# 《概率论与数理统计》期末速通 - 要点

# 1. 基本概念

## 1.1 随机事件

**[定理1.1.1]** **[事件的运算法则]**

(1) **[吸收律]** 若 $A \subset B$ , 则 $A \bigcup B = B, AB = A, \overline{B} \subset \overline{A}$ .

(2) **[交换律]**

① $A \bigcup B = B \bigcup A$ .

② $AB = BA$ .

(3) **[结合律]**

① $(A \bigcup B) \bigcup C = A \bigcup (B \bigcup C)$ .

② $(AB)C = A(BC)$ .

(4) **[分配律]**

① $A(B \bigcup C) = AB \bigcup AC$ .

② $A \bigcup BC = (A \bigcup B)(A \bigcup C)$ .

③ $A(B - C) = AB - AC$ .

(5) **[对偶律,De Morgan律]**

① $\overline{A \bigcup B} = \overline{A}\ \overline{B}, \overline{A \bigcup B \bigcup C} = \overline{A}\ \overline{B}\ \overline{C}$ .

② $\overline{AB} = \overline{A} \bigcup \overline{B}, \overline{ABC} = \overline{A} \bigcup \overline{B} \bigcup \overline{C}$ .

---

## 1.2 频率与概率

概率的公理化定理中的可加性是可列可加性, 性质中的可加性是有限可加性.

**[定理1.2.1]**

(1) **[有限可加性]** 设 $A_1, \cdots, A_n$ 是一列两两互斥的事件, 则 $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ .

(2) **[减法公式]** 对事件 $A$ 和事件 $B$ , 有 $P\left(A\overline{B}\right) = P(A - B) = P(A) - P(AB)$ .

(3) **[单调性]** 对事件 $A$ 和事件 $B$ , 若 $B \subset A$ , 则 $P(A) \geq P(B)$ .

(4) **[加法公式]**

① 对事件 $A$ 和事件 $B$ , 有 $P(A \bigcup B) = P(A) + P(B) - P(AB)$ .

② 对事件 $A, B, C$ , 有

$P(A \bigcup B \bigcup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$ .

---

# 1.3 条件概率

[**定义1.3.1**] 设 $A$ 和 $B$ 是两个事件, 且 $P(A)>0$ . 称 $P(B|A)=\dfrac{P(AB)}{P(A)}$ 为在事件 $A$ 发生的条件下事件 $B$ 发生的**条件概率**.

[**定理1.3.1**] 设随机试验 $E$ 的样本空间为 $S$ , $A$ 、 $B$ 、 $C$ 是三个事件, 且 $P(A)>0$ .

(1) $P(B\bigcup C\mid A)=P(B\mid A)+P(C\mid A)-P(BC\mid A)$ .

(2) 若 $B$ 与 $C$ 互斥, 则 $P(B\bigcup C\mid A)=P(B\mid A)+P(C\mid A)$ .

(3) $P(B-C\mid A)=P(B\mid A)-P(BC\mid A)$ .

(4) $P\left(\overline{B}\mid A\right)=1-P(B\mid A)$ .

[**定理1.3.2**] [**乘法公式**] 设 $A$ 和 $B$ 是两个事件.

(1) 若 $P(A)>0$ , 则 $P(AB)=P(A)P(B\mid A)$ .

(2) 若 $P(B)>0$ , 则 $P(AB)=P(B)P(A\mid B)$ .

[**推广**]

(3) 设 $A,B,C$ 是三个事件, 则 $P(ABC)=P(A)P(B\mid A)P(C\mid AB)$ .

(4) 设 $A_1,\cdots,A_n\ \ (n\geq 2)$ 是一列事件, $A_i$ 先于 $A_{i+1}$ 发生, 且 $P(A_1\cdots A_{n-1})>0$ , 则 $P(A_1\cdots A_n)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_n\mid A_1\cdots A_{n-1})$ .

[**注**] 乘法公式用于计算无独立性的若干个事件的积事件的概率. 若各事件独立, 则不能用乘法公式计算.

[**定理1.3.3**] [**全概率公式**] 设随机试验 $E$ 的样本空间为 $S$ , $A$ 是 $E$ 的一个事件, $B_1,\cdots,B_n$ 是 $S$ 的一个划分, 且 $P(B_i)>0$ , 则 $P(A)=\sum\limits_{i=1}^n P(A\mid B_i)P(B_i)$ .

[**推论**] $n=2$ 时, 有 $P(A)=P(A\mid B)P(B)+A\left(A\mid\overline{B}\right)P\left(\overline{B}\right)$ .

[**注**] 全概率公式用于已知因的概率, 求果的概率.

[**定理1.3.4**] [**Bayes公式**] 设随机试验 $E$ 的样本空间为 $S$ , $A$ 是 $E$ 的一个事件, $B_1,\cdots,B_n$ 是 $S$ 的一个划分, 且 $P(B_i)>0$ , 则 $P(B_i\mid A)=\dfrac{P(AB_i)}{P(A)}=\dfrac{P(A\mid B_i)P(B_i)}{\sum\limits_{j=1}^n P(A\mid B_j)P(B_j)}$ .

[**推论**] $n=2$ 时, $P(B\mid A)=\dfrac{P(AB)}{P(A)}=\dfrac{P(A\mid B)P(B)}{P(A\mid B)P(B)+P\left(A\mid\overline{B}\right)P\left(\overline{B}\right)}$ .

[**注**] Bayes公式用于已知果的概率, 求因的概率.

## 1.4 独立性

**[定理1.4.1]** 设 $A$ 和 $B$ 是两个相互独立的事件, 且 $P(A)>0$ , 则 $P(B \mid A)=P(B)$ .

**[定理1.4.2]** 设 $A$ 和 $B$ 是两个事件, 且 $P(A),P(B)>0$ , 则 $A$ 与 $B$ 独立、 $A$ 与 $B$ 互斥不能同时成立.

**[定义1.4.1]** 设 $A,B,C$ 是三个事件 .

(1) 若 $\begin{cases} P(AB)=P(A)P(B) \\ P(AC)=P(A)P(C) \\ P(BC)=P(B)P(C) \end{cases}$ , 则称 $A,B,C$ **两两独立**.

(2) 若 $\begin{cases} P(AB)=P(A)P(B) \\ P(AC)=P(A)P(C) \\ P(BC)=P(B)P(C) \\ P(ABC)=P(A)P(B)P(C) \end{cases}$ , 则称 $A,B,C$ **相互独立**.

---

# 2. 一维随机变量及其分布

## 2.1 常用分布

**[定理2.1.1]** 离散型随机变量的分布:

| 分布 | 分布律 |
| --- | --- |
| 0-1分布 $(0-1)(p)\ \ (0<p<1)$ | $P\{X=k\}=p^k(1-p)^{1-k}\ \ (k=0,1)$ |
| 二项分布 $b(n,p)\ \ (n\geq 1, 0<p<1)$ | $P\{X=k\}=C_n^k\cdot p^k(1-p)^{n-k}\ \ (k=0,1,\cdots,n)$ |
| Poisson分布 $\pi(\lambda)\ \ (\lambda>0)$ | $P\{X=k\}=\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}\ \ (k=0,1,2,\cdots)$ |

**[定理2.1.2]** 连续型随机变量的分布:

| 分布 | 概率密度 | 分布函数 |
| --- | --- | --- |
| 均匀分布 $U(a,b)\ \ (a<b)$ | $f(x)=\begin{cases} \dfrac{1}{b-a}, a<x<b \\ 0, otherwise \end{cases}$ | $F(x)=\begin{cases} 0, x<a \\ \dfrac{x-a}{b-a}, a\leq x<b \\ 1, x\geq b \end{cases}$ |
| 指数分布 $Exp(\theta)\ \ (\theta>0)$ | $f(x)=\begin{cases} \dfrac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, x>0 \\ 0, x\leq 0 \end{cases}$ | $F(x)=\begin{cases} 1-\mathrm{e}^{-\frac{x}{\theta}}, x\geq 0 \\ 0, x<0 \end{cases}$ |
| 正态分布 $N(\mu,\sigma^2)\ \ (\sigma>0)$ | $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$ | $F(x)=\displaystyle\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$ |

**[定理2.1.3]** **[概率积分]** $\displaystyle\int_{-\infty}^{+\infty}\mathrm{e}^{-x^2}\mathrm{d}x=\sqrt{\pi}$ .

## 2.2 分布函数与概率密度

[**定理2.2.1**] 对随机变量 $X$ , $F(x)$ 是 $X$ 的分布函数的充要条件为如下三条性质: ① $F(x)$ 单调不减; ② $0 \leq F(x) \leq 1$ , 且 $\lim\limits_{x \to -\infty} F(x) = 0, \lim\limits_{x \to +\infty} F(x) = 1$ ; ③ $F(x)$ 是右连续的, 即对 $\forall x$ , 都有 $\lim\limits_{x \to x_0^+} F(x) = F(x_0)$ .

[**定理2.2.2**] 设随机变量 $X$ 的概率密度函数为 $f(x)$ , 分布函数为 $F(x) = \int_{-\infty}^{x} f(t) \mathrm{d}t$ , 则:

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

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

(3) $P\{x_1 < X \leq x_2\} = P\{X \leq x_2\} - P\{X \leq x_1\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) \mathrm{d}x$ .

(4) 若 $f(x)$ 可积, 则 $F(x)$ 连续.

(5) 若 $f(x)$ 连续, 则 $F(x)$ 可导, 且 $F'(x_0) = f(x_0)$ .

[**注**] 若函数 $f(x)$ 有性质(1)和(2), 则 $F(x) = \int_{-\infty}^{x} f(t) \mathrm{d}t$ 是某一随机变量 $X$ 的分布函数, $f(x)$ 是 $X$ 的概率密度.

[**定理2.2.3**] 设连续型随机变量 $X$ 的概率密度函数为 $f(x)$ , 分布函数为 $F(x) = \int_{-\infty}^{x} f(t) \mathrm{d}t$ , 则:

(1) $F(x)$ 连续.

(2) 若 $f(x)$ 在点 $x = x_0$ 处连续, 则 $F'(x_0) = f(x_0)$ .

(3) 对 $\forall$ 常数 $a \in \mathbb{R}$ , 有 $P\{X = a\} = 0$ .

(4)

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

## 2.3 正态分布

[**定义2.3.1**] 若连续型随机变量 $X$ 的概率密度 $f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$ , 其中 $\sigma > 0, \mu$ 都是常数, 则称 $X$ 服从参数为 $\mu, \sigma$ 的**正态分布**或**Gauss分布**, 记作 $X \sim N(\mu, \sigma^2)$ , 其分布函数 $F(x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t$ .

