$$Z=\frac{\sqrt{2}(Y_1-Y_2)}{S}=\frac{U}{\sqrt{\chi^2/2}}$$

服从自由度为 2 的 t 分布.

例6.12 设$X_1,X_2,\cdots,X_n$ $(n>2)$为来自总体 $N(0,1)$的简单随机样本，$\overline{X}$ 为样本均值，记 $Y_i=X_i-\overline{X}$，$i=1,2,\cdots,n$.

求：(Ⅰ) Y_i 的方差 $D(Y_i)$，$i=1,2,\cdots,n$；

(Ⅱ) Y_1 与 Y_n 的协方差 $\mathrm{Cov}(Y_1,Y_n)$.

[提示] 本题考查的是方差与协方差的计算，本质上与统计无关.

[解] (Ⅰ) $D(Y_i)=D(X_i-\overline{X})$

$$=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum_{\substack{k=1\\k\neq i}}^{n}X_k\right]=\frac{n-1}{n}\quad(i=1,2,\cdots,n).$$

(Ⅱ) $\mathrm{Cov}(Y_1,Y_n)=E[Y_1-E(Y_1)][Y_n-E(Y_n)]=E(X_1-\overline{X})(X_n-\overline{X})$

$$=E(X_1X_n)+E(\overline{X}^2)-E(X_1\overline{X})-E(X_n\overline{X})$$

$$=E(X_1)E(X_n)+D(\overline{X})-\frac{1}{n}E(X_1^2)-\frac{1}{n}\sum_{k=2}^{n}E(X_1X_k)-\frac{1}{n}E(X_n^2)-\frac{1}{n}\sum_{k=1}^{n-1}E(X_kX_n)$$

$$=-\frac{1}{n}.$$

七、参数估计

·考试内容与要求·

考试内容

点估计的概念　估计量与估计值　矩估计法　最大似然估计法　估计量的评选标准　区间估计的概念　单个正态总体的均值和方差的区间估计　两个正态总体的均值差和方差比的区间估计

考试要求

1. 理解参数的点估计、估计量与估计值的概念.

2. 掌握矩估计法(一阶、二阶距)和最大似然估计法.

3. 了解估计量的无偏性、有效性(最小方差性)和一致性(相合性)的概念，并会验证估计量的无偏性.

4. 理解区间估计的概念，会求单个正态总体的均值和方差的置信区间，会求两个正态总体的均值差和方差比的置信区间.

·考试内容解析·

(一) 点估计

1. 点估计的概念

设总体 X 的分布形式已知，但含有未知参数 θ；或者总体的某数字特征(例如数学期望或方差)存在但未知，从总体 X 中抽取样本$X_1,X_2,\cdots,X_n$，相应的样本值为$x_1,x_2,\cdots,x_n$. 借助于样本来估计未知参数的值的问题就是参数的点估计问题. 要解决点估计问题，就是要构造一个统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$，用它来估计未知参数 θ，用它的观测值$\hat{\theta}(x_1,x_2,\cdots,x_n)$作为未知参数 θ 的近似值. 称 $\hat{\theta}(X_1,X_2,\cdots,X_n)$为未知参数 θ 的估计量，$\hat{\theta}(x_1,x_2,\cdots,x_n)$为 θ 的估计值. 在不至于混淆的情况下将 θ 的估计量和估计值统称为θ 的估计.

求点估计有两种常用的方法——矩估计法和最大似然估计法.

2. 矩估计法

设总体 X 为连续型随机变量，概率密度为 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$，或设总体 X 为离散型随机变量，概率分

布为

$$P\{X=x_i\}=p(x_i;\theta_1,\theta_2,\cdots,\theta_k),i=1,2,\cdots,$$

其中$\theta_1,\theta_2,\cdots,\theta_k$是待估参数．假设总体$X$的$l$阶原点矩

$$\mu_l=E(X^l)=\int_{-\infty}^{+\infty}x^lf(x;\theta_1,\theta_2,\cdots,\theta_k)\,\mathrm{d}x$$

或

$$\mu_l=E(X^l)=\sum_{i=1}^{\infty}x_i^lp(x_i;\theta_1,\theta_2,\cdots,\theta_k),\quad l=1,2,\cdots,k$$

均存在，一般来说，它们是$\theta_1,\theta_2,\cdots,\theta_k$的函数，记作

$$\mu_l=\mu_l(\theta_1,\theta_2,\cdots,\theta_k).$$

设$X_1,X_2,\cdots,X_n$是来自总体X的样本，根据辛钦大数定律可知样本l阶原点矩

$$A_l=\frac{1}{n}\sum_{i=1}^{n}X_i^l$$

依概率收敛于总体l阶原点矩$\mu_l\ (l=1,2,\cdots,k)$．根据依概率收敛的性质可知样本矩的连续函数依概率收敛于相应的总体矩的同一连续函数．因此，我们用样本矩作为相应的总体矩的估计量，用样本矩的连续函数作为相应的总体矩的连续函数的估计量，由此求得未知参数$\theta_1,\theta_2,\cdots,\theta_k$的估计量．这种求估计量的方法称为矩估计法．

