# 目录

# 一、 概率论基本概念

## 1. 事件的关系与运算

### 1.1 随机试验

对随机现象的观察、记录、实验统称为随机试验. 它具有以下特性:

(1) 可以在相同条件下重复进行;

(2) 事先知道可能出现的结果 $(\geq 2)$;

(3) 进行试验前并不知道哪个试验结果会发生.

### 1.2 样本空间

**定义 1** 随机试验 $E$ 的所有结果构成的集合称为 $E$ 的样本空间, 记为 $S = \{e\}$, 称 $S$ 中的元素 $e$ 为样本点, 一个元素的单点集称为基本事件.

### 1.3 随机事件

一般我们称 $S$ 的子集 $A$ 为 $E$ 的随机事件 $A$, 简称事件 $A$. 当且仅当 $A$ 所包含的一个样本点发生称事件 $A$ 发生.

### 1.4 运算与关系

和、差、积、逆、包含、不相容

## 2. 概率定义与性质

### 2.1 频率

**定义 2** $f_n(A) = \frac{n_A}{n}$, 在 $n$ 次中 $A$ 发生 $n_A$ 次.

性质:

(1) $0 \leq f_n(A) \leq 1$

(2) $f_n(S) = 1$

(3) 若 $A_1, A_2, \ldots, A_k$ 两两不相容, 则

$$f_n\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} f_n(A_i)$$

### 2.2 概率

**定义 3** 对样本空间 $S$ 中任一事件 $A$, 定义一个实数 $P(A)$, 满足以下三条公理:

*(1)* 非负性: $P(A) \geq 0$

*(2)* 规范性: $P(S) = 1$

*(3) 可列可加性：若 $A_1, A_2, \dots, A_k, \dots$ 两两不相容，则*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

称 $P(A)$ 为事件 $A$ 的概率.

性质：

(1) $P(\phi) = 0$

(2) $A_1, A_2, \dots, A_n$, $A_iA_j = \phi, i \neq j$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

(3) $P(A) = 1 - P(\bar{A})$

(4) $P(B - A) = P(B) - P(AB) = P(B\bar{A})$

(5) $P(A \cup B) = P(A) + P(B) - P(AB)$

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_iA_j) \\ &+ \sum_{1 \le i < j < k \le n} P(A_iA_jA_k) + \cdots \\ &+ (-1)^{n-1} P(A_1A_2 \cdots A_n) \end{aligned}$$

(6) $P(AB) = P(A)P(B|A)$ 若 $P(A) > 0$

$$P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cdots A_{n-1})$$

## 3. 等可能概型

**定义 4** 若试验 $E$ 满足：

*(1) $S$ 中样本点有限 (有限性)*

*(2) 出现每一样本点的概率相等 (等可能性)*

称这种试验为等可能概型 (或古典概型).

$$P(A) = \frac{A\text{所包含的样本点数}}{S\text{中的样本点数}}$$

## 4. 条件概率

**定义 5**

$$P(B|A) = \frac{P(AB)}{P(A)}, \; P(A) > 0$$

### 4.1 全概率公式与 Bayes 公式

**定义 6** 称 $B_1, B_2, \dots, B_n$ 为 $S$ 的一个划分，若

*(1) 不漏 $B_1 \cup B_2 \cup \cdots \cup B_n = S$*

*(2) 不重 $B_iB_j = \phi, \; i \neq j$*

**定理 1** 设 $B_1, B_2, \dots, B_n$ 为样本空间 $S$ 的一个划分，$P(B_i) > 0, \; i = 1, 2, \dots, n$，则称

$$P(A) = \sum_{j=1}^{n} P(B_j) \cdot P(A|B_j)$$

为全概率公式.

**定理 2** 接全概率公式的条件，且 $P(A) > 0$，则

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j) \cdot P(A|B_j)}$$

称为 *Bayes* 公式.

## 5. 事件独立性与独立试验

**定义 7** 设 $A, B$ 为两随机事件，如果 $P(AB) = P(A)P(B)$，则称 $A$ 与 $B$ 相互独立. 若 $P(A) \neq 0, P(B) \neq 0$，则

$$\begin{aligned} P(AB) = P(A)P(B) &\Longleftrightarrow P(A|B) = P(A) \\ &\Longleftrightarrow P(B|A) = P(B) \end{aligned}$$

**定义 8** 设 $A_1, A_2, \dots, A_n$ 为 $n$ 个随机时间，若对 $2 \le k \le n$，均有

$$P(A_{i_1}A_{i_2} \cdots A_{i_n}) = \prod_{j=1}^{k} P(A_{i_j})$$

则称 $A_1, A_2, \dots, A_n$ 相互独立.