[**定义2.3.2**] 对连续型随机变量 $X$ , 若 $X \sim N(0, 1)$ , 则称 $X$ 服从**标准正态分布**, 其概率密度 $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}$ , 分布函数 $\varPhi(x) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t$ .

[**定理2.3.3**] 设连续型随机变量 $X$ 服从标准正态分布, 其概率密度和分布函数分别为 $\varphi(x)$ 和 $\varPhi(x)$ , 则: $\varPhi(-x) = 1 - \varPhi(x)$ .

**[定理2.3.4]** 任一正态分布可经一线性变换转化为标准正态分布. 具体地, 设连续型随机变量 $X \sim N(\mu, \sigma^2)$ , 则随机变量 $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$ .

**[定理2.3.5]** 设连续型随机变量 $X \sim N(\mu, \sigma^2)$ , 则:

(1) $F(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right)$ .

(2) $P\{x_1 < X \leq x_2\} = \Phi\left(\dfrac{x_2-\mu}{\sigma}\right) - \Phi\left(\dfrac{x_1-\mu}{\sigma}\right)$ .

---

## 2.4 随机变量的函数

**[定理2.4.1]** 设随机变量 $X$ 有概率密度 $f_X(x)$ . 若 $y=g(x)$ 是关于 $x$ 的严格单调且可导的函数, 即恒有 $g'(x)>0$ 或 $g'(x)<0$ , 则 $Y=g(X)$ 是连续型随机变量, 其概率密度 $f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)|, \alpha < y < \beta \\ 0, otherwise \end{cases}$ , 其中 $x=h(y)$ 是 $y=g(x)$ 的反函数, $\alpha = \min\{g(-\infty), g(+\infty)\}, \beta = \max\{g(-\infty), g(+\infty)\}$ .

(1) $g'(x)>0$ 时, 有 $f_Y(y) = \begin{cases} f_X(h(y)) \cdot [h'(y)], \alpha < y < \beta \\ 0, otherwise \end{cases}$ .

(2) $g'(x)<0$ 时, 有 $f_Y(y) = \begin{cases} f_X(h(y)) \cdot [-h'(y)], \alpha < y < \beta \\ 0, otherwise \end{cases}$ .

**[推论]** 在本定理的条件下, 若 $f_X(x)$ 在有限区间 $[a,b]$ 以外的其他点处为 $0$ , 且 $x \in [a,b]$ 时, 有 $g'(x)>0$ 或 $g'(x)<0$ , 则 $f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)|, \alpha < y < \beta \\ 0, otherwise \end{cases}$ , 其中 $\alpha = \min\{g(a), g(b)\}, \beta = \max\{g(a), g(b)\}$ .

**[注]** 连续型随机变量的函数未必是连续型随机变量.

---

---

# 3. 多维随机变量及其分布

## 3.1 二维随机变量

**[定理3.1.1]** **[分布函数的性质]** 设二维随机变量 $(X,Y)$ 的联合分布函数为 $F(x,y)$ .

(1) **[单调性]** $F(x,y)$ 是分别关于 $x$ 和 $y$ 的不减函数, 即:

① 对固定的 $y$ , 若 $x_2 > x_1$ , 则 $F(x_2,y) \geq F(x_1,y)$ .

② 对固定的 $x$ , 若 $y_2 > y_1$ , 则 $F(x,y_2) \geq F(x,y_1)$ .

(2) **[有界性]** $0 \leq F(x,y) \leq 1$ .

(3) ① 对固定的 $y$ , 有$F(-\infty,y) = \lim\limits_{x\to-\infty} F(x,y) = 0$ .

② 对固定的 $x$ , 有 $F(x,-\infty) = \lim\limits_{y\to-\infty} F(x,y) = 0$ .

③ $F(-\infty,-\infty) = \lim\limits_{\substack{x\to-\infty \\ y\to-\infty}} F(x,y) = 0, F(+\infty,+\infty) = \lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(x,y) = 1$ .

④ $\lim\limits_{x\to+\infty} F(x,y), \lim\limits_{y\to+\infty} F(x,y)$ 都无法确定.

(4) ① 对固定的 $y$, $F(x,y)$ 关于 $x$ 右连续, 即 $F(x_0+0,y)=\lim\limits_{x\to x_0^+}F(x,y)=F(x_0,y)$ .

② 对固定的 $x$ , $F(x,y)$ 关于 $y$ 右连续, 即 $F(x,y_0+0)=\lim\limits_{y\to y_0^+}F(x,y)=F(x,y_0)$ .

(5)对 $\forall x_1<x_2,y_1<y_2$ , 有
$P\{x_1<X\leq x_2,y_1<Y\leq y_2\}=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geq 0$ .

[**定理3.1.2**] 二元函数 $f(x,y)$ 是二维连续型随机变量 $(X,Y)$ 的概率密度的充要条件为:

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

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

[**定理3.1.3**] 设二维连续型随机变量 $(X,Y)$ 的联合分布函数为 $F(x,y)$ , 概率密度为 $f(x,y)$ , 则:

(1) $F(x,y)$ 在 $\mathbb{R}^2$ 上连续.

(2) 在 $f(x,y)$ 的连续点处, 有 $\dfrac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y)$ .

(3) 对平面区域 $G$ , 有 $P\{(X,Y)\in G\}=\displaystyle\iint_G f(x,y)\mathrm{d}x\mathrm{d}y$ .

[**推论**] 设二维随机变量 $(X,Y)$ 的分布函数为 $F(x,y)$ . 若 $F(x,y)$ 可导, 则 $(X,Y)$ 是二维连续型随机变量, 且 $\dfrac{\partial^2F(x,y)}{\partial x\partial y}$ 是它的一个概率密度.

---

## 3.2 边缘分布

[**定理3.2.1**]

(1) 设二维随机变量 $(X,Y)$ 的联合分布函数为 $F(x,y)$ , 则:

① 关于 $X$ 的边缘分布函数 $F_X(x)=\lim\limits_{y\to+\infty}F(x,y)$ .

② 关于 $Y$ 的边缘分布函数 $F_Y(y)=\lim\limits_{x\to+\infty}F(x,y)$ .

(2) 设二维连续型随机变量 $(X,Y)$ 的概率密度为 $f(x,y)$ , 则:

① 关于 $X$ 的边缘分布函数 $F_X(x)=\displaystyle\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$ .

② 关于 $Y$ 的边缘分布函数 $F_Y(y)=\displaystyle\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$ .

[**注**] 联合分布函数可唯一确定两个边缘分布函数, 但两个边缘分布函数不能唯一确定联合分布函数.

[**定理3.2.2**] 设二维随机变量 $(X,Y)$ 的联合概率密度为 $f(x,y)$ , 则:

(1) 关于 $X$ 的边缘概率密度 $f_X(x)=\displaystyle\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$ .

(2) 关于 $Y$ 的边缘概率密度 $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$ .

**[注]** 联合概率密度可唯一确定两个边缘概率密度, 但两个边缘概率密度不能唯一确定联合概率密度, 如二维正态分布.

## 3.3 条件分布

**[定义3.3.1]** 设二维离散型随机变量 $(X,Y)$ 的联合分布律为 $P\{X=x_i,Y=y_j\}=p_{ij}\ \ (i,j=1,2,\cdots)$ , 关于 $X$ 的边缘分布律 $P\{X=x_i\}=\sum\limits_{j=1}^{+\infty}P\{X=x_i,Y=y_i\}=p_{i\cdot}=\sum\limits_{j=1}^{+\infty}p_{ij}\ \ (i=1,2,\cdots)$ , 关于 $Y$ 的边缘分布律 $P\{Y=y_j\}=\sum\limits_{j=1}^{+\infty}P\{X=x_i,Y=y_j\}=p_{\cdot j}=\sum\limits_{i=1}^{+\infty}p_{ij}\ \ (i=1,2,\cdots)$ .

(1) 对固定的 $j$ , 若 $P\{Y=y_j\}>0$ , 则称 $P\{X=x_i\mid Y=y_j\}=\dfrac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\dfrac{p_{ij}}{p_{\cdot j}}\ \ (i=1,2,\cdots)$ 为在 $Y=y_j$ 条件下随机变量 $X$ 的**条件分布律**.

(2) 对固定的 $i$ , 若 $P\{X=x_i\}>0$ , 则称 $P\{Y=y_j\mid X=x_i\}=\dfrac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\dfrac{p_{ij}}{p_{i\cdot}}\ \ (j=1,2,\cdots)$ 为在 $X=x_i$ 条件下随机变量 $Y$ 的**条件分布律**.

**[注1]** 条件分布律的分子为联合分布, 分母为边缘分布, 而边缘分布可由联合分布唯一确定, 故条件分布律由联合分布唯一确定.

**[注2]** 对二维连续型随机变量 $(X,Y)$ 和常数 $x_i,y_j\in\mathbb{R}$ , 有 $P\{X=x_i\}=P\{Y=y_j\}=0$ , 故无法用上述方法定义条件分布.

**[定义3.3.2]** 设二维连续型随机变量 $(X,Y)$ 的概率密度为 $f(x,y)$ , 关于 $X$ 的概率密度为 $f_X(x)$ , 关于 $Y$ 的概率密度为 $f_Y(y)$ .

(1) 对固定的 $y$ , 若 $f_Y(y)>0$ , 则称 $\dfrac{f(x,y)}{f_Y(y)}$ 为在 $Y=y$ 条件下随机变量 $X$ 的**条件概率密度**, 记作 $f_{X|Y}(x\mid y)=\dfrac{f(x,y)}{f_Y(y)}$ . 在 $Y=y$ 的条件下 $X$ 的**条件分布函数** $F_{X|Y}(x\mid y)=P\{X\le x\mid Y=y\}=\int_{-\infty}^{x}\dfrac{f(x,y)}{f_Y(y)}\mathrm{d}x$ .

(2) 对固定的 $x$ , 若 $f_X(x)>0$ , 则称 $\dfrac{f(x,y)}{f_X(x)}$ 为在 $X=x$ 条件下随机变量 $Y$ 的**条件概率密度**, 记作 $f_{Y|X}(y\mid x)=\dfrac{f(x,y)}{f_X(x)}$ . 在 $X=x$ 条件下 $X$ 的**条件分布函数** $F_{Y|X}(y\mid x)=P\{Y\le y\mid X=x\}=\int_{-\infty}^{y}\dfrac{f(x,y)}{f_X(x)}\mathrm{d}y$ .

## 3.4 独立性

**[定义3.4.1]** 设二维随机变量 $(X,Y)$ 的联合分布函数为 $F(x,y)$ , 两个边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$ . 若对 $\forall x,y\in\mathbb{R}$ , 都有 $P\{X\le x,Y\le y\}=P\{X\le x\}\cdot P\{Y\le y\}$ , 即 $F(x,y)=F_X(x)\cdot F_Y(y)$ , 则称随机变量 $X$ 与随机变量 $Y$ **相互独立**.