矩估计法的具体做法是：求出总体X的$1,2,\cdots,k$阶原点矩

$$\mu_l=\mu_l(\theta_1,\theta_2,\cdots,\theta_k),\quad l=1,2,\cdots,k;$$

解此方程组，求得未知参数$\theta_1,\theta_2,\cdots,\theta_k$关于$\mu_1,\mu_2,\cdots,\mu_k$的函数表达式

$$\theta_j=\theta_j(\mu_1,\mu_2,\cdots,\mu_k),\quad j=1,2,\cdots,k;$$

用样本矩A_l代替总体矩$\mu_l\ (l=1,2,\cdots,k)$，得到未知参数$\theta_j\ (j=1,2,\cdots,k)$的矩估计量

$$\hat{\theta}_j=\theta_j(A_1,A_2,\cdots,A_k),\quad j=1,2,\cdots,k,$$

矩估计量的观测值

$$\hat{\theta}_j=\theta_j(a_1,a_2,\cdots,a_k),\quad j=1,2,\cdots,k$$

就是未知参数$\theta_j\ (j=1,2,\cdots,k)$的矩估计值．

3．最大似然估计法

(1) 设总体X是离散型随机变量，概率分布为

$$P\{X=t_i\}=p(t_i;\theta),\quad i=1,2,\cdots,$$

其中θ是待估参数，θ的取值范围是Θ．设$X_1,X_2,\cdots,X_n$是来自总体X的样本，$x_1,x_2,\cdots,x_n$是样本值，称函数

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

为样本$x_1,x_2,\cdots,x_n$的似然函数．如果$\hat{\theta}\in\Theta$使得

$$L(\hat{\theta})=\max_{\theta\in\Theta}L(\theta),$$

这样的$\hat{\theta}$与$x_1,x_2,\cdots,x_n$有关，记作$\hat{\theta}(x_1,x_2,\cdots,x_n)$，称为未知参数$\theta$的最大似然估计值，相应的统计量$\hat{\theta}(X_1,X_2,\cdots,X_n)$称为$\theta$的最大似然估计量．一般统称为$\theta$的最大似然估计．

(2) 设总体X是连续型随机变量，概率密度为$f(x;\theta)$，θ是待估参数，则似然函数为

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

(3) 如果$L(\theta)$或$\ln L(\theta)$关于θ可微，则θ往往可以从方程

$$\frac{\mathrm{d}L(\theta)}{\mathrm{d}\theta}=0 \quad 或 \quad \frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=0$$

中解得.

(4) 如果总体 X 的分布中含有 k 个未知参数 $\theta_1, \theta_2, \cdots, \theta_k$. 则似然函数是这些参数的函数: $L=L(\theta_1, \theta_2, \cdots, \theta_k)$. 分别令

$$\frac{\partial L}{\partial \theta_j}=0 \text{ 或 } \frac{\partial \ln L}{\partial \theta_j}=0, \quad j=1,2,\cdots,k,$$

解上面的方程组, 可得各个待估参数 θ_j 的最大似然估计值 $\hat{\theta}_j(x_1,x_2,\cdots,x_n)$, 从而得到最大似然估计量 $\hat{\theta}_j(X_1,X_2,\cdots,X_n)(j=1,2,\cdots,k)$.

(5) 最大似然估计具有如下性质:设 θ 的函数 $\varphi=\varphi(\theta)(\theta\in\Theta)$ 具有单值反函数 $\theta=\theta(\varphi)(\varphi\in\Phi)$, $\hat{\theta}$ 是总体 X 的分布中未知参数 θ 的最大似然估计.则 $\hat{\varphi}=\varphi(\hat{\theta})$ 是 $\varphi=\varphi(\theta)$ 的最大似然估计.例如,设总体方差 σ^2 的最大似然估计为 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$,则总体标准差的最大似然估计为 $\hat{\sigma}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}$.

4. 估计量的评选标准

(1) 无偏性　如果 θ 的估计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 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 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 是未知参数 θ 的无偏估计量.

(2) 有效性　设 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$ 都是未知参数 θ 的无偏估计量, 如果对于任意 $\theta\in\Theta$, 有

$$D(\hat{\theta}_1)\leqslant D(\hat{\theta}_2),$$

且至少有一个 $\theta\in\Theta$, 使上式中的不等号成立, 则称 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 比 $\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$ 更有效.

(3) 一致性(相合性)　设 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为未知参数 θ 的估计量, 如果对于任意 $\theta\in\Theta$, 当 $n\to\infty$ 时, $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 依概率收敛于 θ, 则称 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为未知参数 θ 的一致估计量或相合估计量.

