概率论基础讲义

管华

2017-09-02 更新

目录

第一章　随机事件与概率

1.1　基本概念

1.2　事件的运算与关系

1.3　概率的定义与基本性质

1.4　概率基本公式

1.4.1　条件概率

在事件 A 发生的情况下 B 发生的概率

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

1.4.2　乘法公式

$$\begin{aligned} P(AB) &= P(A)P(B|A) \\ P(A_1A_2\cdots A_n) &= P(A_1)\,P(A_2|A_1)\,P(A_3|A_1A_2)\cdots P(A_n|A_1\cdots A_{n-1}) \end{aligned}$$

1.5 事件的独立性

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

1.6 全概率公式与贝叶斯公式

1.6.1 完备事件组

$$A_1 + A_2 + \cdots + A_n = \Omega$$

1.6.2 全概率公式

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

1.6.3 贝叶斯公式

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum P(A_i)P(B|A_i)}$$

1.7　三种常见的概型

1.7.1　古典概型

1.7.2　几何概型

1.7.3　n 重贝努利试验

$B_k = \{n\text{次试验中}A\text{出现}k\text{次}\}$

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

第二章　一维随机变量及其分布

2.1　随机变量、分布函数及性质

1. 随机变量 $X(\omega)$，也就是概率 $p = X(\omega)$

2. 分布函数

2.2　离散型

$P\{X = x_i\} = p_i$

表 2.1 离散型随机变量 X 的分布律

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

$$F(x) = P\{X \leqslant x\} = \sum_{x_i \leqslant x} p_i$$

2.3　连续型

$$F(x)=\int_{-\infty}^{x} f(x)\,\mathrm{d}x$$

2.4　常见的

2.4.1　离散型

2.4.1.1　二项分布

定义 2.4.1 ($X\sim B(n,p)$)

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

2.4.1.2　Poisson 分布

定义 2.4.2 ($X\sim \pi(k)$)

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

2.4.1.3　几何分布

定义 2.4.3 ($X\sim G(p)$)

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

2.4.1.4　超几何分布

定义 2.4.4 ($X\sim H(N,M,n)$)

$$P\{X=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}$$

2.4.2　连续型

2.4.2.1　均匀分布 $X \sim U(a,b)$

$$f(x) = \begin{cases} \dfrac{1}{b-a} \\ 0, \text{other} \end{cases} \Bigg| \; F(x) = \begin{cases} 0 \\ \dfrac{x-a}{b-a} \\ 1 \end{cases}$$

2.4.2.2　指数分布 $X \sim E(\lambda)$

$$f(x) = \begin{cases} \lambda \exp\left(-\lambda x\right) \\ 0 \end{cases} \Bigg| \; F(x) = \begin{cases} 1 - \exp\left(-\lambda x\right) \\ 0 \end{cases}$$

2.4.2.3　正态分布 $X \sim N(\mu, \sigma^2)$

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

$$\varphi = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{2.4.1}$$

$$\varphi(x) = \int_{-\infty}^{x} \varphi(t)\,\mathrm{d}t \tag{2.4.2}$$

2.5　函数分布

$$Y = \varphi(X)$$

2.5.1　X 为离散型

$$P\{Y=y_j\}=P\{\varphi(X)=y_j\}=\sum_{\varphi(x_i)=y_j}P\{X=x_i\}$$

2.5.2　X 为连续型

2.5.2.1　Y 为离散型

先求出 Y 的可能取值，在通过 X 的概率分布求出 Y 的可能取值对应的概率。

2.5.2.2　Y 为连续型

$$\begin{cases} F_Y(y)=P\{Y\leqslant y\}=P\{\varphi(X)\leqslant y\}=\int_{\varphi(x)\leqslant y}f_X(x)\,\mathrm{d}x, \quad f_y(y)=F_Y'(y) \\ f_Y(y)=f_X\left[h(y)\right]|h'|, \quad y=\varphi(x)x=h(y) \end{cases}$$

第三章　二维随机变量及其分布

3.1　联合分布与边缘分布

定义 3.1.1 (二维随机变量)

$$(X, Y)$$

定义 3.1.2 (联合分布函数)

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

定义 3.1.3

$$\begin{cases} F_X(x) = P\{X \leqslant x\} \\ F_Y(y) = P\{Y \leqslant y\} \end{cases}$$