(1) 对二维离散型随机变量 $(X,Y)$ , 称随机变量 $X$ 与随机变量 $Y$ **相互独立**, 如果对 $(X,Y)$ 的所有取值 $(x_i,y_j)\ \ (1\le i\le n, 1\le j\le m)$ , 都有 $P\{X=x_i,Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\}$ .

(2) 设二维连续型随机变量 $(X,Y)$ 的联合概率密度为 $f(x,y)$ , 两个边缘概率密度分别为 $f_X(x)$ 和 $f_Y(y)$ . 称随机变量 $X$ 与随机变量 $Y$ **相互独立**, 如果 $f(x,y)=f_X(x)\cdot f_Y(y)$ 在 $\mathbb{R}^2$ 上几乎处处成立(不成立的点构成零测集).

---

## 3.5 随机变量函数的分布

[**定理3.5.1**] 设二维连续型随机变量 $(X,Y)$ 的联合概率密度为 $f(x,y)$ , 随机变量 $Z=g(X,Y)$ . 若可从函数 $z=g(x,y)$ 中解出 $y=h(x,z)$ , 则 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f(x,h(x,z))\cdot\left|\frac{\partial h(x,z)}{\partial z}\right|\mathrm{d}x$ .

[**定理3.5.2**] 设二维连续型随机变量 $(X,Y)$ 的联合概率密度为 $f(x,y)$ , 则 $Z=X+Y$ 是连续型随机变量, 其概率密度 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x$ 或 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f(z-y,y)\mathrm{d}y$ . 若随机变量 $X$ 与随机变量 $Y$ 独立, 设 $(X,Y)$ 的两个边缘分布分别为 $f_X(x)$ 和 $f_Y(y)$ , 则 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f_X(x)\cdot f_Y(z-x)\mathrm{d}x$ 或 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f_X(z-y)\cdot f_Y(y)\mathrm{d}y$ , 其中 $f_X*f_Y=\displaystyle\int_{-\infty}^{+\infty}f_X(x)\cdot f_Y(z-x)\mathrm{d}x=\int_{-\infty}^{+\infty}f_X(z-y)\cdot f_Y(y)\mathrm{d}y$ 称为 $f_X$ 与 $f_Y$ 的**卷积公式**.

[**定理3.5.3**] 设 $X_i\sim N(\mu_i,\sigma_i^2)\ \ (1\le i\le n)$ , 且 $X_1,\cdots,X_n$ 相互独立, 则 $X_1+\cdots+X_n\sim N\left(\displaystyle\sum_{i=1}^n\mu_i,\sum_{i=1}^n\sigma_i^2\right)$ .

[**推论**] $n$ 个独立的、服从正态分布的随机变量的线性组合也服从正态分布.

[**定理3.5.4**] 设二维连续型随机变量 $(X,Y)$ 的联合概率密度为 $f(x,y)$ , 则 $Z=\dfrac{Y}{X}$ 是连续型随机变量, 其概率密度 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}|x|\cdot f(x,xz)\mathrm{d}x$ . 若随机变量 $X$ 与随机变量 $Y$ 独立, 则 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}|x|\cdot f_X(x)\cdot f_Y(xz)\mathrm{d}x$ .

[**证**] 设函数 $z=g(x,y)=\dfrac{y}{x}$ , 则 $y=h(x,z)=xz$ , 此时 $\left|\dfrac{\partial h(x,z)}{\partial z}\right|=|x|$ . 由**定理3.5.1**即证.

[**定理3.5.5**] 设二维连续型随机变量 $(X,Y)$ 的联合概率密度为 $f(x,y)$ , 则 $Z=X\cdot Y$ 是连续型随机变量, 其概率密度 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}\frac{1}{|x|}\cdot f\left(x,\frac{z}{x}\right)\mathrm{d}x$ . 若随机变量 $X$ 与随机变量 $Y$ 独立, 则 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}\frac{1}{|x|}\cdot f_X(x)\cdot f_Y\left(\frac{z}{x}\right)\mathrm{d}x$ .

[**证**] 设函数 $z=g(x,y)=xy$ , 则 $y=h(x,z)=\dfrac{z}{x}$ , 此时 $\left|\dfrac{\partial h(x,z)}{\partial z}\right|=\dfrac{1}{|x|}$ . 由**定理3.5.1**即证.

**[定理3.5.6]** 设两相互独立的随机变量 $X$ 和$Y$ 的分布函数分别为 $F_X(x)$ 和 $F_Y(y)$ , 则:

(1) ① 随机变量 $Z=\max\{X,Y\}$ 的分布函数 $F_Z(z)=F_X(z)\cdot F_Y(z)$ .

② 若 $X$ 与 $Y$ 同分布, 设它们的分布函数都为 $F(x)$ , 概率密度都为 $f(x)$ , 则 $F_Z(z)=[F(x)]^2$ , $Z$ 的概率密度 $f_Z(z)=2\cdot F(z)\cdot f(z)$ .

(2) ① 随机变量 $Z=\min\{X,Y\}$ 的分布函数 $F_Z(z)=1-[1-F_X(z)]\cdot[1-F_Y(z)]$ .

② 若 $X$ 与 $Y$ 同分布, 设它们的分布函数都为 $F(x)$ , 概率密度都为 $f(x)$ , 则 $F_Z(z)=1-[1-F_Z(z)]^2$ , $Z$ 的概率密度 $f_Z(z)=2\cdot[1-F(z)]\cdot f(z)$ .

**[推广]** 设随机变量 $X_1,\cdots,X_n$ 相互独立, 其中 $X_i\ \ (1\le i\le n)$ 的分布函数为 $F_{X_i}(x_i)$ .

(1) ① 随机变量 $Z=\max\{X_1,\cdots,X_n\}$ 的分布函数 $F_Z(z)=\prod\limits_{i=1}^n F_{X_i}(z)$ .

② 若 $X_1,\cdots,X_n$ 同分布, 设它们的分布函数都为 $F(x)$ , 则 $F_Z(z)=[F(z)]^n$ .

(2) ① 随机变量 $Z=\min\{X_1,\cdots,X_n\}$ 的分布函数 $F_Z(z)=1-\prod\limits_{i=1}^n[1-F_{X_i}(x_i)]$ .

② 若 $X_1,\cdots,X_n$ 同分布, 设它们的分布函数都为 $F(x)$ , 则 $F_Z(z)=1-[1-F(z)]^n$ .

---

# 4. 数字特征

## 4.1 常见分布的期望和方差

| 分布 | 期望 | 方差 |
| --- | --- | --- |
| 0-1分布 $(0-1)(p)\ \ (0<p<1)$ | $p$ | $p(1-p)$ |
| 二项分布 $b(n,p)\ \ (n\ge 1,0<p<1)$ | $np$ | $np(1-p)$ |
| Poisson分布 $\pi(\lambda)\ \ (\lambda>0)$ | $\lambda$ | $\lambda$ |
| 均匀分布 $U(a,b)\ \ (a<b)$ | $\dfrac{a+b}{2}$ | $\dfrac{(b-a)^2}{12}$ |
| 指数分布 $Exp(\theta)\ \ (\theta>0)$ | $\theta$ | $\theta^2$ |
| 正态分布 $N(\mu,\sigma^2)\ \ (\sigma>0)$ | $\mu$ | $\sigma^2$ |

---

## 4.2 期望

期望的定义中, 级数和反常积分要求绝对收敛.

**[定理4.2.1]** 设随机变量 $X$ , 函数 $g(x)$ 连续, $Y=g(X)$ , 则:

(1) 若 $X$ 为离散型随机变量, 且其分布律为 $P\{X=x_k\}=p_k\ \ (k=1,2,\cdots)$ . 若级数 $\sum\limits_{k=1}^n g(x_k)\cdot p_k$ 绝对收敛, 则 $E(Y)=E[g(X)]=\sum\limits_{k=1}^{+\infty} g(x_k)\cdot p_k$ .

(2) 若 $X$ 为连续型随机变量, 且其概率密度为 $f(x)$ . 若积分 $\int_{-\infty}^{+\infty} g(x) \cdot f(x) \mathrm{d} x$ 绝对收敛, 则 $E(Y)=E[g(X)]=\int_{-\infty}^{+\infty} g(x) \cdot f(x) \mathrm{d} x$ .

**[定理4.1.5]** 设二维随机变量 $(X, Y)$ , 二元函数函数 $g(x, y)$ 连续, 随机变量$Z=g(X, Y)$ . 若 $(X, Y)$ 为连续型随机变量, 且其联合概率密度为 $f(x, y)$ . 若积分 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \cdot f(x, y) \mathrm{d} x \mathrm{d} y$ 绝对收敛, 则 $E(Z)=E[g(X, Y)]=\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \cdot f(x, y) \mathrm{d} x \mathrm{d} y$ .

**[定理4.1.6]** **[数学期望的性质]**

(1) 对有限个随机变量 $X_1, \cdots, X_n$ , 有 $E\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n E(X_i)$ .

(2) 对有限个相互独立的随机变量 $X_1, \cdots, X_n$ , 有 $E\left(\prod_{i=1}^n X_i\right)=\prod_{i=1}^n E(X_i)$ .

**[注]** 随机变量 $X$ 与 $Y$ 相互独立是 $E(XY)=E(X) \cdot E(Y)$ 的充分条件, 但非必要条件.

---

## 4.3 方差

**[定义4.3.1]** 对随机变量 $X$ , 若期望 $E\{[X-E(X)]^2\}$ 存在, 则称该期望为 $X$ 的**方差**, 记作 $D(X)$ 或 $\operatorname{Var}(X)$ , 即 $D(X)=\operatorname{Var}(X)=E\{[X-E(X)]^2\}$ . 称 $\sigma(X)=\sqrt{D(X)}$ 为 $X$ 的**标准差**或**均方差**.

(1) 离散型随机变量 $X$ 的方差 $D(X)=\sum_{k=1}^{+\infty}[x_k-E(X)]^2 \cdot p_k$ .

(2) 连续型随机变量 $X$ 的方差 $D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 \cdot f(x) \mathrm{d} x$ , 其中 $f(x)$ 为 $X$ 的概率密度.

**[定理4.3.1]** 若随机变量 $X$ 的方差存在, 则 $D(X)=E(X^2)-[E(X)]^2$ .

**[定理4.3.2]** 设随机变量 $X$ 的期望 $E(X)=\mu$ , 方差 $D(X)=\sigma^2$ . 称随机变量 $X^*=\frac{X-\mu}{\sigma}$ 为 $X$ 的**标准化变量**, 其期望 $E(X^*)=0$ , 方差 $D(X^*)=1$ .

**[定理4.4.3]** **[方差的性质]**

(1) 对随机变量 $X$ 和常数 $C$ , 方差 $D(C \cdot X)=C^2 \cdot D(X), D(C+X)=D(X)$ .