(二) 区间估计

1. 置信区间

设总体 X 的分布函数为 $F(x;\theta)$, 其中 θ 是未知参数, 从总体 X 中抽取样本 $X_1, X_2, \cdots, X_n$, 对于给定的 α $(0<\alpha<1)$, 如果两个统计量 $\theta_1=\theta_1(X_1,X_2,\cdots,X_n)$ 和 $\theta_2=\theta_2(X_1,X_2,\cdots,X_n)$, 对于任意 $\theta\in\Theta$ 满足

$$P\{\theta_1(X_1,X_2,\cdots,X_n)<\theta<\theta_2(X_1,X_2,\cdots,X_n)\}=1-\alpha,$$

则称随机区间 (θ_1, θ_2) 为未知参数 θ 的置信水平是 $1-\alpha$ 的置信区间. $\theta_1(X_1,X_2,\cdots,X_n)$ 和 $\theta_2(X_1,X_2,\cdots,X_n)$ 分别称为这一双侧置信区间的置信下限和置信上限, $1-\alpha$ 称为置信水平(或置信度).

2. 单个正态总体均值和方差的置信区间

设 $X\sim N(\mu,\sigma^2)$, 从总体 X 中抽取样本 $X_1,X_2,\cdots,X_n$, 样本均值为 $\overline{X}$, 样本方差为 S^2.

(1) σ^2 已知, μ 的置信水平是 $1-\alpha$ 的置信区间为

$$\left(\overline{x}-\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}},\ \overline{x}+\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\right).$$

(2) σ^2 未知, μ 的置信水平是 $1-\alpha$ 的置信区间为

$$\left(\overline{x}-\frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\ \overline{x}+\frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right).$$

(3) μ 已知，σ^2 的置信水平是 $1-\alpha$ 的置信区间为

$$\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)},\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right].$$

(4) μ 未知，σ^2 的置信水平是 $1-\alpha$ 的置信区间为

$$\left[\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right].$$

标准差 σ 的置信水平是 $1-\alpha$ 的置信区间为

$$\left[\frac{\sqrt{n-1}s}{\sqrt{\chi^2_{\frac{\alpha}{2}}(n-1)}},\frac{\sqrt{n-1}s}{\sqrt{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}\right].$$

3. 两个正态总体均值差和方差比的置信区间

设 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，从总体 X 中抽取样本$X_1,X_2,\cdots,X_{n_1}$，样本均值为$\overline{X}$，样本方差为 S_1^2；从总体 Y 中抽取样本$Y_1,Y_2,\cdots,Y_{n_2}$，样本均值为$\overline{Y}$，样本方差为 S_2^2，并且两个样本是相互独立的.

(1) σ_1^2 和 σ_2^2 已知，$\mu_1-\mu_2$ 的置信水平是 $1-\alpha$ 的置信区间为

$$\left(\bar{x}-\bar{y}-u_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ \bar{x}-\bar{y}+u_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right).$$

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$ 未知，$\mu_1-\mu_2$ 的置信水平是 $1-\alpha$ 的置信区间为

$$\left[\bar{x}-\bar{y}-t_{\frac{\alpha}{2}}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}-\bar{y}+t_{\frac{\alpha}{2}}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right],$$

其中

$$s_w^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2},\quad s_w=\sqrt{s_w^2}.$$

(3) μ_1 和 μ_2 已知，$\frac{\sigma_1^2}{\sigma_2^2}$的置信水平是 $1-\alpha$ 的置信区间为

$$\left[\frac{n_2\sum_{i=1}^{n_1}(x_i-\mu_1)^2}{n_1\sum_{j=1}^{n_2}(y_j-\mu_2)^2}\cdot\frac{1}{F_{\frac{\alpha}{2}}(n_1,n_2)},\frac{n_2\sum_{i=1}^{n_1}(x_i-\mu_1)^2}{n_1\sum_{j=1}^{n_2}(y_j-\mu_2)^2}\cdot F_{\frac{\alpha}{2}}(n_2,n_1)\right].$$

(4) μ_1 和 μ_2 未知，$\frac{\sigma_1^2}{\sigma_2^2}$的置信水平是 $1-\alpha$ 的置信区间为

$$\left(\frac{s_1^2}{s_2^2}\cdot\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\frac{s_1^2}{s_2^2}\cdot F_{\frac{\alpha}{2}}(n_2-1,n_1-1)\right).$$

•例 题 详 解•

例 7.1 已知一批零件的长度 X (单位 cm)服从正态分布 $N(\mu,1)$，从中随机地抽取 16 个零件，得到长度的平均值为 40 (cm)．则 μ 的置信度为 0.95 的置信区间是________．(注:标准正态分布函数值 $\Phi(1.96)=0.975$，$\Phi(1.645)=0.95$)

[答案] (39.51,40.49).

[提示] 本题考查分位数概念及置信区间的求法.

