拒绝 H_0

索引

边缘分布概率, 7
边缘分布函数, 7
边缘分布律, 7
伯努力大数定理, 11
伯努力试验, 3

独立同分布的中心极限定理, 11

方差, 9
分布函数, 3
分布律, 3
分布律表格, 3

概率密度函数, 4

假设检验, 15

离散性随机变量
- (0 - 1) 分布, 3
- 二项分布, 3
- 泊松分布, 3

联合分布函数, 7
连续性随机变量
- 标准正态分布, 5
- 均匀分布, 4
- 正态分布, 5
- 指数分布, 5

n 重伯努力试验, 3

泊松定理, 4

切比雪夫不等式, 9

弱大数定理, 11

数学期望, 9

条件分布律, 7
条件概率密度, 7

协方差, 9
辛钦大数定理, 11

重点
- 泊松定理证明, 4
- 弱大数定理证明, 11