(2) 设 $X, Y$ 是两个随机变量, 则:

① 方差 $D(X \pm Y)=D(X)+D(Y) \pm 2 \cdot E\{[X-E(X)] \cdot[Y-E(Y)]\}$

$=D(X)+D(Y) \pm 2 \cdot[E(XY)-E(X) \cdot E(Y)]$ .

② 若 $X$ 与 $Y$ 相互独立, 则方差 $D(X \pm Y)=D(X)+D(Y)$ .

③ 对有限个相互独立的随机变量 $X_1,\cdots,X_n$ , 方差 $D\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}D(X_i)$ .

④ 方差 $D(aX+bY)=a^2\cdot D(X)+b^2\cdot D(Y)$ .

(3) 方差 $D(X)=0$ iff 随机变量 $X$ 以概率 $1$ 取常数 $E(X)$ , 即 $P\{X=E(X)\}=1$ .

[**注**] $X$ 与 $Y$ 相互独立是 $D(X+Y)=D(X)+D(Y)$ 的充分条件, 但非必要条件.

[**定理4.4.4**]

(1) 若随机变量 $X\sim N(\mu,\sigma^2)$ , 则随机变量 $a\cdot X+b\sim N(a\mu+b,a^2\sigma^2)$ .

(2) 若随机变量 $X_1,\cdots,X_n$ 相互独立, 且 $X_i\sim N(\mu_i,\sigma_i^2)\ \ (1\leq i\leq n)$ , 则随机变量 $\sum_{i=1}^{n}c_i\cdot X_i\sim N\left(\sum_{i=1}^{n}c_i\cdot\mu_i,\sum_{i=1}^{n}c_i^2\cdot\sigma_i^2\right)$ .

---

## 4.4 Chebyshev不等式

[**定理4.4.1**] [**Chebyshev不等式**] 若随机变量 $X$ 的期望 $E(X)=\mu$ , 方差 $D(X)=\sigma^2$ , 则对 $\forall\varepsilon>0$ , 有 $P\{|X-\mu|\geq\varepsilon\}\leq\dfrac{\sigma^2}{\varepsilon^2}$ .

---

## 4.5 协方差与相关系数

[**定义4.5.1**] 对随机变量 $X$ 和 $Y$ , 称期望 $E\{[X-E(X)]\cdot[Y-E(Y)]\}$ 为 $X$ 与 $Y$ 的**协方差**, 记作 $\mathrm{Cov}(X,Y)$ , 即 $\mathrm{Cov}(X,Y)=E\{[X-E(X)]\cdot[Y-E(Y)]\}=E(XY)-E(X)\cdot E(Y)$ ; 称 $\rho_{X,Y}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}$ 为 $X$ 与 $Y$ 的**线性相关系数**, 简称**相关系数**.

[**注1**] 若 $X$ 与 $Y$ 相互独立, 则 $E\{[X-E(X)]\cdot[Y-E(Y)]\}=0$ , 则该值不为 0 时, 反映了 $X$ 与 $Y$ 间的相关关系.

[**注2**] $\mathrm{Cov}(X,Y)=E\{[X-E(X)]\cdot[Y-E(Y)]\}$ 为协方差的定义式, $\mathrm{Cov}(X,Y)=E(XY)-E(X)\cdot E(Y)$ 为协方差的计算式.

[**定理4.5.1**] [**协方差的性质**]

(1) 对随机变量 $X$ , 协方差 $\mathrm{Cov}(X,X)=D(X)$ .

(2) 对随机变量 $X$ 和常数 $C$ , 协方差 $\mathrm{Cov}(X,C)=0$ .

(3) 对两个随机变量 $X$ 和 $Y$ , 协方差 $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$ .

(4) 对两个随机变量 $X,Y$ 和两个常数 $a,b$ , 协方差 $\mathrm{Cov}(aX,bY)=ab\cdot\mathrm{Cov}(X,Y)$ .

(5) 对三个随机变量 $X_1,X_2$ 和 $Y$ , 协方差 $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$ .

(5) 对两个随机变量 $X$ 和 $Y$ , 方差 $D(X\pm Y)=D(X)+D(Y)\pm2\cdot\mathrm{Cov}(X,Y)$ .

[**定理4.5.2**] [**相关系数的性质**]

(1) 随机变量 $X$ 与 $Y$ 的相关系数 $\rho_{X,Y}$ 有界, 且 $|\rho_{X,Y}| \leq 1$ .

(2) 设随机变量 $X$ 与 $Y$ 的相关系数为 $\rho_{X,Y}$ , 则 $|\rho_{X,Y}| = 1$ iff $\exists$ 常数 $a, b\ s.t.\ P\{Y = a + b \cdot X\} = 1$ , 即 $X$ 与 $Y$ 以概率 $1$ 存在线性关系.

① $\rho_{X,Y} = 1$ iff $\exists$ 常数 $a, b\ s.t.\ P\{Y = a + b \cdot X\} = 1$ , 其中 $b > 0$ , 此时称 $X$ 与 $Y$ **正相关**.

② $\rho_{X,Y} = -1$ iff $\exists$ 常数 $a, b\ s.t.\ P\{Y = a + b \cdot X\} = 1$ , 其中 $b < 0$ , 此时称 $X$ 与 $Y$ **负相关**.

③ $\rho_{X,Y} = 0$ , 则称 $X$ 与 $Y$ **不相关**.

[**注**] 相关系数 $\rho_{X,Y}$ 的概率意义: 描述 $X$ 与 $Y$ 间线性关系的强弱的量, $|\rho_{X,Y}|$ 越大, 则 $X$ 与 $Y$ 间的线性关系越强, 否则越差. $|\rho_{X,Y}| = 1$ 时, $X$ 与 $Y$ 以概率 $1$ 存在线性关系.

[**定理4.5.3**] 随机变量 $X$ 与 $Y$ 相互独立是 $X$ 与 $Y$ 不相关的充分条件.

[**证**] 因 $X$ 与 $Y$ 相互独立, 则 $E(XY) = E(X) \cdot E(Y)$ , 进而 $\mathrm{Cov}(X, Y) = 0$ .

由 $\rho_{X,Y} = \dfrac{\mathrm{Cov}(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$ 即证.

[**注1**] 本定理的必要性不成立, 因为 $X$ 与 $Y$ 不相关只表示 $X$ 与 $Y$ 间无线性关系, 不能推出 $X$ 与 $Y$ 间完全无关, 即相互独立. 对服从二维正态分布的随机变量, 相关与独立等价.

[**注2**]

(1) 证明随机变量 $X$ 与 $Y$ 相互独立的方法:

① 对 $\forall x, y$ , 联合分布函数 $F(x, y) = F_X(x) \cdot F_Y(y)$ .

② 对 $\forall i, j$ , 联合分布律 $p_{i,j} = p_{i\cdot} \cdot p_{\cdot j}$ .

③ 对 $\forall x, y$ , 联合概率密度 $f(x, y) = f_X(x) \cdot f_Y(y)$ .

(2) 证明随机变量 $X$ 与 $Y$ 不相关的方法:

① 相关系数 $\rho_{X,Y} = 0$ .

② 协方差 $\mathrm{Cov}(X, Y) = 0$ .

③ 期望 $E(XY) = E(X) \cdot E(Y)$ . (优先使用)

④ 方差 $D(X + Y) = D(X) + D(Y)$ .

[**定理4.5.4**] 设二维随机变量 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ , 其概率密度
$f(x, y) = \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}} \exp\left\{\dfrac{-1}{2(1 - \rho^2)}\left[\dfrac{(x - \mu_1)^2}{\sigma_1^2} - 2\rho\dfrac{(x - \mu_1)(y - \mu_2)}{\sigma_1\sigma_2} + \dfrac{(y - \mu_2)^2}{\sigma_2^2}\right]\right\}$ , 则:

(1) $X$ 与 $Y$ 的相关系数 $\rho_{X,Y} = \rho$ .

(2) $X$ 与 $Y$ 不相关 iff $X$ 与 $Y$ 独立.

---

---

# 5. 大数定律与中心极限定理

## 5.1 大数定律

[**定义5.1.1**] 设 $\{X_n\}\ \ (n=1,2,\cdots)$ 是一列随机变量序列. 对常数 $a$ , 若对 $\forall \varepsilon>0$ , 有 $\lim\limits_{n\to+\infty}P\{|X_n-a|<\varepsilon\}=1$ , 即事件 $|X_n-a|<\varepsilon$ 几乎必然发生, 则称 $\{X_n\}$ **依概率收敛**于 $a$ , 记作 $X_n\stackrel{P}{\rightarrow}a$ .

[**定理5.1.1**] 对随机变量序列 $\{X_n\},\{Y_n\}$ 和常数 $a,b$ , 若 $X_n\stackrel{P}{\rightarrow}a,Y_n\stackrel{P}{\rightarrow}b$ , 则对 $\forall$ 二元连续函数, 随机变量序列 $g(X_n,Y_n)\stackrel{P}{\rightarrow}g(a,b)$ .

[**定理5.1.2**] [**弱大数定律**, **Khinchin大数定理**] 设随机变量 $X_1,X_2,\cdots$ 独立同分布, 且存在数学期望 $E(X_k)=\mu\ \ (k=1,2,\cdots)$ . 对 $n$ 个变量的算术平均值 $\dfrac{1}{n}\sum\limits_{i=1}^n X_i$ , 对 $\forall\varepsilon>0$ , 有 $\lim\limits_{n\to+\infty}P\left\{\left|\dfrac{1}{n}\sum\limits_{i=1}^n X_i-\mu\right|<\varepsilon\right\}=1$ 或 $\lim\limits_{n\to+\infty}P\left\{\left|\dfrac{1}{n}\sum\limits_{i=1}^n X_i-\mu\right|\geq\varepsilon\right\}=0$ , 或记作 $\dfrac{1}{n}\sum\limits_{k=1}^n X_k\stackrel{P}{\rightarrow}\mu=E(X_k)\ \ (n\to+\infty)$ .

[**定理5.1.3**] [**Bernoulli大数定理**, **弱大数定律的推论**] 设 $f_A$ 为 $n$ 次独立重复试验中事件 $A$ 发生的次数, $p$ 为 $A$ 在每次试验中发生的概率, 则对 $\forall\varepsilon>0$ , 有 $\lim\limits_{n\to+\infty}P\left\{\left|\dfrac{f_A}{n}-p\right|<\varepsilon\right\}=1$ 或 $\lim\limits_{n\to+\infty}P\left\{\left|\dfrac{f_A}{n}-p\right|\geq\varepsilon\right\}=0$ , 或记作 $\dfrac{f_A}{n}\stackrel{P}{\rightarrow}p$ .