[解] 记 $\bar{x}$ 为样本均值，置信度为 0.95 (即 $\alpha=0.05$)的双侧置信区间为

$$\left(\overline{x}-\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}},\overline{x}+\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\right).$$

由于 $u_{\frac{\alpha}{2}}=u_{0.025}$，$1-0.025=0.975=\Phi(1.96)$，所以 $u_{0.025}=1.96$．数据代入，得置信区间为

$$\left(40-\frac{1}{\sqrt{16}}\times 1.96,40+\frac{1}{\sqrt{16}}\times 1.96\right)=(39.51,40.49).$$

［典型错误］ 将 $u_{\frac{\alpha}{2}}$写成 $u_{\alpha}=u_{0.05}$，从而 $1-0.05=0.95=\Phi(1.645)$，得到

$$\left(\overline{x}-\frac{\sigma}{\sqrt{n}}\times 1.645,\overline{x}+\frac{\sigma}{\sqrt{n}}\times 1.645\right)=(39.59,40.41).$$

例 7.2 设总体 X 的概率密度为

$$f(x;\theta)=\begin{cases}e^{-(x-\theta)}, & 若\ x\geqslant\theta,\\ 0, & 若\ x<\theta;\end{cases}$$

而 X_1，X_2，…，X_n 是来自总体 X 的简单随机样本，则未知参数 θ 的矩估计量为________.

［答案］ $\frac{1}{n}\sum_{i=1}^{n}X_i-1$ 或 $\overline{X}-1$.

［提示］ 本题考查矩估计量的概念与计算．先求期望 $E(X)$.

［解］
$$E(X)=\int_{-\infty}^{+\infty}xf(x;\theta)\,dx=\int_{\theta}^{+\infty}xe^{-(x-\theta)}dx=\theta+1.$$

根据矩估计量的定义，满足 $E(X)=\overline{X}$ 的 $\hat{\theta}$ 即为 θ 的矩估计量．因此 $\hat{\theta}=\overline{X}-1$.

例 7.3 设一批零件的长度服从正态分布 $N(\mu,\sigma^2)$，其中 μ，σ^2 均未知．现从中随机抽取 16 个零件，测得样本均值 $\overline{x}=20$ (cm)，样本标准差 $s=1$ (cm)，则 μ 的置信度为 0.90 的置信区间是

(A) $\left(20-\frac{1}{4}t_{0.05}(16),20+\frac{1}{4}t_{0.05}(16)\right)$.　　(B) $\left(20-\frac{1}{4}t_{0.1}(16),20+\frac{1}{4}t_{0.1}(16)\right)$.

(C) $\left(20-\frac{1}{4}t_{0.05}(15),20+\frac{1}{4}t_{0.05}(15)\right)$.　　(D) $\left(20-\frac{1}{4}t_{0.1}(15),20+\frac{1}{4}t_{0.1}(15)\right)$.

［答案］ (C).

［提示］ 本题考查置信度、置信区间、分位数等概念．按定义不难得到正确选项．

［解］ 由于 μ 的置信度为 $1-\alpha$ 的置信区间为$\left(\overline{x}\pm\frac{s}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$，其中 $1-\alpha=0.90$，$\alpha=0.10$，$n=16$，$\overline{x}=20$，$s=1$，故应选(C).

例 7.4 设总体 X 的概率密度为

$$f(x)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1,\\ 0, & 其他,\end{cases}$$

其中 $\theta>-1$ 是未知参数，X_1，X_2，…，X_n 是来自总体 X 的一个容量为 n 的简单随机样本，分别用矩估计法和极大似然估计法求 θ 的估计量．

［提示］ 利用矩估计法和极大似然估计法的基本原理进行计算．

［解］ 总体 X 的数学期望为

$$E(X)=\int_{-\infty}^{+\infty}xf(x)\,dx=\int_{0}^{1}(\theta+1)x^{\theta+1}dx=\frac{\theta+1}{\theta+2}.$$

设 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 为样本均值．令

$$\frac{\theta+1}{\theta+2}=\overline{X},$$

解得未知参数 θ 的矩估计量为

$$\hat{\theta} = \frac{2\overline{X} - 1}{1 - \overline{X}}.$$

设 $x_1, x_2, \cdots, x_n$ 是相应于样本 $X_1, X_2, \cdots, X_n$ 的样本值. 则似然函数为

$$L = \begin{cases} (\theta + 1)^n \left(\prod_{i=1}^{n} x_i\right)^{\theta}, & 0 < x_i < 1\ (i = 1,2,\cdots,n), \\ 0. & \text{其他}. \end{cases}$$

当 $0 < x_i < 1\ (i = 1,2,\cdots,n)$时，$L > 0$. 且

$$\ln L = n\ln(\theta + 1) + \theta\sum_{i=1}^{n}\ln x_i, \quad \frac{\mathrm{d}\ln L}{\mathrm{d}\theta} = \frac{n}{\theta + 1} + \sum_{i=1}^{n}\ln x_i.$$

令

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

解得 θ 的最大似然估计值为

$$\hat{\theta} = -1 - \frac{n}{\sum_{i=1}^{n}\ln x_i}.$$

从而得 θ 的最大似然估计量为

$$\hat{\theta} = -1 - \frac{n}{\sum_{i=1}^{n}\ln X_i}.$$

[典型错误] (1)$E(X)$不会求；(2)矩估计和最大似然估计概念不清；(3)似然函数 $L = \prod_{i=1}^{n} f(x_i)$ 中 x 的下标不写，写成 $L = \prod_{i=1}^{n} f(x)$.