## 二、 随机变量及分布函数

### 1. 随机变量

**定义 9** 设随机试验的样本空间 $S=\{e\}$，若 $X=X(e)$ 为定义在样本空间 $S$ 上的实值单值函数，则称 $X=X(e)$ 为随机变量.

### 2. 分布函数

**定义 10** 随机变量 $X$ 的分布函数:

$$\forall x \in R,\, F(x)=P(X \leq x)$$

性质:

(1) $0 \leq F(x) \leq 1$

(2) $F(x)$ 单调不减，且 $F(-\infty)=0, F(+\infty)=1$

(3) $F(x)$ 右连续，即 $F(x+0)=F(x)$

(4) $F(x)-F(x-0)=P(X=x)$

### 3. 离散型随机变量

**定义 11** 取值至多可数的随机变量为离散型的随机变量.

$$\text{概率分布律性质}: p \geq 0, \sum_{i=1}^{\infty} p_i=1$$

$$\text{分布函数}: F(x)=\sum_{x_k \leq x} p_k$$

3.1 0-1 分布 $X \sim B(1,p)$

| X | 0 | 1 |
|---|---|---|
| P | $1-p$ | $p$ |

$$F(x)=P\{X \leq x\}=\begin{cases}0 & x<0 \\ 1-p & 0 \leq x<1 \\ 1 & x \geq 1\end{cases}$$

$$E(X)=p,\, Var(X)=p(1-p)$$

3.2 n 重贝努里试验与二项分布 $X \sim B(n,p)$

进行 $n$ 次独立重复观测，每次 $A$ 或 $\bar{A}$ 发生，$P(A)=p$，$X$ 表示 $A$ 发生的次数.

$$P(X=k)=C_n^k p^k(1-p)^{n-k},\, k=0,1,\ldots,n$$

$$E(x)=np,\, Var(X)=np(1-p)$$

3.3 泊松分布 $X \sim P(\lambda)$

$$P(X=k)=\frac{e^{-\lambda}\lambda^k}{k!},\, k=0,1,\ldots$$

$$E(X)=Var(X)=\lambda$$

泊松定理:

**定理 3** 当 $n>10, p<0.1$ 时,

$$C_n^k p^k(1-p)^{n-k} \approx \frac{e^{-\lambda}\lambda^k}{k!},\, \lambda=np$$

### 4. 连续型随机变量

**定义 12** 对于随机变量 $X$ 的分布函数 $F(x)$，若对 $\forall x \in R$，$\exists f(x) \geq 0$,

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

则称 $X$ 为连续型随机变量，其中 $f(x)$ 称为 $X$ 的概率密度函数，简称密度函数.

性质:

(1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x=1$

(3) $P(a \leq X \leq b)=P(a<X<b)=F(b)-F(a)=\int_b^a f(x)\,\mathrm{d}x$

4.1 均匀分布 $X \sim U(a,b)$

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

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

$$E(X)=\frac{a+b}{2},\, Var(X)=\frac{(b-a)^2}{12}$$

$$a<c<d<b,\, P(c \leq X \leq d)=\frac{d-c}{b-a}$$

4.2 正态分布 $X \sim N(\mu, \sigma^2)$

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

$$E(X) = \mu, Var(X) = \sigma^2$$

标准正态分布 $N(0,1)$:

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

性质:

(1) $\Phi(-x) + \Phi(x) = 1$, $\Phi(0) = \frac{1}{2}$

(2) $\Phi(z_\alpha) = 1 - \alpha$, $z_{1-\alpha} = -z_\alpha$

(3) 当 $X \sim N(\mu, \sigma^2)$ 时: 令 $t = \frac{x-\mu}{\sigma} \sim N(0,1)$,

$$P(x \le b) = \Phi(\frac{b-\mu}{\sigma})$$

(4) $aX + b \sim N(a\mu + b, a^2\sigma^2)$

4.3 指数分布 $X \sim Exp(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$$
$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

$$E(X) = \frac{1}{\lambda}, Var(X) = \frac{1}{\lambda^2}$$

无记忆性:

$$s > 0, t > 0, P(X > s + t | X > s) = P(X > t) = e^{-\lambda t}$$

**5. 随机变量函数的分布**

已知 $X$ 的概率分布, $Y = g(X)$, 求 $Y$ 的概率分布.

(1) 若 Y 为离散型随机变量, 则先写出 $Y$ 的可能取值: $y_1, y_2, \dots, y_j, \dots$, 再找出 $(Y = y_i)$ 的等价事件 $(X \in D_j)$, 得 $P(Y = y_i) = P(X \in D_j)$.

(2) 若 Y 为连续型随机变量, 则先写出 $Y$ 的概率分布函数: $F_Y(y) = P(Y \le y)$, 找出 $(Y \le y)$ 的等价事件 $(X \in D_y)$, 得 $F_Y(y) = P(X \in D_y)$; 再求出 $Y$ 的概率密度函数 $f_Y(y)$.

