概率论与数理统计

高教社编第四版 单册

解题方法及规律总结

大学考试系列复习资料-07

大学生干货坊 发布

◎知识点目录索引：

一、各个概率公式

1. 加法公式

对于任意两事件 A, B，有

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

2. 减法公式

设 A, B 是两个事件，若 $A \subset B$，则有

$$P(B - A) = P(B) - P(A)，\quad P(B) \geq P(A).$$

$$（P(B - AB) = P(B) - P(AB)）$$

3. 乘法公式

设 $P(A) > 0$，则有

$$P(AB) = P(B \mid A) \cdot P(A).$$

4. 条件概率

设 A, B 是两个事件，且 $P(A) > 0$，称

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

为在事件 A 发生的条件下事件 B 发生的**条件概率**.

5. 独立性

设 A, B 是两个事件，若满足

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

则称事件 A, B **相互独立**，简称 A, B 独立.

6. 全概率公式

设试验 E 的样本区间为 S，A 为 E 的事件，$B_1, B_2, \cdots, B_n$ 为 S 的一个划分，且 $P(B_i) > 0(i = 1, 2, \cdots, n)$，则

$$P(A)=P(A\mid B_1)P(B_1)+P(A\mid B_2)P(B_2)+\cdots+P(A\mid B_n)P(B_n).$$

当$n=2$时，将B_1记为B，B_2即为$\overline{B}$，则**全概率公式**为

$$P(A)=P(A\mid B)P(B)+P(A\mid \overline{B})P(\overline{B}).$$

7. 贝叶斯公式

设试验E的样本区间为S，A为E的事件，$B_1, B_2, \cdots, B_n$为S的一个划分，且$P(A)>0$，$P(B_i)>0(i=1, 2, \cdots, n)$，则

$$P(B_i\mid A)=\frac{P(A\mid B_i)P(B_i)}{\sum_{j=1}^{i}P(A\mid B_j)P(B_j)}\quad (i=1,\ 2,\ \cdots,\ n).$$

当$n=2$时，则**贝叶斯公式**为

$$P(B\mid A)=\frac{P(AB)}{P(A)}=\frac{P(A\mid B)P(B)}{P(A\mid B)P(B)+P(A\mid \overline{B})P(\overline{B})}.$$

二、九大基本分布的数学期望和方差

分布名称		符号	分布律或密度函数	数学期望	方差
离散型	两点分布	$X\sim(0-1)$	$P(X=k)=p^k(1-p)^{1-k},\ k=0,1$	p	$1-p$
	二项分布	$X\sim b(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np	$np(1-p)$
	泊松分布	$X\sim\pi(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\cdots,n$	λ	λ
连续型	均匀分布	$X\sim U(a,b)$	$f(x)=\begin{cases}\dfrac{1}{b-a}, & x\in(a,b)\\ 0, & \text{其他}\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
	指数分布	$X\sim E(\theta)$	$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0\\ 0, & \text{其他}\end{cases}$	θ	θ^2
	正态分布	$X\sim N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
常用统计分布	χ^2分布	$X\sim\chi^2(n)$		n	$2n$
	t分布	$X\sim t(n)$		0 $(n>1)$	$\dfrac{n}{n-2}$ $(n>2)$
	F分布	$X\sim F(n_1,n_2)$		$\dfrac{n_2}{n_1-2}$ $(n_1>2)$	

三、随机变量及其分布

1. 随机变量的分布函数 $F(x)$

定义 设 X 是一个随机变量，x 是任意实数，函数

$$F(x)=P(X\le x),\ -\infty<x<\infty$$

称为 X 的**分布函数**.

对于任意实数 $a,b(a<b)$，有

$$P(a<X\le b)=P(X\le b)-P(X\le a)=F(b)-F(a)$$

$$P(X>a)=1-P(X\le a)=1-F(a).$$

2. 连续型随机变量的概率密度函数 $f(x)$

（1）**定义** 若对随机变量 X 的分布函数 $F(x)$，存在非负可积函数 $f(x)$，使对任意实数 x 有

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

则称 X 为连续型随机变量，$f(x)$ 称为 X 的**概率密度**函数.