[**注1**] 频率的稳定性: 对 $\forall\varepsilon>0$ , 当试验次数 $n$ 充分大时, 事件"频率 $\dfrac{f_A}{n}$ 与概率 $p$ 的偏差 $<\varepsilon$"几乎必然发生.

[**注2**] 由实际推断原理: 实际应用中, 试验次数很大时, 可用事件的频率代替其概率.

---

## 5.2 中心极限定理

[**定理5.2.1**] [**独立同分布的中心极限定理**] 若随机变量 $X_1,\cdots,X_n,\cdots$ 独立同分布, 且期望 $E(X_k)=\mu\ \ (k=1,2,\cdots)$ , 方差 $D(X_k)=\sigma^2\ \ (k=1,2,\cdots)$ , 则随机变量 $\sum\limits_{k=1}^n X_k$ 的标准化变量 $Y_n=\dfrac{\sum\limits_{k=1}^n X_k-E\left(\sum\limits_{k=1}^n X_k\right)}{\sqrt{D\left(\sum\limits_{k=1}^n X_k\right)}}=\dfrac{\sum\limits_{k=1}^n X_k-n\mu}{\sqrt{n}\sigma}$ 的分布函数 $F_n(x)$ 对 $\forall x$ 满足

$\lim\limits_{n\to+\infty}F_n(x)=\lim\limits_{n\to+\infty}P\left\{\dfrac{\sum\limits_{k=1}^n X_k-n\mu}{\sqrt{n}\sigma}\leq x\right\}=\displaystyle\int_{-\infty}^x\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\varPhi(x)$ , 即 $Y_n\stackrel{\text{近似}}{\sim}N(0,1)$ .

**[注]** 本定理的另一形式: $\dfrac{\sum\limits_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} = \dfrac{\dfrac{1}{n}\sum\limits_{k=1}^{n} X_k - \mu}{\dfrac{\sigma}{\sqrt{n}}} = \dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} \overset{\text{近似}}{\sim} N(0,1)$ , 其中 $\overline{X} = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$ 为 $X_1, \cdots, X_n$ 的算术平均值.

因 $Z = \dfrac{\overline{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} \overset{\text{近似}}{\sim} N(0,1)$ , 则 $\overline{X} = \dfrac{\sigma}{\sqrt{n}} Z + \mu \overset{\text{近似}}{\sim} N\left(\mu, \dfrac{\sigma^2}{n}\right)$ .

**[定理5.2.2]** **[De Moivre-Laplace定理]** 若随机变量 $\eta_n \sim b(n,p) \ \ (n = 1, 2, \cdots; 0 < p < 1)$ , 则对 $\forall x$ , 有 $\lim\limits_{n \to +\infty} P\left\{\dfrac{\eta_n - np}{\sqrt{np(1-p)}} \le x\right\} = \displaystyle\int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t = \varPhi(x)$ .

**[注]** 本定理表明: 二项分布的极限分布是正态分布.

随机变量 $X \sim b(n,p)$ 的概率的计算:

① $n \le 5$ 时, 直接计算: $P\{X = k\} = C_n^k \cdot p^k \cdot (1-p)^{n-k} \ \ (k = 0, 1, \cdots, n)$ .

② $n$ 很大, $p$ 很小, 且 $\lambda = np < 10$ 时, 由Poisson定理: $P\{X = k\} \approx \dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!} \ \ (k = 0, 1, \cdots, n)$ .

③ $n \ge 50$ , $p$ 不是很小时, 由**De Moivre-Laplace定理**: $\dfrac{\eta_n - np}{\sqrt{np(1-p)}} \overset{\text{近似}}{\sim} N(0,1)$ .

---

---

# 6. 样本与抽样分布

## 6.1 样本

**[定理6.1.1]** 设 $X_1, \cdots, X_n$ 是取自总体 $X$ 且容量为 $n$ 的样本, 则:

(1) 随机变量 $(X_1, \cdots, X_n)$ 的联合分布函数 $F^*(x_1, \cdots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n) = \prod\limits_{i=1}^{n} F(x_i)$ .

(2) 离散型随机变量的样本 $X_1, \cdots, X_n$ 的联合分布律

$$P\{X_1 = x_1, \cdots, X_n = x_n\} = P\{X_1 = x_1\} \cdots P\{X_n = x_n\} = \prod_{i=1}^{n} P\{X = x_i\} .$$

(3) 连续型随机变量的样本 $X_1, \cdots, X_n$ 的联合概率密度

$$f^*(x_1, \cdots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n) = \prod_{i=1}^{n} f(x_i) .$$

---

## 6.2 统计量与经验分布函数

**[定义6.2.1]** 设 $X_1, \cdots, X_n$ 是取自总体 $X$ 的一个样本, $g(X_1, \cdots, X_n)$ 是 $X_1, \cdots, X_n$ 的一个函数. 若 $g$ 中不含未知参数, 则称 $g(X_1, \cdots, X_n)$ 是一个**统计量**. 若 $x_1, \cdots, x_n$ 是样本 $X_1, \cdots, X_n$ 的样本值, 则称 $g(x_1, \cdots, x_n)$ 是统计量 $g(X_1, \cdots, X_n)$ 的**观察值**.

[**定义6.2.2**] 设 $X_1,\cdots,X_n$ 的取自总体 $X$ 的一个样本, $x_1,\cdots,x_n$ 是样本的观察值.

| 定义 | 统计量 | 观察值 |
| --- | --- | --- |
| **样本均值** | $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ | $\overline{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$ |
| **样本方差** | $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\dfrac{1}{n-1}\left(\sum\limits_{i=1}^{n}X_i^2-n\overline{X}^2\right)$ | $s^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2=\dfrac{1}{n-1}\left(\sum\limits_{i=1}^{n}x_i^2-n\overline{x}^2\right)$ |
| **样本标准差** | $S=\sqrt{S^2}=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2}$ | $s=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}$ |
| **样本 $k$ 阶(原点)矩** | $A_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^k\ \ (k=1,2,\cdots)$ | $a_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^k\ \ (k=1,2,\cdots)$ |
| **样本 $k$ 阶中心矩** | $B_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^k\ \ (k=2,3,\cdots)$ | $b_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i-\overline{x})^k\ \ (k=2,3,\cdots)$ |

[**注**] 注意方差和标准差中的系数为 $\dfrac{1}{n-1}$ 而非 $\dfrac{1}{n}$ .

[**定理6.2.1**] 设 $X_1,\cdots,X_n$ 的取自总体 $X$ 的一个样本. 若 $X$ 的 $k\ \ (k=1,2,\cdots)$ 阶矩 $E(X^k)=\mu_k$ 存在, 则 $n\to+\infty$ 时, 样本的 $k$ 阶矩 $A_k\xrightarrow{P}\mu_k$ .

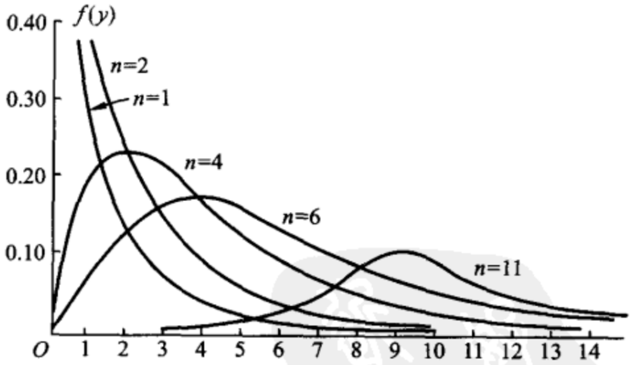
[**定义6.2.3**] 设 $X_1,\cdots,X_n$ 是总体 $F$ 的一个样本. 设 $S(x)$ 为 $X_1,\cdots,X_n$ 中 $\leq x$ 的随机变量的个数, 定义**经验分布函数** $F_n(x)=\dfrac{1}{n}S(x)$ . 对一个样本值, 易得 $F_n(x)$ 的观察值, 仍用 $F_n(x)$ 表示.

## 6.3 抽样分布

[**定义6.3.1**] 统计量的分布称为**抽样分布**. 总体的分布函数已知时, 抽样分布确定.

[**定义6.3.2**] $X_1, \cdots, X_n$ 是取自正态总体 $N(0,1)$ 的样本, 称统计量 $\chi^2 = X_1^2 + \cdots + X_n^2$ 服从自由度为 $n$ 的 $\chi^2$ **分布**或**卡方分布**, 记作 $\chi^2 \sim \chi^2(n)$ , 其中自由度指上式右端包含的独立的随机变量的个数.

[**注1**] $\chi^2(n)$ 分布的概率密度的图象如下图所示:




[**注2**] 若随机变量 $X \sim N(0,1)$ , 则 $X \sim \chi^2(1)$ .

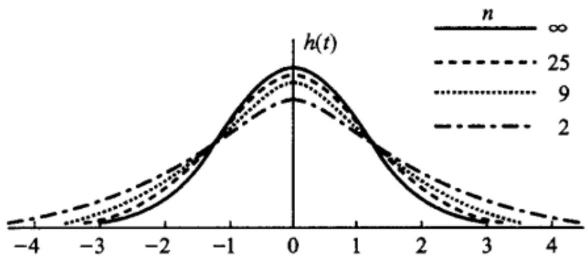
[**定理6.3.1**] [ $\chi^2$ **分布的性质**]

(1) [ $\chi^2$ **分布的可加性**] 若随机变量 $\chi_1^2 \sim \chi^2(n_1), \chi_2^2 \sim \chi^2(n_2)$ 且相互独立, 则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$ .

(2) 若随机变量 $\chi^2 \sim \chi^2(n)$ , 则期望 $E(\chi^2) = n$ , 方差 $D(\chi^2) = 2n$ .

[**定义6.3.3**] 设随机变量 $X \sim N(0,1), Y \sim \chi^2(n)$ 且相互独立, 则称统计量 $t = \dfrac{X}{\sqrt{\dfrac{Y}{n}}}$ 服从自由度为 $n$ 的 $t$ **分布**或**Student分布**, 记作 $t \sim t(n)$ .

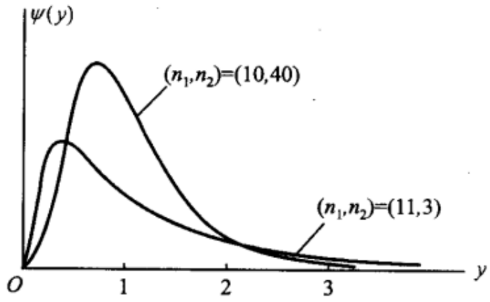
[**注**] $t(n)$ 分布的概率密度的图象如下图所示:




$h(t)$ 的图象关于 $t = 0$ 对称.