3.2　离散型

3.2.1　联合分布律

$$P\{X = x_i, Y = y_j\} = p_{ij}$$

3.2.2　边缘分布律

$$\begin{cases} P\{X = x_i\} = \sum_{j=1}^{n} p_{ij} = p_{i\cdot} \\ P\{Y = y_i\} = \sum_{i=1}^{m} p_{ij} = p_{\cdot j} \end{cases}$$

3.3　连续型

定义 3.3.1 (联合密度函数 $f(x,y)$)

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

3.3.1　边缘分布

$$\begin{cases} f_X(x) = \int_{-\infty}^{+\infty} dy, \quad \text{边际密度函数} \\ F_X(x) = \int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y, \quad \text{边缘分布函数} \end{cases}$$

3.4　条件分布

$$\begin{cases} P\{Y = y_j | X = x_i\} = \dfrac{p_{ij}}{p_{i\cdot}} \\ f_{\frac{Y}{X}}(y|x) = \dfrac{f(x,y)}{f_X(x)} \end{cases}$$

3.5 随机变量的独立性

等价于

$$F(x,y)=F_X(x)F_Y(y)$$

$$\begin{cases} p_{ij}=p_{i:}\times p_{:j} \\ f(x,y)=f_X(x)f_Y(y) \end{cases}$$

3.6 常见的

3.7 函数分布

$U=\varphi(X,Y)$ U 的分布函数为 $F_U(u)=P\{U\leqslant u\}=P\{\varphi(X,Y)\leqslant u\}$

3.7.1 离散型

$$P\{U=u_k\}=\sum_{\varphi(x_i,y_j)=u_k}p_{ij}$$

3.7.2 连续型

$$F_U(u)=P\{U\leqslant u\}=P\{\varphi(X,Y)\leqslant u\}=\iint_{\varphi(x,y)\leqslant u}f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

3.7.3 常见的

3.7.3.1 max/min

$$U = \min\{X, Y\} \tag{3.7.1}$$

$$V = \max\{X, Y\} \tag{3.7.2}$$

$$F_U(u) = 1 - \big[1 - F_x(u)\big] \times \big[1 - F_Y(u)\big] \tag{3.7.3}$$

$$F_V(v) = F_X(v) \times F_Y(v) \tag{3.7.4}$$

3.7.3.2 $X + Y$

$$U = X + Y$$

$$f_U(u) = \int_{-\infty}^{+\infty} f(x, u - x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f(u - y, y) dy$$

特别地,

$$X \sim B(m, p), Y \sim B(n, p) \qquad U \sim B(m + n, p) \tag{3.7.5}$$

$$X \sim P(\lambda_1), Y \sim P(\lambda_2) \qquad U \sim P(\lambda_1 + \lambda_2) \tag{3.7.6}$$

$$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2) \qquad U \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \tag{3.7.7}$$

第四章　随机变量的数字特征

4.1　数学期望

4.1.1　概念

4.1.1.1　一维

$Y = \phi(X)$

1. 离散型

$$E\,X = \sum x_i p_i$$
$$E\,Y = \sum \phi(x_i) p_i$$

2. 连续型

$$E\,X = \int_{-\infty}^{\infty} x f(x) \, \mathrm{d}x$$
$$E\,Y = \int_{-\infty}^{\infty} \phi(x) f(x) \, \mathrm{d}x$$

$f(x)\ \mathrm{d}x = x_i$

4.1.1.2　二维

$Z = \phi(X, Y)$

$$
E\,Z = \sum\sum \phi(x_i, y_i) p_{ij}
$$
$$
E\,Z = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \phi(x,y) f(x,y)\,\mathrm{d}y
$$

4.1.2 性质

1. $E(aX \pm bY) = aEX \pm bEY$

2. 独立 $\Rightarrow E\,XY = (E\,X)(E\,Y)$ &uad 反之不对!