（2）性质：

①非负性：$f(x)\ge 0$；

②规范性：$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$；

③对任意实数 $a,b\ (a\le b)$，$P(a<X\le b)=F(b)-F(a)=\int_{a}^{b}f(x)\mathrm{d}x$；

④若 $f(x)$ 在点 x 处连续，则有 $F'(x)=f(x)$；

⑤连续型随机变量在某一点取值的概率为 0，即 $P(X=x_0)=0$.

3. 联合分布函数 $F(x,y)$

$$F(x,y)=P\{(X\le x)\cap(Y\le y)\}\xlongequal{\text{记成}}P(X\le x,Y\le y).$$

性质：

（1）$F(x,y)$ 是关于 x 和 y 的不减函数；

（2）$0 \le F(x,y) \le 1$，且

对于任意固定的 y，$F(-\infty, y) = 0$；

对于任意固定的 x，$F(x, -\infty) = 0$；

$F(-\infty, -\infty) = 0$，$F(\infty, \infty) = 1$.

4. 联合概率密度 $f(x,y)$

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

性质：

（1）$f(x,y) \ge 0$；

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

（3）设 G 是关于 xOy 平面上的区域，点 (X,Y) 落在 G 内的概率为

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

（4）若 $f(x,y)$ 在点 (x,y) 处连续，则

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

5. 边缘分布函数

$$F_X(x) = \int_{-\infty}^{x}\left[\int_{-\infty}^{\infty} f(x,y)\mathrm{d}y\right]\mathrm{d}x .$$

$$F_X(x) = F(x,\infty), \quad F_Y(x) = F(\infty, y).$$

6. 边缘概率密度

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y, \quad f_Y(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x .$$

【例 1】已知概率密度 $f(x,y)$，求 $F(x,y)$.

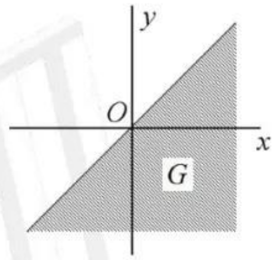
设二维随机变量 (X,Y) 具有概率密度

$$f(x,y)=\begin{cases}2e^{-(2x+y)}, & x>0,\ y>0\\ 0, & \text{其他}\end{cases}.$$

（1）求分布函数 $F(x,y)$；（2）求概率 $P(Y\le X)$.

解：（1）$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(x,y)\mathrm{d}x\mathrm{d}y$

$$=\begin{cases}\int_0^y\int_0^x 2e^{-(2x+y)}\mathrm{d}x\mathrm{d}y, & x>0,\ y>0\\ 0, & \text{其他}\end{cases}$$



即有 $$F(x,y)=\begin{cases}(1-e^{-2x})(1-e^{-y}), & x>0,\ y>0\\ 0, & \text{其他}\end{cases}.$$

（2）将 (X,Y) 看作是平面上随机点的坐标，即有

$$\{Y\le X\}=\{(X,Y)\in G\}.$$

其中 G 为 xOy 平面上直线 $y=x$ 及其下方的部分，如图所示，于是

$$P(Y\le X)=P\{(X,Y)\in G\}=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{\infty}\int_y^{\infty}2e^{-(2x+y)}\mathrm{d}x\mathrm{d}y=\frac{1}{3}.$$

7. 相互独立性

若 $F(x,y)=F_X(x)F_Y(y)$，则称随机变量 X 和 Y 是相互独立的.

等价于 $f(x,y)=f_X(x)f_Y(y)$.

8. 正态分布的 3σ 准则

设 $X\sim N(\mu,\sigma^2)$，则

$$P(|Y-\mu|\le\sigma)=0.6826$$

$$P(|Y-\mu|\le 2\sigma)=0.9544$$

$$P(|Y-\mu|\le 3\sigma)=0.9974$$

可以认为，Y 的取值几乎全部集中在 $[\mu-3\sigma,\ \mu+3\sigma]$ 区间内.