例 7.5　设总体 X 的概率密度为

$$f(x) = \begin{cases} \frac{6x}{\theta^3}(\theta - x), & 0 < x < \theta. \\ 0, & \text{其他}. \end{cases}$$

$X_1, X_2, \cdots, X_n$ 是取自总体 X 的简单随机样本.

(1) 求 θ 的矩估计量 $\hat{\theta}$；

(2) 求 $\hat{\theta}$的方差 $D(\hat{\theta})$.

[提示] 求 θ 的矩估计量 $\hat{\theta}$，只要令 $E(X) = \overline{X}$ 即可. 求 $\hat{\theta}$ 的方差需计算 $D(X)$. 本题实质上是计算随机变量的均值和方差问题.

[解] (1) $E(X) = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = \int_0^{\theta} \frac{6x^2}{\theta^3}(\theta - x)\,\mathrm{d}x = \frac{\theta}{2}$.

记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，令 $\frac{\theta}{2} = \overline{X}$. 得 θ 的矩估计量为

$$\hat{\theta} = 2\overline{X}.$$

(2) 由于

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^{\theta} \frac{6x^3}{\theta^3}(\theta - x)\,\mathrm{d}x = \frac{6\theta^2}{20},$$

$$D(X) = E(X^2) - [E(X)]^2 = \frac{6\theta^2}{20} - \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{20},$$

所以　$\hat{\theta} = 2\overline{X}$ 的方差为

$$D(\hat{\theta}) = D(2\overline{X}) = 4D(\overline{X}) = \frac{4}{n}D(X) = \frac{\theta^2}{5n}.$$

例 7.6　设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta\left(0<\theta<\frac{1}{2}\right)$是未知参数，利用总体 X 的如下样本值

$$3, 1, 3, 0, 3, 1, 2, 3,$$

求 θ 的矩估计值和最大似然估计值.

［提示］　本题考查：矩估计值与最大似然估计值的计算，属基本题，没有什么综合性.

［解］
$$E(X) = 0\times\theta^2 + 1\times 2\theta(1-\theta) + 2\times\theta^2 + 3\times(1-2\theta) = 3-4\theta.$$

$$\bar{x} = \frac{1}{8}(3+1+3+0+3+1+2+3) = 2.$$

令 $E(X) = \bar{x}$，即 $3-4\theta = 2$. 得 θ 的矩估计值 $\hat{\theta} = \frac{1}{4}$.

对于给定的样本值，似然函数为

$$L(\theta) = \theta^2[2\theta(1-\theta)]^2\theta^2(1-2\theta)^4 = 4\theta^6(1-\theta)^2(1-2\theta)^4.$$

$$\ln L(\theta) = \ln 4 + 6\ln\theta + 2\ln(1-\theta) + 4\ln(1-2\theta).$$

$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta} = \frac{6}{\theta} - \frac{2}{1-\theta} - \frac{8}{1-2\theta} = \frac{6-28\theta+24\theta^2}{\theta(1-\theta)(1-2\theta)},$$

令$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta} = 0$，解得 $\theta_{1,2} = \frac{1}{12}(7\pm\sqrt{13})$. $\frac{1}{12}(7+\sqrt{13}) > \frac{1}{2}$，不合题意. $0 < \frac{1}{12}(7-\sqrt{13})$合乎题意. 故 θ 的最大似然估计值为 $\hat{\theta} = \frac{1}{12}(7-\sqrt{13})$.

［典型错误］　①不知道矩估计值为何物，基本式子也没有写对；②$E(X)$计算错；③不会写 $L(\theta)$，有的甚至同底幂相乘也乘错. 本题考的实际是几个认知点，有的考生连记也没有记住.

例 7.7　设总体 X 的概率密度为

$$f(x;\lambda) = \begin{cases} \lambda\alpha x^{\alpha-1}\mathrm{e}^{-\lambda x^{\alpha}}, & 若\ x > 0, \\ 0. & 若\ x \leqslant 0, \end{cases}$$

其中 $\lambda>0$ 是未知参数. $\alpha>0$ 是已知常数，试根据来自总体 X 的简单随机样本$X_1, X_2, \cdots, X_n$，求 λ 的最大似然估计量$\hat{\lambda}$.

［提示］　本题考查未知参数的最大似然估计法. 关键是正确写出似然函数.