**定义 13** *设 $X \sim f_X(x)$, $-\infty < x < \infty$, $y = g(x)$ 是严增(减)可微函数, 即 $g'(x) > 0$(或 $g'(x) < 0$). $Y = g(X)$, 则 $Y$ 具有概率密度函数为*

$$f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)| & \alpha < y < \beta \\ 0 & otherwise \end{cases}$$

*这里 $(\alpha, \beta)$ 是 $Y$ 的取值范围, $h$ 是 $g$ 的反函数, $h(y) = x \iff y = g(x)$.*

# 三、 多元随机变量及其分布

## 1. 二元离散型随机变量

### 1.1 联合概率分布

**定义 14** 若二元随机变量 $(X,Y)$ 全部可能取到的不同值是有限对或可列无限对, 则称 $(X,Y)$ 是离散型随机变量.

联合概率分布律:

$$
\begin{aligned}
(x_i, y_j), i, j = 1, 2, \ldots \\
P(X = x_i, Y = y_j) = p_{ij}, i, j = 1, 2, \ldots
\end{aligned}
$$

性质:

(1) $p_{ij} \geq 0,\ i, j = 1, 2, \ldots$

(2) $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$

### 1.2 边际分布律

$$
\begin{aligned}
P(Y = y_j) =& P(X < +\infty, Y = y_j) \\
=& \sum_{i=1}^{\infty} p_{ij} \equiv p_{\bullet j}\ j = 1, 2, \ldots \\
P(X = x_i) =& P(X = x_i, Y < +\infty) \\
=& \sum_{j=1}^{\infty} p_{ij} \equiv p_{i\bullet}\ i = 1, 2, \ldots
\end{aligned}
$$

### 1.3 条件分布律

(1) 对 $Y = y_j$, 若 $P(Y = y_j) > 0$

$$
\begin{aligned}
P(X = x_i | Y = y_j) =& \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} \\
=& \frac{p_{ij}}{p_{\bullet j}}\ i = 1, 2, \ldots
\end{aligned}
$$

(2) 对 $X = x_i$, 若 $P(X = x_i) > 0$

$$
\begin{aligned}
P(Y = y_j | X = x_i) =& \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} \\
=& \frac{p_{ij}}{p_{i\bullet}}\ j = 1, 2, \ldots
\end{aligned}
$$

## 2. 随机变量 (X,Y) 的联合分布函数

$$
\begin{aligned}
F(x, y) =& P\{(X \leq x) \cup (Y \leq y)\} \\
\equiv& P(X \leq x, Y \leq y)
\end{aligned}
$$

### 2.1 边际分布函数

$$
\begin{aligned}
F_X(x) = F(x, +\infty) = P(X \leq x, Y < +\infty) \\
F_Y(y) = F(+\infty, y) = P(X < +\infty, Y \leq y)
\end{aligned}
$$

### 2.2 条件分布函数

$$
\begin{aligned}
F_{X|Y}(x|y) = P(X \leq x | Y = y) = \frac{P(X \leq x, Y = y)}{P(Y = y)} \\
F_{Y|X}(y|x) = P(Y \leq y | X = x) = \frac{P(X = x, Y \leq y)}{P(X = x)}
\end{aligned}
$$

## 3. 二元连续型随机变量

### 3.1 联合概率密度函数

**定义 15** $F(x, y)$, $\exists f(x, y) > 0, \forall x, y,$

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

性质:

(1) $f(x, y) \geq 0$

(2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \, \mathrm{d}x \mathrm{d}y = 1$

(3) 设 $G$ 是平面

$$
P\{(X, Y) \in G\} = \iint_G f(x, y) \, \mathrm{d}x \mathrm{d}y
$$

(4) 在 $(x, y)$ 连续 $f(x, y)$,

$$
\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)
$$

### 3.2 边际概率密度函数

$$
\begin{aligned}
f_X(x) = \int_{-\infty}^{+\infty} f(x, y) \, \mathrm{d}y \\
f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) \, \mathrm{d}x
\end{aligned}
$$

$$
\begin{aligned}
F_X(x) = \int_{-\infty}^{x} f_X(u) \, \mathrm{d}u \\
F_Y(y) = \int_{-\infty}^{y} f_Y(v) \, \mathrm{d}v
\end{aligned}
$$

3.3 条件概率密度函数

$$\text{在 } Y=y \text{ 下}, f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}, f_Y(y)>0$$
$$\text{在 } X=x \text{ 下}, f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}, f_X(x)>0$$

$$F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}\,\mathrm{d}u$$
$$F_{Y|X}(y|x)=\int_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}\,\mathrm{d}v$$

$$P(Y\in D|X=x)=\int_D f_{Y|X}(y|x)\,\mathrm{d}y$$

## 4. 二元随机变量常见分布及性质

4.1 二元均匀分布

$(X,Y)$ 在二维有界 $D$ 上,

$$f(x,y)=\begin{cases}\frac{1}{D\text{面积}} & (x,y)\in D\\ 0 & otherwise\end{cases}$$

若 $D_1\subset D$,

$$P\{(X,Y)\in D_1\}=\frac{D_1\text{面积}}{D\text{面积}}$$

4.2 二元正态 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

性质:

(1)

$$X\sim f_X(x)\sim N(\mu_1,\sigma_1^2)$$
$$Y\sim f_Y(y)\sim N(\mu_2,\sigma_2^2)$$

(2)

在 $\{X=x\}$ 条件下,

$$f_{Y|X}(y|x)\sim N(\mu_2+\rho\frac{\sigma_1}{\sigma_2}(x-\mu_1),(1-\rho^2)\sigma_2^2)$$

在 $\{Y=y\}$ 条件下,

$$f_{X|Y}(x|y)\sim N(\mu_1+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_2),(1-\rho^2)\sigma_1^2)$$

(3) $X$ 与 $Y$ 相互独立 $\Longleftrightarrow \rho=0$

(4) 若 $\rho=0$, $aX+bY+c\sim N(a\mu_1+n\mu_2+c, a^2\sigma_1^2+b^2\sigma_2^2)$

## 5. 事件独立性与随机变量独立性

**定义 16** $\forall x,y\in R$,

$$P(X\le x,Y\le y)=P(X\le x)P(y\le y)$$

*or*

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

$A$ 发生与否都不会改变 $B$ 发生的概率.

$A,B$ 相互独立 $\Longleftrightarrow \bar{A},B$ 相互独立 $\Longleftrightarrow A,\bar{B}$ 相互独立 $\Longleftrightarrow \bar{A},\bar{B}$ 相互独立.

(1) 离散型独立 $\Longleftrightarrow$

$$\forall i,j,\ P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$$

(2) 连续型独立 $\Longleftrightarrow$

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

除去 0 面积后处处成立.

## 6. 二元随机变量函数的分布

6.1 离散型

$(X,Y)$, 有 $P(X=x_i,Y=y_i)=p_{ij}$.

(1) 若 $U=u(x,y),V=v(x,y)$, 先定 $u_i,v_j$, 找 $(U=u_i,V=V_j)=\{(X,Y)\in D\}$.

(2) 若 $Z=z(x,y)$, 先定 $z_i$, 找 $(Z=z_i)=\{(X,Y)\in D\}$.

6.2 连续型

分布函数

$$F_Z(z)=P(g(X,Y)\le z)=\iint_{g(x,y)\le z}f(x,y)\,\mathrm{d}x\mathrm{d}y$$

(1) $Z=X+Y$

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

若 $X,Y$ 独立,

$$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)\,\mathrm{d}x$$
$$=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)\,\mathrm{d}y$$

(2) $M = \max\{X, Y\}$, $N = \min\{X, Y\}$, $X, Y$ 独立

$$
\begin{aligned}
F_M(z) =& P(X \le z, Y \le z) = F_X(z) F_Y(z) \\
F_N(z) =& 1 - P(X > z, Y > z) \\
=& 1 - (1 - F_X(z))(1 - F_Y(z))
\end{aligned}
$$

# 四、 随机变量数字特征

## 1. 数学期望

### 1.1 离散型 X 分布律

$P(X = x_k) = p_k$, $k = 1, 2, \ldots,$

$$E(X) = \sum_{k=1}^{\infty} x_k p_k$$

$$E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k$$

### 1.2 离散型 (X,Y) 联合分布律

$P(X = x_i, Y = y_j) = p_{ij}$, $i, j = 1, 2, \ldots,$

$$E[h(X, Y)] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h(x_i, y_j) p_{ij}$$

### 1.3 连续型 X 概率密度函数

$f(x)$,

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

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

### 1.4 连续型 (X,Y) 联合密度函数

$f(x, y)$

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

### 1.5 性质

(1) $E(aX + bY + c) = aE(X) + bE(Y) + c$

(2) 若 $X, Y$ 相互独立, 则 $E(XY) = E(X)E(Y)$

## 2. 方差

**定义 17**

$$Var(x) = E\left\{[X - E(x)]^2\right\}$$

(1) $Var(X) = E(X^2) - [E(X)]^2$
or $E(X^2) = Var(X) + [E(X)]^2$

(2) 离散型

$$Var(x) = \sum_{i=1}^{\infty} [x_i - E(x)]^2 p_i$$

(3) 连续型

$$Var(x) = \int_{-\infty}^{+\infty} [x - E(x)]^2 f(x) \, \mathrm{d}x$$

### 2.1 性质

(1) $a, c \in \mathbb{C}$, $Var(aX + c) = a^2 Var(X)$.

(2) $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$.
若 $X, Y$ 相互独立, $Var(X + Y) = Var(X) + Var(Y)$.

(3) $Var(X) = 0 \Longleftrightarrow P(X = C) = 1, C = E(X)$.