9. 正态近似

当$n \to \infty$时，若$Y \sim B(n, p)$，则

$$\frac{Y - np}{\sqrt{np(1-p)}} \sim N(0,1).$$

四、数学期望和方差

1. 数学期望（均值）

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

（2）连续型：$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x$.

2. 数学期望的性质

设C为常数，

（1）$E(C) = C$.

（2）$E(CX) = CE(X)$.

（3）$E(X+Y) = E(X) + E(Y)$.

（4）若X, Y相互独立，则$E(XY) = E(X)E(Y)$.

3. 方差

简化公式：$D(X) = E(X^2) - [E(X)]^2$.

推论：$E(X^2) = D(X) + [E(X)]^2$.

4. 方差的性质

设C为常数，

（1）$D(C) = 0$.

（2）$D(CX) = C^2 D(X)$，$D(X+C) = D(X)$.

（3）若 X,Y 相互独立，则 $D(X+Y)=D(X)+D(Y)$.

*5. 切比雪夫不等式

设随机变量 X 具有数学期望 $E(X)=\mu$，方差 $D(x)=\sigma^2$，则对于任意正数 ε，不等式

$$P\{|x-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$$

成立，该不等式称为切比雪夫不等式.

6. 协方差及相关系数

（1）协方差

$$\begin{aligned}Cov(X,Y)&=E\{[X-E(X)][Y-E(Y)]\}\\&=E(XY)-E(X)E(Y).\end{aligned}$$

（2）相关系数

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

（3）协方差性质

① $Cov(aX,bY)=abCov(X,Y)$，a,b 为常数.

② $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$.

③ $Cov(X,Y)=Cov(Y,X)$，$Cov(X,X)=D(X)$.

◎注：X 和 Y 相互独立时，$Cov(X,Y)=\rho_{XY}=0$，即 X,Y 不相关. 反之，若 X,Y 不相关，X 和 Y **不一定**相互独立.

五、样本及抽样分布

1. 基本概念

（1）总体：试验的全部可能的**观察值**成为总体.

（2）容量：总体中所包含的个体的**个数**称为总体的容量.

（3）样本：总体 X 中相互独立的随机变量 $X_1, X_2, \cdots, X_n$ 为容量为 n 的简单随机事件.

（4）统计量：样本里的随机变量对应的数称为统计量.

2. 抽样分布定理

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本，$\overline{X}$，S^2 分别是样本均值和样本方差，则有

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

（2）$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，$\overline{X}$ 与 S^2 相互独立.

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

3. χ^2 分布

（1）定义

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(0,1)$ 的样本，则称统计量

$$\chi^2 = X_1^{\ 2} + X_2^{\ 2} + \cdots + X_n^{\ 2}$$

服从自由度为 n 的 χ^2 分布，记为 $\chi^2 \sim \chi^2(n)$.

自由度即上式右端包含独立变量的个数.

（2）性质

①可加性

设 $\chi_1^{\ 2} \sim \chi^2(n_1)$， $\chi_2^{\ 2} \sim \chi^2(n_2)$，且 $\chi_1^{\ 2}$， $\chi_2^{\ 2}$ 相互独立，则有

$$\chi_1^{\ 2} + \chi_2^{\ 2} \sim \chi^2(n_1 + n_2).$$

②数学期望和方差

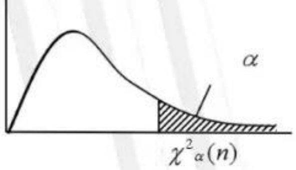
若 $\chi^2 \sim \chi^2(n)$，则有 $E(\chi^2) = n$， $D(\chi^2) = 2n$.

③若 $\chi^2 \sim \chi^2(n)$，则当 $n \to \infty$ 时， $\dfrac{X-n}{\sqrt{2n}}$ （由性质②得）的分布近似为标准正态分布 $N(0,1)$.