**[定义6.3.4]** 设随机变量 $U \sim \chi^2(n_1), V \sim \chi^2(n_2)$ 且相互独立, 则称随机变量 $F = \dfrac{\dfrac{U}{n_1}}{\dfrac{V}{n_2}}$ 服从自由度为 $(n_1, n_2)$ 的 $F$ **分布**, 记作 $F \sim F(n_1, n_2)$ .

**[注]** $F(n_1, n_2)$ 分布的概率密度的图象如下图所示:




**[定理6.3.2]** 若随机变量 $F \sim F(n_1, n_2)$ , 则随机变量 $\dfrac{1}{F} \sim F(n_2, n_1)$ .

---

## 6.4 分位点

**[定义6.4.1]** 设随机变量 $X$ 的概率密度为 $f(x)$ . 若对固定的 $\alpha \in (0, 1)$ , $\exists z_\alpha \in \mathbb{R}\ s.t.$ $P\{X > z_\alpha\} = \displaystyle\int_{z_\alpha}^{+\infty} f(x)\mathrm{d}x = \alpha$ , 则称点 $z_\alpha$ 为 $X$ 的**上 $\alpha$ 分位点**.

**[定义6.4.2]** 设随机变量 $X \sim N(0, 1)$ . 对固定的 $\alpha \in (0, 1)$ , 称 $s.t.\ P\{X > z_\alpha\} = \displaystyle\int_{z_\alpha}^{+\infty} \varphi(x)\mathrm{d}x = \alpha$ 的点 $z_\alpha$ 为 $N(0, 1)$ 的**上 $\alpha$ 分位点**, 其中 $\varphi(x)$ 是 $N(0, 1)$ 的概率密度.

**[定理6.4.1]** 设随机变量 $X \sim N(0, 1)$ . 对固定的 $\alpha \in (0, 1)$ , 设 $z_\alpha$ 是 $X$ 的上 $\alpha$ 分位点, 则 $z_{1-\alpha} = -z_\alpha$ .

**[证]** 由正态分布的概率密度的图象关于 $y$ 轴对称, 结合几何意义即证.

**[定义6.4.3]** 设随机变量 $\chi^2 \sim \chi^2(n)$ . 对固定的 $\alpha \in (0, 1)$ , 称 $s.t.\ P\{\chi^2 > \chi^2_\alpha(n)\} = \displaystyle\int_{\chi^2_\alpha(n)}^{+\infty} f(y)\mathrm{d}y = \alpha$ 的点 $\chi^2_\alpha(n)$ 为 $\chi^2(n)$ 的**上 $\alpha$ 分位点**, 其中 $f(x)$ 是 $\chi^2(n)$ 的概率密度.

**[定义6.4.4]** 设随机变量 $t \sim t(n)$ . 对固定的 $\alpha \in (0, 1)$ , 称 $s.t.\ P\{t > t_\alpha(n)\} = \displaystyle\int_{t_\alpha(n)}^{+\infty} h(t)\mathrm{d}t = \alpha$ 的点 $t_\alpha(n)$ 为 $t(n)$ 的**上 $\alpha$ 分位点**, 其中 $h(t)$ 是 $t(n)$ 的概率密度.

**[定理6.4.2]** 设随机变量 $X \sim t(n)$ . 对固定的 $\alpha \in (0, 1)$ , 设 $t_\alpha$ 是 $X$ 的上 $\alpha$ 分位点, 则 $t_{1-\alpha} = -t_\alpha$ .

**[证]** 由 $t$ 分布的概率密度的图象关于 $y$ 轴对称, 结合几何意义即证.

**[定义6.4.5]** 设随机变量 $F \sim F(n)$ . 对固定的 $\alpha \in (0,1)$ , 称 $s.t.\ P\{F > F_\alpha(n_1,n_2)\} = \int_{F_\alpha(n_1,n_2)}^{+\infty} \psi(y) \mathrm{d}y = \alpha$ 的点 $F_\alpha(n_1,n_2)$ 为 $F(n_1,n_2)$ 的**上 $\alpha$ 分位点**, 其中 $\psi(x)$ 是 $F(n)$ 的概率密度.

**[定理6.4.3]** 设随机变量 $X \sim F(n_1,n_2)$ . 对固定的 $\alpha \in (0,1)$ , 设 $F_\alpha$ 是 $X$ 的上 $\alpha$ 分位点, 则 $F_{1-\alpha}(n_1,n_2) = \dfrac{1}{F_\alpha(n_2,n_1)}$ .

---

## 6.5 正态总体的样本均值与样本方差的分布

**[定理6.5.1]** 设 $X_1,\cdots,X_n$ 是取自总体 $X$ 的样本, 期望 $E(X)=\mu$ 和方差 $D(X)=\sigma^2$ 都存在, 则:

(1) 样本均值的期望 $E\left(\overline{X}\right)=\mu$ .

(2) 样本均值的方差 $D\left(\overline{X}\right)=\dfrac{\sigma^2}{n}$ .

(3) 样本方差的期望 $E(S^2)=\sigma^2$ .

**[注]** 本定理与 $X$ 服从何分布无关, 只需保证期望和方差存在.

**[定理6.5.2]** 设 $X_1,\cdots,X_n$ 是取自总体 $X \sim N(\mu,\sigma^2)$ 的样本, 样本均值为 $\overline{X}$ , 样本方差为 $S^2$ , 则:

(1) $\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ .

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

(3) $\overline{X}$ 与 $S^2$ 相互独立.

(4) $\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$ .

(5) $\displaystyle\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \sim \chi^2(n-1)$ .

[**定理6.5.3**] 设 $X_1,\cdots,X_{n_1}$ 是取自总体 $N(\mu_1,\sigma_1^2)$ 的样本, $Y_1,\cdots,Y_{n_2}$ 是取自总体 $N(\mu_2,\sigma_2^2)$ 的样本, 且这两个样本相互独立. 设两样本的样本方差分别为 $S_1^2=\dfrac{1}{n_1-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2,S_2^2=\dfrac{1}{n_2-1}\sum\limits_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2$ , 则:

(1) $\dfrac{\dfrac{S_1^2}{\sigma_1^2}}{\dfrac{S_2^2}{\sigma^2}}\sim F(n_1-1,n_2-1)$ .

(2) 两总体的方差相同, 即 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时, 有 $\dfrac{\left(\overline{X}-\overline{Y}\right)-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$ , 其中 $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2},S_w=\sqrt{S_w^2}$ .

---

---

# 7. 参数估计

## 7.1 矩估计

[**定义7.1.1**] [**矩估计的步骤**] 设总体 $X$ 的分布中有 $m$ 个未知参数 $\theta_1,\cdots,\theta_m$ .

(1) 求总体的各阶矩 $E(X^k)\ \ (k=1,\cdots,m)$ .

(2) 令样本的各阶矩等于总体的各阶矩, 得到 $m$ 个方程 $\begin{cases}\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i=E(X)\\\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2=E(X^2)\\\cdots\\\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^m=E(X^m)\end{cases}$ .

(3) 上述方程的解 $\hat{\theta}_k(X_1,\cdots,X_m)$ 为 $\theta_k$ 的**矩估计量**, 简称**矩估计**.

[**定理7.1.1**] 设总体 $X$ 的均值 $\mu$ 和方差 $\sigma^2\ \ (\sigma^2>0)$ 都存在且未知. 设 $X_1,\cdots,X_n$ 是取自 $X$ 的样本, 则矩估计 $\begin{cases}\hat{\mu}=\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i\\\widehat{\sigma^2}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2-\overline{X}^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2\end{cases}$ .

---

# 7.2 最大似然估计

[**定义7.2.1**] 求最大似然估计的步骤:

(1) 写出似然函数:

① 若总体为离散型, 则似然函数 $L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)=\prod\limits_{i=1}^{n}P\{X=x_i\}$ .

② 若总体为连续型, 则似然函数 $L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)=\prod\limits_{i=1}^{n}f(x_i;\theta_1,\cdots,\theta_m)$ .

(2) 对似然函数两边取对数, 得到**对数似然函数**:

① 若总体为离散型, 则 $L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)=\sum\limits_{i=1}^{n}\ln P\{X=x_i\}$ .

② 若总体为连续型, 则 $L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)=\sum\limits_{i=1}^{n}\ln f(x_i;\theta_1,\cdots,\theta_m)$ .

(3) 对数似然函数关于各未知参数求偏导, 得对数似然方程: $\begin{cases}\dfrac{\partial}{\partial\theta_1}\ln L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)=0\\ \cdots\\ \dfrac{\partial}{\partial\theta_m}\ln L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)=0\end{cases}$ .

(4) 解对数似然方程, 若有解 $\begin{cases}\theta_1=\theta_1(x_1,\cdots,x_n)\\ \cdots\\ \theta_m=\theta_m(x_1,\cdots,x_n)\end{cases}$ , 则最大似然估计量 $\begin{cases}\hat{\theta}_1=\theta_1(X_1,\cdots,X_n)\\ \cdots\\ \hat{\theta}_m=\theta_m(X_1,\cdots,X_n)\end{cases}$ .

(5) 若对数似然方程无解, 则用单调性直接观察 $L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_m)$ 取得最大值时的 $\theta_i(x_1,\cdots,x_n)\ \ (i=1,\cdots,m)$ .

[**定理7.2.1**] 设 $\hat{\theta}$ 是总体 $X$ 的概率分布中参数 $\theta$ 的最大似然估计. 设参数 $u=u(\theta)$ 有单值反函数 $\theta=\theta(u)$ , 则 $u(\theta)$ 的最大似然估计 $\hat{u}=u(\theta)$ .

---

# 7.3 估计量的评选标准

[**定义7.3.1**] 设 $X_1,\cdots,X_n$ 是取自总体 $X$ 的一组样本, $\hat{\theta}=\theta(X_1,\cdots,X_n)$ 是待估参数 $\theta$ 的估计量. 若 $E\left(\hat{\theta}\right)=\theta$ , 则称 $\hat{\theta}=\theta(X_1,\cdots,X_n)$ 是 $\theta$ 的**无偏估计**.

[**定理7.3.1**] 设总体 $X$ 的均值 $\mu$ 和方差 $\sigma^2>0$ 都未知, 则:

(1) 样本均值 $\overline{X}$ 是总体均值 $\mu$ 的无偏估计.

(2) 样本方差 $S^2$ 是总体方差 $\sigma^2$ 的无偏估计.

(3) 估计量 $\dfrac{1}{n}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ 不是 $\sigma^2$ 的无偏估计.

[**定理7.3.2**] 设总体 $X$ 的 $k\ \ (k\geq 1)$ 阶矩 $\mu_k=E(X^k)$ 存在, $X_1,\cdots,X_n$ 是 $X$ 的一个样本, 则无论总体服从何分布, 样本的 $k$ 阶矩 $A_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^k$ 都是总体的 $k$ 阶矩 $\mu_k$ 的无偏估计量.