## 3. 协方差与相关系数

### 3.1 协方差

**定义 18**

$$Cov(X, Y) = E\left\{[X - E(X)][Y - E(Y)]\right\}$$

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)E(Y) \\ \text{or } E(XY) &= Cov(X, Y) + E(X)E(Y) \end{aligned}$$

性质:

(1) $Cov(X, Y) = Cov(Y, X)$

(2) $Cov(X, X) = Var(X)$

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

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

(5) $Cov(aX + bY, cX + dY) = acVar(X) + bdVar(Y)$
$+ (ad + bc)Cov(X, Y)$

(6) $Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y)$
$+ 2abCov(X, Y)$

### 3.2 相关系数

**定义 19**

$$\begin{aligned} \rho_{XY} =& \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} \\ =& Cov\left(\frac{X - E(X)}{\sqrt{Var(X)}} \frac{Y - E(Y)}{\sqrt{Var(Y)}}\right) \end{aligned}$$

可看作标准化后的协方差

性质:

(1) $|\rho_{XY}| \leq 1$.

(2) $|\rho_{XY}| = 1 \Longleftrightarrow \exists a, b, P(Y = a + bX) = 1$
特别的, $\rho_{XY} = 1, b > 0$; $\rho_{XY} = -1, b < 0$.

不相关: $\rho_{XY} = 0$

(1) $X, Y$ 独立 $\Longrightarrow X, Y$ 不相关.

(2) 若 $(X, Y)$ 服从正太分布,
则 $X, Y$ 独立 $\Longleftrightarrow X, Y$ 不相关

## 4. 正态分布性质

若 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

(1) $aX + bY + c \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2)$

(2) $Cov(aX + bY, cX + dY) = ac\sigma_1^2 + bd\sigma_2^2 + (ad + bc)\rho\sigma_1\sigma_2$

(3) $aX + bY, cX + dY$ 相互独立 $\Longleftrightarrow Cov(aX + bY, cX + dY) = 0$

## 5. 上 $\alpha$ 分位数

**定义 20** $X$ 连续随机变量, $\exists F(x), f(x)$,

$$P\left\{X > x_\alpha\right\} = 1 - F(x_\alpha) = \int_{x_\alpha}^{=\infty} f(x) \, \mathrm{d}x = \alpha$$

$x_\alpha$ 为 $X$ 上 $\alpha$ 分位数.

# 五、 大数定律和中心极限定理

## 1. 依概率收敛

**定义 21** $Y_1, \dots, Y_n$, $\exists c \in R, \forall \epsilon > 0$,

$$\lim_{n\to+\infty} P\left\{|Y_n - c| \geq \epsilon\right\} = 0$$

称 $\{Y_n, n \geq 1\}$ 依概率收敛于 $c$, 记 $Y_n \xrightarrow{P} c, (n \to +\infty)$.

性质: 若 $X_n \xrightarrow{P} a, Y_n \xrightarrow{P} b$, $g \in \mathbb{C}(a,b)$,

$$g(X_n, Y_n) \xrightarrow{P} g(a,b)$$

## 2. 不等式

### 2.1 马尔可夫不等式

**定理 4** 对随机变量 $Y$, $\exists k$ 阶矩 $(k \geq 1)$, $\forall \epsilon > 0$

$$P\left\{|Y| \geq \epsilon\right\} \leq \frac{E\left(|Y|^k\right)}{\epsilon^k}$$

$$or\ P\left\{|Y| < \epsilon\right\} \geq 1 - \frac{E\left(|Y|^k\right)}{\epsilon^k}$$

### 2.2 切比雪夫不等式

**定理 5** 对随机变量 $X$, $\exists Var(x)$, $\forall \epsilon > 0$

$$P\left\{|X - E(X)| \geq \epsilon\right\} \leq \frac{Var(x)}{\epsilon^2}$$

$$P\left\{|X - E(X)| < \epsilon\right\} \geq 1 - \frac{Var(x)}{\epsilon^2}$$

## 3. 大数定律

### 3.1 弱大数定律

**定理 6** 随机变量序列 $Y_1, \dots, Y_n, \dots$, 若 $\exists \{c_n, n \geq 1\}$, 则当 $n \to +\infty$,

$$\frac{1}{n}\sum_{i=1}^{n} Y_i - c_n \xrightarrow{P} 0$$

即 $\forall \epsilon > 0$,

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Y_i - c_n\right| \geq \epsilon\right\} = 0$$

特别地, 当 $c_n \equiv c, n = 1, 2, \dots$, $n \to +\infty$,

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{P} c$$

### 3.2 切比雪夫大数定律

**定理 7** $X_1, X_2, \dots, X_n$ 相互独立, 且 $\exists E(X_i) = \mu$, $Var(X_i) = \sigma^2$, 则 $n \to +\infty$,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$$

### 3.3 辛钦大数定律

**定理 8** $X_1, \dots, X_n$ 独立同分布, 且 $\exists E(X_i) = \mu$, 则当 $n \to +\infty$,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$$