④设 $X_1, X_2, \cdots, X_n$ 相互独立，都服从正态分布 $N(\mu, \sigma^2)$，则

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

（3）上分位点



对于给定的正数 α， $0 < \alpha < 1$，满足条件

$$P(\chi^2 > \chi_\alpha^{\ 2}(n)) = \int_{\chi_\alpha^{\ 2}(n)}^{\infty} f(y)\mathrm{d}y = \alpha$$

的点即为 $\chi^2(n)$ 的上 α 分位点，即图中阴影部分.

4. t 分布

（1）定义

设 $X \sim N(0,1)$， $Y \sim \chi^2(n)$，且 X, Y 相互独立，则称随机变量

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

服从自由度为 n 的 t 分布，即为 $t \sim t(n)$.

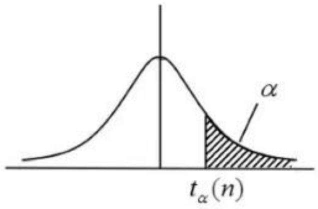
（2）性质

①数学期望和方差

$$E(t) = 0，\quad D(t) = \frac{n}{n-2}\ (n > 2).$$

②当 $n \to \infty$ 时， $t \sim N(0,1)$.

（3）上分位点



对于给定的α，$0<\alpha<1$，满足条件

$$P(t>t_{\alpha}(n))=\int_{t_{\alpha}(n)}^{\infty}h(t)\mathrm{d}t=\alpha$$

的点即为t的上α分位点，即图中阴影部分.

由$h(t)$的对称性可知$t_{1-\alpha}(n)=-t_{\alpha}(n)$.

t分布的上α分位点可查附表 4 获得，在$n>45$时，有正态近似

$$t_{\alpha}(n)\approx z_{\alpha}.$$

5. F 分布

（1）定义

设$U\sim\chi^2(n_1)$，$V\sim\chi^2(n_2)$且U,V相互独立，则称随机变量

$$F=\frac{U/n_1}{V/n_2}$$

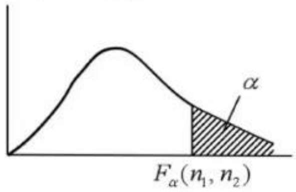
服从自由度为n的F分布，记为$F\sim F(n_1,n_2)$.

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

（2）数学期望

$$E(F)=\frac{n_2}{n_1-2}，\quad n_1>2.$$

（3）上分位点



对于给定的α，$0<\alpha<1$，满足条件

$$P(F>F_{\alpha}(n_1,n_2))=\int_{F_{\alpha}(n_1,n_2)}^{\infty}\varphi(y)\mathrm{d}y=\alpha$$

的点即为$F(n_1,n_2)$的上α分位点，即图中阴影部分.

◎重要性质：$F_{1-\alpha}(n_1,n_2)=\dfrac{1}{F_{\alpha}(n_2,n_1)}$，该式常用来求$F$分布表中未列出的常用的上$\alpha$分位点.

六、参数估计

1. 点估计

（1）矩估计法求矩估计值

设

$$\begin{cases} \mu_1 = \mu_1(\theta_1, \theta_2, \cdots, \theta_n) \\ \mu_2 = \mu_2(\theta_1, \theta_2, \cdots, \theta_n) \\ \quad \vdots \\ \mu_n = \mu_n(\theta_1, \theta_2, \cdots, \theta_n) \end{cases}$$

由该方程组反解出$\theta_1, \theta_2, \cdots, \theta_n$，得到

$$\begin{cases} \theta_1 = \theta_1(\mu_1, \mu_2, \cdots, \mu_n) \\ \theta_2 = \theta_2(\mu_1, \mu_2, \cdots, \mu_n) \\ \quad \vdots \\ \theta_n = \theta_n(\mu_1, \mu_2, \cdots, \mu_n) \end{cases}$$

以A_i分别代替上式中的μ_i，$i = 1, 2, \cdots, n$，就以

$$\hat{\theta}_i = \hat{\theta}_i(A_1, A_2, \cdots, A_n)，\quad i = 1, 2, \cdots, n$$

分别作为θ_i，$i = 1, 2, \cdots, n$的估计量，这种估计量称为**矩估计量**，其观察值称为**矩观察值**. 在后续例题中，一般替换的A_i就是$E(X) = \overline{X}$.