[**定义7.3.2**] 设 $X_1,\cdots,X_n$ 是取自总体 $X$ 的一组样本, $\widehat{\theta_1}=\widehat{\theta_1}(X_1,\cdots,X_n)$ 和 $\widehat{\theta_2}=\widehat{\theta_2}(X_1,\cdots,X_n)$ 都在待估参数 $\theta$ 的无偏估计量. 若对 $\forall\theta$ , 都有 $D\left(\widehat{\theta_1}\right)\leq D\left(\widehat{\theta_2}\right)$ , 且至少对某个 $\theta$ , 上式的不等号严格成立, 则称 $\widehat{\theta_1}$ 较 $\widehat{\theta_2}$ 更有效.

[**定义7.3.3**] 设 $X_1,\cdots,X_n$ 是取自总体 $X$ 的一组样本, $\hat{\theta}=\theta(X_1,\cdots,X_n)$ 是待估参数 $\theta$ 的估计量. 若对 $\forall\varepsilon>0$ , 都有 $\lim\limits_{n\to+\infty}P\left\{\left|\hat{\theta}-\theta\right|<\varepsilon\right\}$ , 即 $\hat{\theta}$ 依概率收敛于 $\theta$ , 则称 $\hat{\theta}=\theta(X_1,\cdots,X_n)$ 是 $\theta$ 的**相合估计量**或**一致估计量**.

[**注**] 相合性是对估计量的基本要求, 无相合性的估计量是不可取的.

[**定理7.3.3**]

(1) 样本的 $k\ \ (k\geq 1)$ 阶矩 $A_k$ 是总体的 $k$ 阶矩 $\mu_k=E(X^k)$ 的相合估计量, 即 $A_k\overset{P}{\to}\mu_k=E(X^k)\ \ (n\to+\infty)$ .

(2) 若待估参数 $\theta=g(\mu_1,\cdots,\mu_k)$ , 其中 $g$ 为连续函数, 则 $\theta$ 的矩估计量 $\hat{\theta}=g\left(\widehat{\mu_1},\cdots,\widehat{\mu_k}\right)=g(A_1,\cdots,A_k)$ 是 $\theta$ 的相合估计量.

(3) 样本均值 $\overline{X}$ 是总体均值 $E(X)$ 的相合估计量, 即 $\overline{X}\overset{P}{\to}E(X)\ \ (n\to+\infty)$ .

(4) 样本方差 $S^2$ 是总体方差 $D(X)$ 的相合估计量, 即 $S^2\overset{P}{\to}D(X)\ \ (n\to+\infty)$ .

---

## 7.4 区间估计

一般用区间长度刻画精确度, 可信程度相同时, 区间越短, 精确度越高.

[**定义7.4.1**] 设总体 $X$ 的分布函数为 $F(x;\theta)$ , 其中未知参数 $\theta\in\Theta$ , $\Theta$ 是 $\theta$ 可能取值的范围. 对给定的 $\alpha\in(0,1)$ , 若由取自 $X$ 的一组样本 $X_1,\cdots,X_n$ 确定的两个统计量 $\underline{\theta}=\underline{\theta}(X_1,\cdots,X_n),\overline{\theta}=\overline{\theta}(X_1,\cdots,X_n)$ 满足 $\underline{\theta}<\overline{\theta}$ , 且对 $\forall\theta\in\Theta$ , 都有 $P\left\{\underline{\theta}(X_1,\cdots,X_n)<\theta<\overline{\theta}(X_1,\cdots,X_n)\right\}\geq 1-\alpha$ , 则称随机区间 $\left(\underline{\theta},\overline{\theta}\right)$ 是 $\theta$ 的置信水平为 $(1-\alpha)$ 的**置信区间**, 称 $\underline{\theta}$ 为置信水平为 $(1-\alpha)$ 的双侧置信区间的**置信下限**, 称 $\overline{\theta}$ 为置信水平为 $(1-\alpha)$ 的双侧置信区间的**置信上限**, 称 $(1-\alpha)$ 为**置信水平**.

[**注1**] 置信水平为 $(1-\alpha)$ 的置信区间可能不唯一.

[**注2**] 求待估参数 $\theta$ 的置信区间的方法:

(1) 构造一个与样本 $X_1,\cdots,X_n$ 和 $\theta$ 有关的函数 $W=W(X_1,\cdots,X_n;\theta)$ , 其分布不依赖于 $\theta$ 和其他未知参数, 称有该性质的函数 $W$ 为**枢轴量**. 枢轴量可用点估计的方法构造.

对取自总体 $X\sim N(\mu,\sigma^2)$ 的样本 $X_1,\cdots,X_n$ , 常用的枢轴量:

① $\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$ .

② $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}\sim N(0,1)$ .

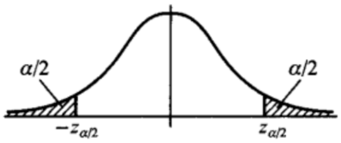
③ $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ .

④ $\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$ .

(2) 对给定的置信水平 $(1-\alpha)$ , 求两个常数 $a, b\ s.t.\ P\{a < W(X_1, \cdots, X_n; \theta) < b\} = 1-\alpha$ . 若能从 $a < W(X_1, \cdots, X_n; \theta) < b$ 中反解得与 $\theta$ 有关的不等式 $\underline{\theta} < \theta < \overline{\theta}$ , 其中 $\underline{\theta} = \underline{\theta}(X_1, \cdots, X_n; \theta), \overline{\theta} = \overline{\theta}(X_1, \cdots, X_n; \theta)$ 都是统计量, 则区间 $\left(\underline{\theta}, \overline{\theta}\right)$ 是 $\theta$ 的一个置信水平为 $(1-\alpha)$ 的置信区间.

$a, b$ 一般取 $W$ 的上分位点, 有如下两种情况:

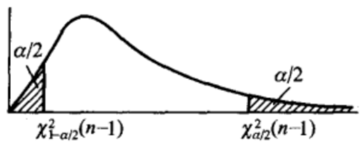
① 概率密度的图象单峰且关于 $y$ 轴对称, 如下图所示:




如 $W \sim N(0,1)$ 或 $W \sim t(n)$ 时, 取关于 $y$ 轴对称的两分位点,

即 $P\{-z_{\frac{\alpha}{2}} < W < z_{\frac{\alpha}{2}}\} = 1-\alpha$ 或 $P\{-t_{\frac{\alpha}{2}} < W < t_{\frac{\alpha}{2}}\} = 1-\alpha$ .

② 概率密度的图象不关于 $y$ 轴对称, 如下图所示:




如 $W \sim \chi^2(n)$ 或 $W \sim F(n_1, n_2)$ 时, 分别取左侧、右侧的面积为 $\dfrac{\alpha}{2}$ 的分位点,

即 $P\left\{\chi^2_{1-\frac{\alpha}{2}} < W < \chi^2_{\frac{\alpha}{2}}\right\} = 1-\alpha, P\left\{F_{1-\frac{\alpha}{2}} < W < F_{\frac{\alpha}{2}}\right\} = 1-\alpha$ .

---

## 7.5 正态总体的均值和方差的区间估计

[**定理7.5.1**] 设 $X_1, \cdots, X_n$ 是取自正态总体 $X \sim N(\mu, \sigma^2)$ 的样本, $\overline{X}$ 和 $S^2$ 分别是样本均值和样本方差.

(1) 总体方差 $\sigma^2$ 已知时, 总体均值 $\mu$ 的置信水平为 $(1-\alpha)$ 的置信区间为 $\left(\overline{X} \pm \dfrac{\sigma}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}}\right)$ .

(2) $\sigma^2$ 未知时, $\mu$ 的置信水平为 $(1-\alpha)$ 的置信区间为 $\left(\overline{X} \pm \dfrac{S}{\sqrt{n}} \cdot t_{\frac{\alpha}{2}}(n-1)\right)$ .

(3) $\mu$ 未知时, $\sigma^2$ 的置信水平为 $(1-\alpha)$ 的置信区间为 $\left(\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$ , 则总体的标准差 $\sigma$ 的置信水平为 $(1-\alpha)$ 的置信区间为 $\left(\dfrac{\sqrt{n-1} \cdot S}{\sqrt{\chi^2_{\frac{\alpha}{2}}(n-1)}}, \dfrac{\sqrt{n-1} \cdot S}{\sqrt{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}\right)$ .

**[定理7.5.2]** 取置信水平为 $(1-\alpha)$ . 设 $X_1,\cdots,X_{n_1}$ 是取自第一个总体 $X\sim N(\mu_1,\sigma_1^2)$ 的一组样本, $Y_1,\cdots,Y_{n_2}$ 是取自第二个总体 $Y\sim N(\mu_2,\sigma_2^2)$ 的一组样本, 且这两个样本相互独立. 设第一、二个总体的样本均值分别为 $\overline{X}$ 和 $\overline{Y}$ , 样本方差分别为 $S_1^2$ 和 $S_2^2$ .

(1) 总体的方差 $\sigma_1^2$ 和 $\sigma_2^2$ 都已知时, 两正态总体的均值差 $(\mu_1-\mu_2)$ 的置信区间为 $\left(\overline{X}-\overline{Y}\pm z_{\frac{\alpha}{2}}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}\right)$ .

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$ , 但 $\sigma^2$ 未知时, $(\mu_1-\mu_2)$ 的置信区间为 $\left(\overline{X}-\overline{Y}\pm t_{\frac{\alpha}{2}}(n_1+n_2-2)\cdot S_w\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\right)$ , 其中 $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}, S_w=\sqrt{S_w^2}$ .

**[注]** 对两正态总体的均值差 $(\mu_1-\mu_2)$ 的置信区间:

① 若置信下限 $>0$ , 则 $\mu_1>\mu_2$ .

② 若置信区间包含 $0$ , 则 $\mu_1$ 较 $\mu_2$ 无显著差别.

**[定理7.5.3]** 取置信水平为 $(1-\alpha)$ . 设 $X_1,\cdots,X_{n_1}$ 是取自第一个总体 $X\sim N(\mu_1,\sigma_1^2)$ 的一组样本, $Y_1,\cdots,Y_{n_2}$ 是取自第二个总体 $Y\sim N(\mu_2,\sigma_2^2)$ 的一组样本, 且这两个样本相互独立. 设第一、二个总体的样本均值分别为 $\overline{X}$ 和 $\overline{Y}$ , 样本方差分别为 $S_1^2$ 和 $S_2^2$ . 总体均值 $\mu_1$ 和 $\mu_2$ 都未知时, 两总体的方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间为 $\left(\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$ .

**[注]** 对两正态总体的方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间, 若置信区间包含 $1$ , 则 $\sigma_1^2$ 较 $\sigma_2^2$ 无显著差别.