若 $h(x) \in \mathbb{C}$, $\exists E(h(X_1)) = a$, 则当 $n \to +\infty$,

$$\frac{1}{n}\sum_{i=1}^{n} h(x_i) \xrightarrow{P} a$$

### 3.4 贝努里大数定律

**定理 9** $n_A$ 为 $n$ 重贝努里实验中 $A$ 发生次数, $P(A) = p$, 则当 $n \to +\infty$,

$$\frac{n_A}{n} \xrightarrow{P} p$$

## 4. 中心极限定律

### 4.1 林德贝格-列维中心极限定理 (独立同分布)

**定理 10** $X_1, X_2, \dots, X_n$ 独立同分布, $E(X_i) = \mu$, $Var(X_i) = \sigma^2$, 当 $n$ 足够大时 $(n > 50)$,

$$\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$$

### 4.2 棣莫弗—拉普拉斯中心极限定理 (二项分布)

**定理 11** $n_A$ 为 $n$ 重贝努里实验中 $A$ 发生次数, $P(A) = p$, 当 $n$ 足够大时,

$$B(n,p) \sim N\left(np, np(1-p)\right)$$

# 六、 统计量与抽样分布

## 1. 随机样本与统计量

### 1.1 总体

随机变量 $X$.

### 1.2 样本 (简单随机样本)

$X_1, X_2, \ldots, X_n$ 相互独立且与 $X$ 同分布.

### 1.3 样本均值

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

### 1.4 样本方差

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$
$$Var(S^2) = \frac{2\sigma^4}{n-1}$$

### 1.5 样本 k 阶 (原点) 矩

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \,(k = 1, 2, \ldots)$$

### 1.6 样本 k 阶中心矩

$$B_k = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^k \,(k = 1, 2, \ldots)$$

## 2. 一些分布

### 2.1 $\chi^2$ 分布

**定义 22** 随机变量 $X_1, X_2, \ldots, X_n$ 相互独立, $X_i \sim N(0,1)$

$$\chi^2 = \sum_{i=1}^{2} X_i^2 \sim \chi^2(n)$$

性质:

(1) $X \sim \chi^2(n)$, 则 $E(X) = n$, $Var(X) = 2n$.

(2) $Y_i \sim \chi^2(n)$, $i = 1, 2$, $Y_1, Y_2$ 相互独立,
则 $Y_1 + Y_2 \sim \chi^2(n_1 + n_2)$.

### 2.2 $t$ 分布

**定义 23** $X, Y$ 相互独立, $X \sim N(0,1)$, $Y \sim \chi^2(n)$,

$$T = \frac{X}{\sqrt{\frac{Y}{n}}} \sim t(n)$$

性质: $t_{1-\alpha}(n) + t_\alpha(n) = 0$

### 2.3 $F$ 分布

**定义 24** $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, $X, Y$ 相互独立,

$$F = \frac{\frac{X}{n_1}}{\frac{Y}{n_2}} \sim F(n_1, n_2)$$

性质:

(1) $F \sim F(n_1, n_2)$, 则 $\frac{1}{F} \sim F(n_2, n_1)$.

(2) $F_{1-\alpha}(n_1, n_2) F_\alpha(n_2, n_1) = 1$.

(3) $T \sim t(n)$, 则 $T^2 \sim F(1, n)$.

## 3. 正态总体下的抽样分布

**定理 12** $X_1, X_2, \ldots, X_n$ 为自正太总体 $N(\mu, \sigma^2)$ 简单随机样本, $\bar{X}$ 为样本均值, $S^2$ 为样本方差, 则

*(1)*

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

*(2)*

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

*(3)*

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

$\bar{X}$ 与 $S^2$ 相互独立.

**定理 13** $(X_1, \ldots, X_{n_1})$, $(Y_1, \ldots, Y_{n_2})$ 来自 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$, 且相互独立, 有 $\bar{X}, \bar{Y}, S_1^2, S_2^2$, 则

*(1)*

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{\frac{S_1^2}{S_2^2}}{\frac{\sigma_1^2}{\sigma_2^2}} \sim F(n_1 - 1, n_2 - 1)$$

*(2)*

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

*(3)* 当 $\sigma_1^2=\sigma_2^2=\sigma^2$,

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

$$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

$E(S_w^2)=\sigma^2$, $Var(S_w^2)=\frac{\sigma^4}{n_1+n_2-2}$

# 七、 参数估计

## 1. 矩估计

总体 $X$ 有 m 个未知参数 $\theta_1,\dots,\theta_m$,
$\exists m$ 阶矩 $\mu_1,\dots,\mu_m$

(1) 计算 $\mu_k=E(X^k)=g_k(\theta_1,\dots,\theta_m)$, $k=1,\dots,m$.

(2) 求反函数, 得 $\theta_k=h_k(\mu_1,\dots,\mu_m)$, $k=1,\dots,m$.