（2）似然函数$L(\theta)$

离散型：$L(\theta) = L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta),\ \theta \in \Theta$.

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

（3）最大似然估计值（量）

离散型：在$\theta \in \Theta$内挑选使似然函数$L(\theta) = \prod_{i=1}^{n} p(x_i; \theta)$达到**最大值**的参数值$\hat{\theta}$，作为参数$\theta$的估计值，则得到的$\hat{\theta}$与样本值$x_1, x_2, \cdots, x_n$有关，故将其记为$\hat{\theta}(x_1, x_2, \cdots, x_n)$，并称其为参数$\theta$的**最大似然估计值**，其对

应的统计量$\hat{\theta}(X_1, X_2, \cdots, X_n)$称为参数$\theta$的**最大似然估计量**.

连续型：使似然函数$L(\theta)=\prod_{i=1}^{n} f(x_i;\theta)$达到**最大值**的$\hat{\theta}$同样与样本值$x_1, x_2, \cdots, x_n$有关，故将其记为$\hat{\theta}(x_1, x_2, \cdots, x_n)$，并称其为参数$\theta$的**最大似然估计值**，其对应的统计量$\hat{\theta}(X_1, X_2, \cdots, X_n)$称为参数$\theta$的**最大似然估计量**.

已知似然函数$L(\theta)$后，可通过**直接求导**，即令$\frac{\mathrm{d}}{\mathrm{d}\theta}L(\theta)=0$求出$\hat{\theta}$，或通过**对数似然方程**，即令$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta)=0$求出$\hat{\theta}$.

【例 2】求矩估计量、最大似然估计量.

设总体X的概率密度为$f(x)=\begin{cases}\theta x^{\theta-1}, 0<x<1 \\ 0, \quad 其他\end{cases}$，其中$\theta>0$是未知参数，$X_1, X_2, \cdots, X_n$为来自总体的一个简单样本$x_1, x_2, \cdots, x_n$为样本值，求$\theta$的矩估计量和最大似然估计量.

解：（1）$\mu_1=E(X)=\int_{-\infty}^{\infty} xf(x)\mathrm{d}x=\int_0^1 \theta x^{\theta}\mathrm{d}x=\frac{\theta}{\theta+1}$，反解得$\theta=\frac{\mu_1}{1-\mu_1}$，

用$\overline{X}(=E(x))$替换，于是θ得矩估计量$\hat{\theta}=\frac{\overline{X}}{1-\overline{X}}$.

（2）似然函数$L(x_1, x_2, \cdots, x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\theta^n\left(\prod_{i=1}^{n} x_i\right)^{\theta-1}$.

两边取对数得$\ln L=n\ln\theta+(\theta-1)\sum_{i=1}^{n}\ln x_i$，令$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L=\frac{\theta}{n}+\sum_{i=1}^{n}\ln x_i=0$得

θ的最大似然估计值和估计量分别为$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln x_i}$，$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln X_i}$.

2. 估计量的评选标准

（1）无偏性

若估计量$\hat{\theta}=\hat{\theta}(X_1, X_2, \cdots, X_n)$的数学期望$E(\hat{\theta})$存在，且对于任意$\theta \in \Theta$有$E(\hat{\theta})=\theta$，则称$\hat{\theta}$是$\theta$的无偏估计量.

◎注：θ即为待估计参数，$\hat{\theta}$即为由样本$X_1, X_2, \cdots, X_n$确定的估计量，用来估计θ.