［解］　似然函数

$$L(x_1, x_2, \cdots, x_n; \lambda) = (\lambda\alpha)^n \mathrm{e}^{-\lambda\sum_{i=1}^{n} x_i^{\alpha}} \prod_{i=1}^{n} x_i^{\alpha-1}.$$

对数似然函数
$$\ln L = n - n\ln\lambda + n\ln\alpha - \lambda\sum_{i=1}^{n} x_i^{\alpha} + \sum_{i=1}^{n}\ln x_i^{\alpha-1}.$$

由
$$\frac{\partial\ln L}{\partial\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i^{\alpha} = 0$$

解得 λ 的最大似然估计值为

$$\hat{\lambda} = n\Big/\sum_{i=1}^{n} x_i^{\alpha}.$$

从而 λ 的最大似然估计量为

$$\hat{\lambda} = n\Big/\sum_{i=1}^{n} x_i^{\alpha}.$$

例 7.8 假设0.50，1.25，0.80，2.00是来自总体 X 的简单随机样本值，已知 $Y=\ln X$ 服从正态分布 $N(\mu,1)$.

(1) 求 X 的数学期望 $E(X)$(记 $E(X)$ 为 b)；

(2) 求 μ 的置信度为0.95的置信区间；

(3) 利用上述结果求 b 的置信度为0.95的置信区间.

[提示] 已知 Y 的分布，$X=e^Y$. 因此 $b=E(X)=E(e^Y)$. 利用求随机变量函数的期望公式即可，由于 μ 是 Y 的均值，因此利用 $\overline{Y}$ 的抽样分布即可求出 μ 的置信区间. 根据所求出的 b ($E(X)$)与 μ 之间的关系即可求出 b 的置信区间。

[解] (1) Y 的概率密度为

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

于是,(令 $t=y-\mu$)

$$\begin{aligned}b&=EX=Ee^Y=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^ye^{-\frac{(y-\mu)^2}{2}}dy\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{t+\mu}e^{-\frac{1}{2}t^2}dt=e^{\mu+\frac{1}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(t-1)^2}dt\\&=e^{\mu+\frac{1}{2}}.\end{aligned}$$

(2) 当置信度 $1-\alpha=0.95$ 时，$\alpha=0.05$ 标准正态分布的水平为 $\alpha=0.05$ 的双侧分位数等于1.96，故由 $\overline{Y}\sim N(\mu,1/4)$. 可得参数 μ 的置信度为0.95的置信区间

$$\left[\overline{y}-1.96\times\frac{1}{\sqrt{4}},\overline{y}+1.96\times\frac{1}{\sqrt{4}}\right]=(\overline{y}-0.98,\overline{y}+0.98)\qquad(*)$$

其中 $\overline{y}$ 表示总体 Y 的样本均值. 于是，将

$$\overline{y}=\frac{1}{4}(\ln 0.5+\ln 0.8+\ln 1.25+\ln 2)=\frac{1}{4}\ln 1=0.$$

代入区间(*)的端点，得参数 μ 的0.95置信区间$(-0.98,0.98)$.

(3) 由 e^x 的严格递增性，可见 b 的置信度为0.95的置信区间为$(e^{-0.48},e^{-1.48})$.

例 7.9 设随机变量 X 的分布函数为

$$F(x;\alpha,\beta)=\begin{cases}1-\left(\frac{\alpha}{x}\right)^\beta, & x>\alpha,\\0, & x\leqslant\alpha,\end{cases}$$

其中参数 $\alpha>0$，$\beta>1$. 设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本.

(Ⅰ) 当 $\alpha=1$ 时，求未知参数 β 的矩估计量；

(Ⅱ) 当 $\alpha=1$ 时，求未知参数 β 的最大似然估计量；

(Ⅲ) 当 $\beta=2$ 时，求未知参数 α 的最大似然估计量.

[提示] 本题考查未知参数的矩估计法和最大似然估计法，是常考题型.

[解] 当 $\alpha=1$ 时，X 的概率密度为

$$f(x;\beta)=\begin{cases}\frac{\beta}{x^{\beta+1}}, & x>1,\\0, & x\leqslant 1.\end{cases}$$

(Ⅰ) 由于

$$E(X)=\int_{-\infty}^{+\infty}xf(x;\beta)dx=\int_1^{+\infty}x\frac{\beta}{x^{\beta+1}}dx=\frac{\beta}{\beta-1},$$

令 $\frac{\beta}{\beta-1}=\overline{X}$，解得 $\beta=\frac{\overline{X}}{\overline{X}-1}$，所以，参数 β 的矩估计量为

$$\hat{\beta} = \frac{\overline{X}}{\overline{X} - 1}.$$

（Ⅱ）对于总体 X 的样本值$x_1, x_2, \cdots, x_n$，似然函数为

$$L(\beta) = \prod_{i=1}^{n} f(x_i;\alpha) = \begin{cases} \dfrac{\beta^n}{(x_1 x_2 \cdots x_n)^{\beta+1}}, & x_i > 1\ (i = 1, 2, \cdots, n), \\ 0, & \text{其他}. \end{cases}$$

当 $x_i > 1\ (i = 1, 2, \cdots, n)$时，$L(\beta) > 0$，取对数得

$$\ln L(\beta) = n \ln \beta - (\beta + 1) \sum_{i=1}^{n} \ln x_i,$$

对 β 求导数，得

$$\frac{\mathrm{d} \ln L(\beta)}{\mathrm{d}\beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \ln x_i,$$