(3) 以样本各阶矩 $A_1,\dots,A_m$ 代替
总体各阶矩 $\theta_1,\dots,\theta_m$, 得各参数的矩估计

$$\hat{\theta}=h_k(A_1,\dots,A_m),\ k=1,\dots,m$$

## 2. 极大似然估计

设总体 $X$ 的分布律为 $p(x;\theta)$ (或密度函数 $f(x;\theta)$), $\theta\in\Theta$. 从总体 $X$ 中取得样本 $X_1,\dots,X_n$, 其观察值为 $x_1,\dots,x_n$. 似然函数 $L(\theta)=\prod_{i=1}^n p(x_i;\theta)$ (或 $L(\theta)=\prod_{i=1}^n f(x_i;\theta)$).

$$\text{极大似然原理} L\left(\hat{\theta}(x_1,\dots,x_n)\right)=\max_{\theta\in\Theta}L(\theta)$$

$\hat{\theta}(x_1,\dots,x_n)$ 称为 $\theta$ 的极大似然估计值, 相应统计量 $\hat{\theta}(X_1,\dots,X_n)$ 称为 $\theta$ 的极大似然估计量 (MLE).

求解:

(1) 令 $\ln L(\theta)=l(\theta)$, 称为对数似然函数, 再令

$$\left.\frac{\partial l(\theta)}{\partial\theta_i}\right|_{\hat{\theta}_i,1\le i\le k}=0$$

解得 $\hat{\theta}_i$, $i=1,2,\dots,k$.

(2) 若 $L(\theta)$ 关于某个 $\theta_i$ 单调增 (减),

$$\theta_i\le(\ge)\hat{\theta}_i(x_1,\dots,x_n)$$

此时 $\hat{\theta}(x_1,\dots,x_n)$ 为 $\theta$ 的极大似然估计值, 相应统计量 $\hat{\theta}(X_1,\dots,X_n)$ 为 $\theta$ 的极大似然估计量.

(3) 若 $\hat{\theta}(X_1,\dots,X_n)$ 为 $\theta$ 的极大似然估计量, 则 $g(\theta)$ 的极大似然估计量为 $g\left(\hat{\theta}(X_1,\dots,X_n)\right)$.

## 3. 估计量评选准则

### 3.1 无偏性准则

**定义 25** 若参数 $\theta$ 的估计量 $\hat{\theta}=\hat{\theta}(X_1,\dots,X_n)$ 满足

$$E(\hat{\theta})=\theta$$

则称 $\hat{\theta}$ 是 $\theta$ 的无偏估计量.

### 3.2 有效性准则

**定义 26** $\hat{\theta}_1, \hat{\theta}_2$ 是 $\theta$ 的两个无偏估计量, 如果 $\forall \theta \in \Theta$,

$$Var(\hat{\theta}_1) \le Var(\hat{\theta}_2)$$

且不等号至少对某一 $\theta$ 成立, 则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

### 3.3 均方误差准则

**定义 27** $Mse(\hat{\theta}) = E\left(\hat{\theta} - \theta\right)^2$ 称为 $\hat{\theta}$ 关于 $\theta$ 的均方误差. $\hat{\theta}_1, \hat{\theta}_2$ 是 $\theta$ 的两个估计量, 如果 $\forall \theta \in \Theta$,

$$Mse(\hat{\theta}_1) \le Mse(\hat{\theta}_2)$$

且不等号至少对某一 $\theta$ 成立，则称 $\hat{\theta}_1$ 优于 $\hat{\theta}_2$.

### 3.4 相合性准则

**定义 28** 设 $\hat{\theta}_n = \hat{\theta}(X_1, \dots, X_n)$ 为参数 $\theta$ 的估计量，若 $\forall \theta \in \Theta$, 当 $n \to +\infty$,

$$\hat{\theta}_n \xrightarrow{P} \theta$$

即 $\forall \epsilon > 0$,

$$\lim_{n \to +\infty} P\left\{\left|\hat{\theta}_n - \theta\right| \ge \epsilon\right\} = 0$$

则称 $\hat{\theta}_n$ 为 $\theta$ 的 相合估计量 或 一致估计量.

## 4. 置信区间

**定义 29** 设总体 $X$ 的分布函数 $F(x;\theta)$, $\theta$ 未知. 对给定值 $\alpha(0 < \alpha < 1)$, 有两个统计量 $\theta_L = \theta_L(X_1, \dots, X_n)$,$\theta_U = \theta_U(X_1, \dots, X_n)$, 使得

$$P\{\theta_L(X_1, \dots, X_n) < \theta < \theta_U(X_1, \dots, X_n)\} \ge 1 - \alpha$$

称 $(\theta_L, \theta_U)$ 是 $\theta$ 的置信水平为 $1 - \alpha$ 的双侧置信区间; $\theta_L, \theta_U$ 分别为双侧置信下限和双侧置信上限.

若

$$P\{\theta_L(X_1, \dots, X_n) < \theta\} \ge 1 - \alpha$$

则称 $\theta_L$ 是 $\theta$ 的置信水平为 $1 - \alpha$ 的单侧置信下限.

若

$$P\{\theta < \theta_U(X_1, \dots, X_n)\} \ge 1 - \alpha$$

则称 $\theta_U$ 是 $\theta$ 的置信水平为 $1 - \alpha$ 的单侧置信上限.