（2）有效性

设$\hat{\theta}_1=\hat{\theta}_1(X_1, X_2, \cdots, X_n)$与$\hat{\theta}_2=\hat{\theta}_2(X_1, X_2, \cdots, X_n)$都是$\theta$的无偏估计量，若对于任意$\theta \in \Theta$，有

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

且至少对于某一个$\theta \in \Theta$使上式中的不等号成立，则称$\hat{\theta}_1$较$\hat{\theta}_2$有效. 即**“方差较小者更有效”**.

3. 区间估计（置信区间）

（1）置信区间的定义

置信区间是指由样本统计量所构造的总体.

在统计学中，一个概率样本的置信区间是对这个样本的某个总体参数的区间估计.

对于给定值α（$0<\alpha<1$），若由来自X的样本$X_1, X_2, \cdots, X_n$确定的两个统计量$\underline{\theta}=\underline{\theta}(X_1, X_2, \cdots, X_n)$和$\overline{\theta}=\overline{\theta}(X_1, X_2, \cdots, X_n)$（$\underline{\theta}<\overline{\theta}$），对于任意$\theta \in \Theta$满足

$$P(\underline{\theta}<\theta<\overline{\theta}) \ge 1-\alpha$$

则$(\underline{\theta}<\overline{\theta})$是$\theta$的置信水平为$1-\alpha$的**置信区间**，$\underline{\theta}$和$\overline{\theta}$分别称为置信下限

和置信上限，$1-\alpha$ 称为**置信水平**（置信度）.

（2）置信区间的意义

置信区间展现的是这个参数的真实值有一定概率落在测量结果周围的程度，其给出的是被测量参数的测量值的可信程度.

（3）窄的置信区间比宽的置信区间能提供更多有关总体参数的信息.

假设全班考试的平均分数为 65 分，则

置信区间	间 隔	宽窄度	表达的意思
0-100 分	100	宽	等于什么也没告诉你
30-80 分	50	较窄	你能估出大概的平均分了（55 分）
60-70 分	10	窄	你几乎能判定全班的平均分了（65 分）

（4）枢轴量

若 $W=W(X_1,X_2,\cdots,X_n;\theta)$ 的分布不依赖于 θ 以及其他未知参数，称具有这样性质的函数 W 为**枢轴量**.

七、假设检验

处理参数的假设检验问题的步骤如下：

1. 根据实际问题的要求，提出原假设 H_0 及备择假设 H_1；

2. 给定显著性水平 α 以及样本容量 n；

3. 确定检验统计量以及拒绝域的形式；

4. 按 $P\{当H_0为真拒绝H_0\} \le \alpha$ 求出拒绝域；

5. 取样，根据样本观察值作出决策，是接受 H_0 还是拒绝 H_0.

【例 3】假设检验问题.

一个小学校长在报纸上看到这样的报道："这一城市的初中学生平均每周看 8 小时电视"。她认为她所领导的学校，学生看电视的时间明显小于该数字. 为此，她向 100 个学生做了调查，得知平均每周看电视的时间 $\bar{x} = 6.5$ 小时，样本标准差为 $s = 2$ 小时. 问是否可以认为这位校长的看法是对的？取 $\alpha = 0.05$.

解：（1）提出假设 H_0： $\mu \le 8$； H_1： $\mu > 8$.

（2）当 n 充分大时，$\dfrac{\bar{x}-\mu}{s/\sqrt{n}}$ 近似的服从 $N(0,1)$ 分布.

（3） H_0 的拒绝域近似为 $\dfrac{\bar{x}-\mu}{s/\sqrt{n}} \ge z_\alpha$.

（4） $n = 100,\ \alpha = 0.05,\ \bar{x} = 6.5,\ S = 2$，由计算知

$$|t| = \left|\frac{6.5-8}{2/\sqrt{100}}\right| = 7.5 \ge z_{0.05} = 1.65.$$

（5）故在 $\alpha = 0.05$ 下，拒绝 H_0，即认为校长的看法是不对的.