令$\dfrac{\mathrm{d}\ln(\beta)}{\mathrm{d}\beta} = 0$，解得

$$\beta = \frac{n}{\sum\limits_{i=1}^{n} \ln x_i},$$

β 的最大似然估计量为

$$\hat{\beta} = \frac{n}{\sum\limits_{i=1}^{n} \ln X_i}.$$

（Ⅲ）当 $\beta = 2$ 时，X 的概率密度为

$$f(x;\alpha) = \begin{cases} \dfrac{2\alpha^2}{x^3}, & x > \alpha, \\ 0, & x \leqslant \alpha. \end{cases}$$

对于总体 X 的样本值$x_1, x_2, \cdots, x_n$，似然函数为

$$L(\alpha) = \prod_{i=1}^{n} f(x_i;\alpha) = \begin{cases} \dfrac{2^n \alpha^{2n}}{(x_1 x_2 \cdots x_n)^3}, & x_i > \alpha\ (i = 1, 2, \cdots, n), \\ 0, & \text{其他}. \end{cases}$$

当 $x_i > \alpha\ (i = 1, 2, \cdots, n)$时，$\alpha$ 越大，$L(\alpha)$越大，因而 α 的最大似然估计值为

$$\hat{\alpha} = \min\{x_1, x_2, \cdots, x_n\},$$

则 α 的最大似然估计量为

$$\hat{\alpha} = \min\{X_1, X_2, \cdots, X_n\}.$$

［典型错误］ ① 在构造似然函数时，误把 X 的分布函数当成了概率密度，把似然函数定义为

$$L(\beta) = \prod_{i=1}^{n} F(x_i;\beta).$$

② 部分考生不会构造似然函数，说明考生对最大似然估计法不甚理解.

③ 在求解(Ⅲ)时，许多考生在写出似然函数后，不会继续求似然函数的极大值点，从而无法求出 α 的最大似然估计量，主要原因是在求解极大值点时不能使用常用的微分法，而只能用最值的定义求得，这正是(Ⅲ)与(Ⅱ)的区别所在.

④ 有些考生求得 α 的最大似然估计为

$$\hat{\alpha} = \max\{x_1, \cdots, x_n\},$$

显然这个答案是错误的，因为不能得证所有的 x_i 都大于$\hat{\alpha}$，所以 $L(\hat{\alpha}) = 0$，0 不可能为似然函数的最大值.

例 7.10 设$X_1, X_2, \cdots, X_n\ (n > 2)$为来自总体 $N(0, \sigma^2)$的简单随机样本，其样本均值为 $\overline{X}$. 记 $Y_i = X_i -$

$\overline{X}$，$i=1,2,\cdots,n$.

（Ⅰ）求 Y_i 的方差 $D(Y_i)$，$i=1,2,\cdots,n$；

（Ⅱ）求 Y_1 与 Y_n 的协方差 $\mathrm{Cov}(Y_1,Y_n)$；

（Ⅲ）若 $c(Y_1+Y_n)^2$ 是 σ^2 的无偏估计量，求常数 c.

［提示］ 本题考查方差、协方差的计算，以及无偏估计量的概念与性质.

［解］ （Ⅰ）$D(Y_i)=D(X_i-\overline{X})=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum_{k\neq i}X_k\right]$

$$=\frac{n-1}{n}\sigma^2, i=1,2,\cdots,n.$$

（Ⅱ）$\mathrm{Cov}(Y_1,Y_n)=E[Y_1-E(Y_1)][Y_n-E(Y_n)]=E(X_1-\overline{X})(X_n-\overline{X})$

$$=E(X_1X_n)+E(\overline{X}^2)-E(X_1\overline{X})-E(X_n\overline{X})$$

$$=E(X_1)E(X_n)+D(\overline{X})-\frac{1}{n}E(X_1^2)-\frac{1}{n}\sum_{i=2}^{n}E(X_1X_i)-\frac{1}{n}E(X_n^2)-\frac{1}{n}\sum_{i=1}^{n-1}E(X_iX_n)$$

$$=-\frac{1}{n}\sigma^2.$$

（Ⅲ）$E[c(Y_1+Y_n)^2]=cD(Y_1+Y_n)=c[D(Y_1)+D(Y_n)+2\mathrm{Cov}(Y_1,Y_n)]$

$$=c\left[\frac{n-1}{n}+\frac{n-1}{n}-\frac{2}{n}\right]\sigma^2=\frac{2(n-2)}{n}c\sigma^2$$

$$=\sigma^2,$$

故

$$c=\frac{n}{2(n-2)}.$$

例 7.11 设总体 X 的概率密度为

$$f(x)=\begin{cases}2e^{-2(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}$$

其中 $\theta>0$ 是未知参数，从总体 X 中抽取简单随机样本 $X_1,X_2,\cdots,X_n$，记 $\hat{\theta}=\min(X_1,X_2,\cdots,X_n)$.