---

## 7.6 单侧置信区间

**[定义7.6.1]** 设总体 $X$ 的分布函数为 $F(x;\theta)$ , 其中未知参数 $\theta\in\Theta$ , $\Theta$ 是 $\theta$ 可能取值的范围. 对给定的 $\alpha\in(0,1)$ :

(1) 若由取自 $X$ 的一组样本 $X_1,\cdots,X_n$ 确定的统计量 $\underline{\theta}=\underline{\theta}(X_1,\cdots,X_n)$ , 对 $\forall\theta\in\Theta$ , 都有 $P\left\{\theta>\underline{\theta}(X_1,\cdots,X_n)\right\}\geq 1-\alpha$ , 则称随机区间 $(\underline{\theta},+\infty)$ 为 $\theta$ 的置信水平为 $(1-\alpha)$ 的**单侧置信区间**, 称 $\underline{\theta}$ 为置信水平为 $(1-\alpha)$ 的**单侧置信下限**.

(2) 若由取自 $X$ 的一组样本 $X_1,\cdots,X_n$ 确定的统计量 $\overline{\theta}=\overline{\theta}(X_1,\cdots,X_n)$ , 对 $\forall\theta\in\Theta$ , 都有 $P\left\{\theta<\overline{\theta}(X_1,\cdots,X_n)\right\}\geq 1-\alpha$ , 则称随机区间 $(-\infty,\overline{\theta})$ 为 $\theta$ 的置信水平为 $(1-\alpha)$ 的**单侧置信区间**, 称 $\overline{\theta}$ 为置信水平为 $(1-\alpha)$ 的**单侧置信上限**.

**[定理7.6.1]** 设 $X_1,\cdots,X_n$ 是取自总体 $X\sim N(\mu,\sigma^2)$ 的一组样本, 其中均值 $\mu$ 和方差 $\sigma^2$ 都未知. 取置信水平为 $(1-\alpha)$ , 则:

(1) $\mu$ 的单侧置信下限 $\underline{\mu}=\overline{X}-\dfrac{S}{\sqrt{n}}\cdot t_\alpha(n-1)$ .

(2) $\sigma^2$ 的单侧置信上限 $\overline{\sigma^2}=\dfrac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$ .

---

# 8. 假设检验

## 8.1 假设检验

[**定义8.1.1**]

(1) 考察**假设检验问题**: 在显著性水平 $\alpha$ 下, 检验假设 $H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$ , 其中 $H_0$ 称为**原假设**或**零假设**, $H_1$ 称为**备选假设**.

(2) 假设检验问题的任务: 根据样本, 用检验方法选择接受 $H_0$ 或接受 $H_1$ .

(3) 根据题设和条件确定一个统计量 $Z$ 并在 $H_0$ 成立的条件下确定其分布, 称 $Z$ 为**检验统计量**.

(4) $Z$ 取某区域 $C$ 中的值时拒绝 $H_0$ , 称 $C$ 为**拒绝域**, 拒绝域的边界点称为**临界点**.

(5) 根据有限的样本值判断 $H_0$ 是否成立时, 不可避免地会发生如下两类错误:

① 第一类错误: $\{H_0\text{为真},\text{拒绝}H_0\}$ , 称为**弃真错误**.

② 第二类错误: $\{H_0\text{为假},\text{接受}H_0\}$ , 称为**取伪错误**.

(6) 上述错误无法排除, 只能控制犯错的概率, 此处只考虑控制犯第一类错误的概率, 称为**显著性检验**, 即令 $P\{H_0\text{为真},\text{拒绝}H_0\}\leq\alpha$ , 其中很小的数 $\alpha$ 称为**显著性水平**.

(7) 根据假设的形式, 假设检验分为三类:

① 假设形如 $H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$ 的假设检验称为**双边检验**.

② 假设形如 $H_0:\mu\leq\mu_0,H_1:\mu>\mu_0$ 的假设检验称为**右边检验**.

③假设形如 $H_0:\mu\geq\mu_0,H_1:\mu<\mu_0$ 的假设检验称为**左边检验**.

左边检验和右边检验统称**单边检验**.

[**注**] 原假设和备选假设的选取原则:

① 将大众普遍认为成立的命题作为原假设, 因为原假设不能轻易拒绝, 除非有足够的证据证明它不真.

② 将想证否的命题作为原假设, 将想证真的命题作为备择假设.

## 8.2 正态总体的均值的假设检验

[**定理8.2.1**] 设 $X_1, \cdots, X_n$ 是取自总体 $X \sim N(\mu, \sigma^2)$ 的一组样本, 其中总体均值 $\mu$ 未知. 取显著性水平为 $\alpha$ .

(1) 若总体方差 $\sigma^2$ 已知, 则用 $Z$ **检验** , 即取检验统计量 $Z = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0, 1)$ , 则检验问题和拒绝域如下:

| 检验问题 | 假设 | 拒绝域 |
| --- | --- | --- |
| 双边检验 | $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$ | $\|z\| \geq z_{\frac{\alpha}{2}}$ |
| 右边检验 | $H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$ | $z \geq z_\alpha$ |
| 左边检验 | $H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0$ | $z \leq -z_\alpha$ |

(2) 若 $\sigma^2$ 未知, 则用 $t$ **检验**, 即取检验统计量 $t = \dfrac{\overline{X} - \mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$ , 则检验问题和拒绝域如下:

| 检验问题 | 假设 | 拒绝域 |
| --- | --- | --- |
| 双边检验 | $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$ | $\|t\| \geq t_{\frac{\alpha}{2}}(n-1)$ |
| 右边检验 | $H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$ | $t \geq t_\alpha(n-1)$ |
| 左边检验 | $H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0$ | $t \leq -t_\alpha(n-1)$ |

[**定理8.2.2**] 设 $X_1,\cdots,X_{n_1}$ 和 $Y_1,\cdots,Y_{n_2}$ 分别是取自总体 $X\sim N(\mu_1,\sigma_1^2)$ 和 $Y\sim N(\mu_2,\sigma_2^2)$ 的一组样本, 其中总体均值 $\mu_1$ 和 $\mu_2$ 、总体方差 $\sigma_1^2=\sigma_2^2=\sigma^2$ 都未知. 设这两个样本相互独立, 样本均值分别为 $\overline{X}$ 和 $\overline{Y}$ , 样本方差分别为 $S_1^2$ 和 $S_2^2$ . 取显著性水平为 $\alpha$ , 给定一个常数 $\delta$ , 用 $t$ **检验**, 即取检验统计量 $t=\dfrac{\overline{X}-\overline{Y}-\delta}{S_w\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$ , 则检验问题和拒绝域如下:

| **检验问题** | **假设** | **拒绝域** |
| --- | --- | --- |
| 双边检验 | $H_0:\mu_1-\mu_2=\delta,H_1:\mu_1-\mu_2\neq\delta$ | $\lvert t\rvert\geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$ |
| 右边检验 | $H_0:\mu_1-\mu_2\leq\delta,H_1:\mu_1-\mu_2>\delta$ | $t\geq t_\alpha(n_1+n_2-2)$ |
| 左边检验 | $H_0:\mu_1-\mu_2\geq\delta,H_1:\mu_1-\mu_2<\delta$ | $t\leq -t_\alpha(n_1+n_2-2)$ |

[**注**] 若 $\sigma_1^2$ 和 $\sigma_2^2$ 都已知, 则用 $Z$ **检验**, 即取检验统计量 $Z=\dfrac{\overline{X}-\overline{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1^2}+\dfrac{\sigma_2^2}{n_2^2}}}\sim N(0,1)$ , 则检验问题和拒绝域如下:

| **检验问题** | **假设** | **拒绝域** |
| --- | --- | --- |
| 双边检验 | $H_0:\mu_1-\mu_2=\delta,H_1:\mu_1-\mu_2\neq\delta$ | $\lvert z\rvert\geq z_{\frac{\alpha}{2}}$ |
| 右边检验 | $H_0:\mu_1-\mu_2\leq\delta,H_1:\mu_1-\mu_2>\delta$ | $z\geq z_\alpha$ |
| 左边检验 | $H_0:\mu_1-\mu_2\geq\delta,H_1:\mu_1-\mu_2<\delta$ | $z\leq -z_\alpha$ |

## 8.3 正态总体的方差的假设检验

[**定理8.3.1**] 设 $X_1,\cdots,X_n$ 是取自总体 $X\sim N(\mu,\sigma^2)$ 的一组样本, 其中总体均值 $\mu$ 和总体方差 $\sigma^2$ 都未知. 取显著性水平为 $\alpha$ . 用 $\chi^2$ **检验**, 即取检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$ , 则检验问题和拒绝域如下:

| **检验问题** | **假设** | **拒绝域** |
| --- | --- | --- |
| 双边检验 | $H_0:\sigma^2=\sigma_0^2,H_1:\sigma^2\neq\sigma_0^2$ | $\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n-1)$ 或 $\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n-1)$ |
| 右边检验 | $H_0:\sigma^2\leq\sigma_0^2,H_1:\sigma^2>\sigma_0^2$ | $\chi^2\geq\chi^2_\alpha(n-1)$ |
| 左边检验 | $H_0:\sigma^2\geq\sigma_0^2,H_1:\sigma^2<\sigma_0^2$ | $\chi^2\leq\chi^2_{1-\alpha}(n-1)$ |

[**定理8.3.2**] 设 $X_1,\cdots,X_{n_1}$ 和 $Y_1,\cdots,Y_{n_2}$ 分别是取自总体 $X\sim N(\mu_1,\sigma_1^2)$ 和 $Y\sim N(\mu_2,\sigma_2^2)$ 的一组样本, 其中总体均值 $\mu_1,\mu_2$ 和总体方差 $\sigma_1^2,\sigma_2^2$ 都未知. 设这两个样本相互独立, 样本方差分别为 $S_1^2$ 和 $S_2^2$ . 取显著性水平为 $\alpha$ , 用 $F$ **检验**, 即取检验统计量 $F=\dfrac{S_1^2}{S_2^2}$ , 则检验问题和拒绝域如下:

| 检验问题 | 假设 | 拒绝域 |
| --- | --- | --- |
| 双边检验 | $H_0:\sigma_1^2=\sigma_2^2,H_1:\sigma_1^2\neq\sigma_2^2$ | $F\geq F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$ 或 $F\leq F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$ |
| 右边检验 | $H_0:\sigma_1^2\leq\sigma_2^2,H_1:\sigma_1^2>\sigma_2^2$ | $F\geq F_\alpha(n_1-1,n_2-1)$ |
| 左边检验 | $H_0:\sigma_1^2\geq\sigma_2^2,H_1:\sigma_1^2<\sigma_2^2$ | $F\leq F_{1-\alpha}(n_1-1,n_2-1)$ |