4.2 方差

4.2.1 性质

1. $D(aX + b) = a^2 DX$

2. 独立 $\Rightarrow D(aX + bY) = a^2 DX + b^2 DY$

4.2.2 常见的

常见随机变量的数学期望及方差

4.3 协方差、相关系数

4.3.1 协方差

4.3.1.1 定义

$$
\operatorname{cov}(X, Y) = E(X - EX)(Y - EY)
$$

表 4.1 常见随机变量的数学期望及方差

分布	分布函数/密度函数	期望	方差
$X \sim B(n,p)$	$P\{X=k\} = C_n^k\, p^k(1-p)^{n-k}$	np	$np(1-p)$
$X \sim \pi(\lambda)$	$P\{X=k\} = \frac{\lambda^k i}{k!}\exp(-\lambda)$	λ	λ
$X \sim G(p)$	$P\{X=k\} = p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$X \sim U(a,b)$	$f(x) = \dfrac{1}{b-a}$ $F(x) = \dfrac{x-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$X \sim E(\lambda)$	$f(x) = \lambda\exp(-\lambda x)$ $F(x) = 1-\exp(-\lambda x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$X \sim N(\mu,\sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$	μ	σ^2

4.3.1.2　计算公式

$$\operatorname{cov}(X,Y) = EXY - EX \cdot EY$$

4.3.1.3　性质

$$\operatorname{cov}(X,X) = DX \tag{4.3.1}$$

$$\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X) \tag{4.3.2}$$

$$\operatorname{cov}(aX+bY,Z) = a\operatorname{cov}(X,Z)b\operatorname{cov}(Y,Z) \tag{4.3.3}$$

$$\operatorname{cov}(aX,bY) = ab\operatorname{cov}(X,Y) \tag{4.3.4}$$

$$\text{不独立} \Rightarrow D(aX+bY) = a^2DX + b^2DY + 2ab\operatorname{cov}(X,Y) \Rightarrow \tag{4.3.5}$$

4.3.2 相关系数

4.3.2.1 定义

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{DX}\sqrt{DY}}$$

4.3.2.2 性质

$|\rho_{XY}| \leqslant 1$

$$\rho_{XY} \leqslant 1 \quad \Longleftrightarrow \quad EXY = EX \times EY \tag{4.3.6}$$

$$\rho_{XY} = 1 \quad \Longleftrightarrow \quad Y = aX + b,\ a > 0 \tag{4.3.7}$$

$$\rho_{XY} = -1 \quad \Longleftrightarrow \quad Y = aX + b,\ a < 0 \tag{4.3.8}$$

(4.3.9)

第五章　大数定理与中心极限定理

5.1　切比雪夫不等式

$$P\{|X-\mu| \geqslant \epsilon\} \leqslant \frac{DX}{\epsilon^2} \tag{5.1.1}$$

$$P\{|X-\mu| < \epsilon\} \geqslant 1-\frac{DX}{\epsilon^2} \tag{5.1.2}$$

5.2　大数定律

- 切比雪夫大数定律
- 独立同分布大数定律
- 辛钦大数定律

定理 5.2.1

所有的大数定律都是指随机变量组在独立的条件下，若干个随笔建立的平均值依概率收敛到其数学期望的平均值，即

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{独立}} E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \mu \tag{5.2.1}$$

更进一步，

$$\frac{1}{n}\sum_{i=1}^{n}X_i^k \xrightarrow{\text{独立}} E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^k\right)=\mu \tag{5.2.2}$$

5.3　中心极限定理

第六章　数理统计的基本概念与基本原理

6.1　基本概念

定义 6.1.1 (样本均值)

$$\bar{X} = \frac{1}{n}\sum X_i \Rightarrow EX = \mu$$

定义 6.1.2 (样本方差)

$$S^2 = \frac{1}{n-1}\sum \left(X_i - \bar{X}\right)^2 \Rightarrow DX = \sigma^2$$

定义 6.1.3 (*k* 阶原点矩)

$$A_k = \frac{1}{n-1}\sum X_i^k \Rightarrow EX^k$$

6.2 三个重要的抽样分布

6.2.1 χ^2 分布

定义 6.2.1 (χ^2 分布)
条件:

1. $X_1, \cdots, X_n$ 相互独立
2. 都服从标准正态分布

$$Z = X_1^2 + \cdots + X_n^2$$

称 Z 服从 χ^2 分布，记为 $X \sim \chi^2(n)$.

定理 6.2.1

$$E\chi^2(n) = n \tag{6.2.1}$$

$$D\chi^2(n) = 2n \tag{6.2.2}$$

若 $X \sim \chi^(m), Y \sim \chi^2(n)$, 则

$$X + Y = \chi^2(m + n)$$

6.2.2 t 分布

定义 6.2.2
条件:

1. $X \sim N(0, 1)$
2. $Y \sim \chi^2(n)$
3. 独立

$$Z = \frac{X}{\sqrt{Y/n}}$$

称 Z 服从 t 分布，记为 $Z \sim t(n)$.