(1) 求总体 X 的分布函数 $F(x)$；

(2) 求统计量 $\hat{\theta}$ 的分布函数 $F_{\hat{\theta}}(x)$；

(3) 如果用 $\hat{\theta}$ 作为 θ 的估计量，讨论它是否具有无偏性.

［提示］ 本题考查由分布密度求分布函数，求统计量 $\hat{\theta}$ 的分布函数，讨论估计量是否有无偏性.

［解］ (1) 分布函数

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\begin{cases}1-e^{-2(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}$$

(2) 统计量 $\hat{\theta}$ 的分布函数：

$$\begin{aligned}F_{\hat{\theta}}(x)&=P\{\hat{\theta}\leqslant x\}=P\{\min(X_1,X_2,\cdots,X_n)\leqslant x\}\\&=1-P\{\min(X_1,X_2,\cdots,X_n)>x\}\\&=1-P\{X_1>x,X_2>x,\cdots,X_n>x\}\\&=1-P\{X_1>x\}\cdot P\{X_2>x\}\cdots P\{X_n>x\}\\&=1-[1-F(x)]^n\\&=\begin{cases}1-e^{-2n(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}\end{aligned}$$

(3) $\hat{\theta}$ 的概率密度为

$$f_{\hat{\theta}}(x)=F'_{\hat{\theta}}(x)=\begin{cases}2ne^{-2n(x-\theta)}, & x>\theta,\\ 0, & x\leqslant\theta.\end{cases}$$

因为

$$E(\hat{\theta}) = \int_{-\infty}^{+\infty} x f_{\hat{\theta}}(x)\mathrm{d}x = \int_{0}^{+\infty} 2nx\mathrm{e}^{-2n(x-\theta)}\mathrm{d}x = \theta + \frac{1}{2n} \neq 0,$$

所以 $\hat{\theta}$ 作为 θ 的估计量不具有无偏性.

[典型错误] (1) 分布函数概念模糊, 例如将分布函数 $F(x)$ 写成

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{\theta}^{x} 2\mathrm{e}^{-2(t-\theta)}\mathrm{d}\theta,$$

或甚至写成

$$F(x) = \int_{0}^{x} 2\mathrm{e}^{-2(x-\theta)}\mathrm{d}\theta,$$

均未注明 $x > \theta$, 未写出当 $x \leqslant \theta$ 时的 $F(x)$.

有的干脆写成 $F(x) = \int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$.[应是 $F(+\infty) = 1$]

(2) 积分 $\int_{\theta}^{x} 2\mathrm{e}^{-2(x-\theta)}\mathrm{d}\theta$ 计算错.

(3) 将 $P\{X_i > x\}$ 写成 $F(x)$. [应是 $P\{X_i > x\} = 1 - P\{X_i \leqslant x\} = 1 - F(x)$].

(4) 数学期望 $E(X)$ 的公式写成 $\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1\left[\text{应是 } E(X) = \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x\right]$. 这些都说明基础不扎实.

例 7.12 设某种元件的使用寿命 X 的概率密度为

$$f(x;\theta) = \begin{cases} 2\mathrm{e}^{-2(x-\theta)}, & x \geqslant \theta, \\ 0, & x < \theta. \end{cases}$$

其中 $\theta > 0$ 为未知参数. 又设 $x_1, x_2, \cdots, x_n$ 是 X 的一组样本观测值. 求参数 θ 的最大似然估计值.

[提示] 根据最大似然估计的定义来求.

[解] 似然函数为

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

$$= \begin{cases} 2^n \mathrm{e}^{-2\sum_{i=1}^{n}(x_i - \theta)}, & x_i \geqslant \theta\ (i = 1, 2, \cdots, n). \\ 0, & \text{其他}. \end{cases}$$

当 $x_i \geqslant \theta\ (i = 1, 2, \cdots, n)$ 时, $L(\theta) > 0$, 取对数, 得

$$\ln L(\theta) = n\ln 2 - 2\sum_{i=1}^{n}(x_i - \theta).$$

因为 $\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta} = 2n > 0$, 所以 $L(\theta)$ 单调增加.

由于 θ 必须满足 $\theta \leqslant x_i\ (i = 1, 2, \cdots, n)$, 因此当 θ 取 $x_1, x_2, \cdots, x_n$ 中的最小值时, $L(\theta)$ 取最大值. 所以 θ 的最大似然估计值为 $\hat{\theta} = \min(x_1, x_2, \cdots, x_n)$.

[典型错误] 由于 $\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta} = 2n > 0$, $L(\theta)$ 单调增加, 有些考生就不知道应如何取估计值了.

八、假设检验

• 考试内容与要求 •

考试内容

显著性检验　假设检验的两类错误　单个及两个正态总体的均值和方差的假设检验

考试要求

1. 理解显著性检验的基本思想, 掌握假设检验的基本步骤, 了解假设检验可能产生的两类错误.