## 5. 正态总体均值, 方差的置信区间与单侧置信限

注意区分是成对数据还是两正态总体数据

# 八、 假设检验

## 1. 假设检验的过程

### 1.1 利用拒绝域

(1) 提出假设 (原假设 $H_0 \longleftrightarrow$ 备择假设 $H_1$).

(2) 提出检验统计量和拒绝域形式.

(3) 在给定显著性水平 $\alpha$ 下, 根据 Neyman-Pearson 原则求出拒绝域的临界值.

(4) 根据实际样本观测值作出判断.

### 1.2 利用 $P\_$ 值

(1) 提出假设 (原假设 $H_0 \longleftrightarrow$ 备择假设 $H_1$).

(2) 提出检验统计量和拒绝域形式.

(3) 计算检验统计量的观测值与 $P\_$ 值.

(4) 根据 $P\_$ 值和显著性水平 $\alpha$, 作出判断

- 若 $P\_ \le \alpha$, 则拒绝原假设.
- 若 $P\_ \ge \alpha$, 则接受原假设

## 2. 两类错误

(1) 第 I 类错误: 拒绝真实的原假设.

(2) 第 II 类错误: 接受错误的原假设.

### 2.1 Neyman-Pearson 原则

首先控制犯第 I 类错误的概率不超过显著性水平 $\alpha$, 再寻找检验, 使得犯第 II 类错误的概率尽可能小.

## 3. $P\_$ 值计算

设检验统计量为 H, 检验统计量的观测值为 h.

(1) (双边检验)$H_0: \theta = \theta_0$, $H_1: \theta \ne \theta_0$
若检验统计量太大或太小时拒绝,
令 $p = P(H \ge h|\theta = \theta_0)$, 则 $P\_ = 2\min(p, 1-p)$.

(2) (左边检验)$H_0: \theta \ge \theta_0$, $H_1: \theta < \theta_0$
若检验统计量太小时拒绝, 则 $P\_ = P(H \le h|\theta = \theta_0)$.

(3) (右边检验)$H_0: \theta \le \theta_0$, $H_1: \theta > \theta_0$
若检验统计量太大时拒绝, 则 $P\_ = P(H \ge h|\theta = \theta_0)$.

| | 待估参数 | 其他参数 | 枢轴量及分布 | 置信区间 | 单侧置信限 |
|---|---|---|---|---|---|
| 一个 | $\mu$ | $\sigma^2$ 已知 | | | |
| 正态 | $\mu$ | $\sigma^2$ 未知 | | | |
| 总体 | $\sigma^2$ | $\mu$ 未知 | | | |
| 两个 | $\mu_1 - \mu_2$ | $\sigma_1^2, \sigma_2^2$ 已知 | | | |
| 正态 | $\mu_1 - \mu_2$ | $\sigma_1^2 = \sigma_2^2$ 未知 | | | |
| 总体 | $\frac{\sigma_1^2}{\sigma_2^2}$ | $\mu_1, \mu_2$ 未知 | | | |

表 1: 正态总体均值, 方差的置信区间与单侧置信限

| | 枢轴量 | 待估参数 | 置信区间 | 原假设 | 拒绝域 (代入真值) | 检验统计量值 | $P\_$ 值 |
|---|---|---|---|---|---|---|---|
| 一个<br>正态<br>总体 | | | | | | | |
| 两个<br>正态<br>总体 | | | | | | | |

表 2: 正态总体均值、方差的置信区间与假设检验

**4. 正态总体均值、方差的置信区间与假设检验**

**5. 拟合优度检验**

步骤:

(1) 提出原假设 $H_0 : X \sim p(x;\theta_1,\dots,\theta_r)$ (或 $X \sim f(x;\theta_1,\dots,\theta_r)$).

(2) 求极大似然估计 $\hat{\theta}_1,\dots,\hat{\theta}_r$, 如果没有未知参数，跳过这一步.

(3) 在 $H_0$ 下, 总体 $X$ 取值的全体分成 $k$ 个两两不相交的子集 $A_1,\dots,A_k$, 以 $n_i(i=1,\dots,k)$ 记样本观察值 $x_1,\dots,x_n$ 中落在 $A_i$ 的个数 (实际频数). (一般题中已给出)

(4) 计算当 $H_0$ 为真 (用第 2 步的 $\hat{\theta}_1,\dots,\hat{\theta}_r$ 或真值) 时 $A_i$ 发生的概率估计值 $\hat{p}_i = P_{H_0}(A_i)$, $i=1,\dots,k$, 称 $n\hat{p}_i$ (或 $np_i$) 为理论频数.
检验: $np_i \geq 5$, 否则合并子集并转到步骤 3.

(5) 检验统计量

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{k} \frac{n_i^2}{np^i} - n$$

(6) 拒绝域

$$\chi^2 \geq \chi_\alpha^2(k-r-1)$$

此时拒绝原假设.