定理 6.2.2

1. t 分布近似服从于正态分布
2. $EZ = 0$
3. $DZ = \frac{n}{n-2}$

6.2.3　F 分布

定义 6.2.3

条件:

1. $X \sim \chi^2(m)$
2. $Y \sim \chi^2(n)$

$$Z = \frac{X/m}{Y/n}$$

记为 $F \sim (m, n)$.

定理 6.2.3

1. $\frac{1}{F} \sim F(n, m)$
2. $F_{1-\alpha/2}(m, n) = \frac{1}{F_{\alpha/2}(n,m)}$

6.3　正态总体下常用的抽样分布

$X_1, \ldots, X_n$ 是来自总体 X 的样本，则

1.
$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

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

3.
$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 \sim \chi^2(n-1)$$

4.
$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$$

5. $\bar{X}$ 和 S^2 相互独立

6.
$$ES^2 = \sigma^2$$

第七章　参数估计与假设检验

7.1　参数估计的种类

- 点估计
- 区间估计

7.2　点估计

7.2.1　矩估计

总体 $X \sim f(x,\theta)$，但参数 θ 未知，需要对参数 θ 进行估计。步骤 1. 取样 $X_1, X_2, \cdots, X_n$ 2. 计算样本均值 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，根据大数定律有

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \frac{1}{n}\sum_{i=1}^{n} EX_i = EX$$

3. 令 $\bar{X} = EX$，在 EX 的结果中包含参数 $\theta \Rightarrow \hat{\theta}$

若含有两个参数 θ_1, θ_2 ，由大数定律知

1. $\bar{X} \to EX, A_2 = \frac{1}{n}\sum X_i^2 \to \frac{1}{n}\sum EX_i^2 = EX^2$
2. 令 $\bar{X} = EX, A_2 = EX^2$ 或令 $\frac{1}{n}\sum \left(X_i - \bar{X}\right)^2 = DX \Rightarrow \theta_1, \theta_2$ 的估计。

7.2.2　极大似然估计

设 $X \sim f(x,\theta)$ 或 $X \sim f(x,\theta_1,\theta_2)$ 步骤：1.

1. 对离散型

$$L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}P\{X=x_i\}$$

2. 对连续型

$$L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}f(x_i;\theta)$$

2. 求

$$\ln L(x_1,x_2,\cdots,x_n;\theta)$$

3.

- 对一个参数令

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L=0$$

- 对两个参数令

$$\frac{\partial}{\partial\theta_i}\ln L=0,\quad i=1,2$$

4. 解似然方程或方程组

7.3　区间估计

1. 设 $X \sim N(\mu,\sigma^2)$,　$X_1,X_2,\cdots,X_n$ 是来自总体样本 X 的一个样本。若存在两个统计量 $\hat{\theta}_1,\hat{\theta}_2$，使得 $P\{\hat{\theta}_1<\theta<\hat{\theta}_1\}=1-\alpha$，称 $\left(\hat{\theta}_1,\hat{\theta}_2\right)$ 为参数 θ 的置信度为 $1-\alpha$ 的置信区间。

- σ^2 已知，μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}},\bar{X}+\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\right)$$

区间长度与样本无关

- σ^2 未知，μ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\bar{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$$

  区间长度与样本有关

- μ 未知，σ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)S^2}{\chi^2_{\frac{1-\alpha}{2}}(n-1)}\right)$$

- μ 已知，σ 的置信度为 $1-\alpha$ 的置信区间

2. 三个评价标准

- 参数估计的无偏性设 $\hat{\theta}$ 为 θ 的一个估计量，若 $E\hat{\theta}=\theta$，称 $\hat{\theta}$ 为 θ 的无偏估计量
- 参数估计的有效性设 $\hat{\theta}_1,\hat{\theta}_2$ 都是 θ 的无偏估计量，若 $D(\hat{\theta}_1)<D(\hat{\theta}_2)$，称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效
- 一致性设 $\hat{\theta}$ 为 θ 的一个估计量，若 $\hat{\theta}\xrightarrow{P}\theta$，称 $\hat{\theta}$ 为 θ 的一致估计