即 $$n \geqslant (1.96 \times 3)^2 \approx 34.57,$$
所以 n 至少应取 35.

［典型错误］ 不清楚样本均值的分布，不清楚正态分布的对称性.

例 3.6.12 设 $X_1, X_2, \cdots, X_9$ 是来自正态总体 X 的简单随机样本，

$$Y_1 = \frac{1}{6}(X_1 + \cdots + X_6), \quad Y_2 = \frac{1}{3}(X_7 + X_8 + X_9),$$

$$S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2, Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S},$$

证明统计量 Z 服从自由度为 2 的 t 分布.

［提示］ 本题主要考查 t 分布、χ^2 分布、随机变量和的分布. 要证明统计量 Z 服从自由度为 2 的 t 分布，必须证明 $Z = \dfrac{U}{\sqrt{\chi^2/2}}$，其中 $U \sim N(0,1)$，χ^2 服从自由度为 2 的 χ^2 分布.

［证］ 记 $DX = \sigma^2$（未知），易见

$$EY_1 = EY_2, \quad DY_1 = \frac{\sigma^2}{6}, \quad DY_2 = \frac{\sigma^2}{3}.$$

由于 Y_1 和 Y_2 独立，可见 $E(Y_1 - Y_2) = 0$,

$$D(Y_1 - Y_2) = \frac{\sigma^2}{6} + \frac{\sigma^2}{3} = \frac{\sigma^2}{2},$$

从而

$$U = \frac{Y_1 - Y_2}{\sigma/\sqrt{2}} \sim N(0,1).$$

由正态总体样本方差的性质，知

$$\chi^2 = \frac{2S^2}{\sigma^2}$$

服从自由度为 2 的 χ^2 分布.

由于 Y_1 与 Y_2，Y_1 与 S^2 以及 Y_2 与 S^2 独立，可见 $Y_1 - Y_2$ 与 S^2 独立.

于是，由服从 t 分布随机变量的结构，知

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

服从自由度为 2 的 t 分布.

［典型错误］

① 没有说明 $Y_1 - Y_2$ 与 S^2 独立.

② 不知 $\dfrac{2S^2}{\sigma^2} \sim \chi^2(2)$.

七、参数估计（数学三）

统计推断，即是用样本推断总体，是数理统计学的核心内容. 估计与检验是其中的两大类问题. 参数估计就是根据样本对总体中的未知参数或数字特征进行估计，常分为点估计与区间估计两种. 本节的重点是理解参数的点估计、估计量与估计值的概念，了解估计量的无偏性、有效性和相合性的概念及其证明方法. 掌握求估计量的矩估计法和最大似然估计法，了解区间估计的概念，掌握建立未知参数置信区间的一般方法，掌握正态总体参数及其相关特征的置信区间的求法.

• 考试内容与要求 •

考试内容

点估计的概念　估计量和估计值　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的区间估计　单个正态总体的方差和标准差的区间估计　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；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.

2. 掌握矩估计法(一阶、二阶矩)和最大似然估计法.

3. 掌握建立未知参数的(双侧和单侧)置信区间的一般方法；掌握正态总体均值、方差、标准差、矩以及与其相联系的数字特征的置信区间的求法.

4. 掌握两个正态总体的均值差和方差比及相关数字特征的置信区间的求法.

·考试内容解析·

(一) 点估计

1. 估计量、估计值

用来做估计的统计量称为估计量. 估计量是一个随机变量，它所取的具体值称为估计值. 点估计即是用估计量的值估计未知参数的值.

点估计问题，就是要构造一个统计量 $\hat{\theta}(X_1,\cdots,X_n)$，用它来估计未知参数 θ，即用 $\hat{\theta}(X_1,\cdots,X_n)$ 的观测值 $\hat{\theta}(x_1,\cdots,x_n)$ 作为未知参数 θ 的值.

2. 估计量的评选标准

对于一个未知参数，可供选择的估计量会有许多，而评选估计量好坏的标准有无偏性、有效性、相合性等.

(1) 无偏性：如 $E\hat{\theta}(X_1,\cdots,X_n)=\theta$，则称 $\hat{\theta}(X_1,\cdots,X_n)$ 为 θ 的无偏估计.

(2) 有效性：对于未知参数 θ，如果其两个估计量 $\hat{\theta}_1(X_1,\cdots,X_n)$ 和 $\hat{\theta}_2(X_1,\cdots,X_n)$ 有 $D\hat{\theta}_1\leqslant D\hat{\theta}_2$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.

(3) 相合性：对于未知参数 θ，如果其一个估计量 $\hat{\theta}(X_1,\cdots,X_n)\xrightarrow{P}\theta$，则称其为 θ 的相合估计量.

3. 矩估计法

矩估计法是用样本矩估计相应的总体矩从而得到参数估计的一种估计方法. 矩估计法不需要知道总体分布.

以 $\theta_1,\theta_2,\cdots,\theta_m$ 记总体 X 的 m 个待估的未知参数，以 α_k 和 μ_k 分别表示总体 X 的 k 阶原点矩和中心矩($k=1,2,\cdots$)且 α_m，μ_m 均存在，则求取 $\theta_1,\theta_2,\cdots,\theta_m$ 的矩估计的一般步骤为：

(1) 计算总体 X 的一阶直到 m 阶原点矩.

$$\alpha_i=\alpha_i(\theta_1,\cdots,\theta_m),\ i=1,\cdots,m.$$

(2) 用样本矩估计总体矩：

$$\alpha_i(\theta_1,\cdots,\theta_m)=\frac{1}{n}\sum_{l=1}^{n}X_l^i=A_i,i=1,2,\cdots,m.$$

(3) 求解上述方程，得到 θ_i 的矩估计为

$$\hat{\theta}_i=\theta_i(A_1,\cdots,A_m),\ i=1,2,\cdots,m.$$

考试大纲只要求涉及一阶矩和二阶矩，即 $m=1$ 或 2 的情形.

4. 最大似然估计法

最大似然估计法要求事先知道总体分布的数学形式. 设总体 X 的概率函数为 $f(x;\theta)$，其中 θ 是一维或二维未知参数. 对于离散型总体，其概率函数为

$$f(x;\theta)=\begin{cases}P\{X=x|\theta\}, & \text{如 } x \text{ 是 } X \text{ 的可能值},\\ 0, & \text{如 } x \text{ 不是 } X \text{ 的可能值};\end{cases}$$

对于连续型总体，其概率函数 $f(x;\theta)$ 即是概率密度.

(1) 似然函数：设总体 X 的概率函数为 $f(x;\theta)$，$X_1,\cdots,X_n$ 为来自总体 X 的一组简单随机样本．则称函数

$$L(\theta)=f(X_1;\theta)\cdot\cdots\cdot f(X_n;\theta)$$

为似然函数．称函数 $\ln L(\theta)$ 为对数似然函数．

(2) 最大似然估计量：对于给定的样本值 $x_1,\cdots,x_n$，使似然函数 $L(\theta)$ 或对数似然函数 $\ln L(\theta)$ 达到最大值的参数值 $\hat{\theta}$，称为未知参数 θ 的最大似然估计值．对于简单随机样本$(X_1,\cdots,X_n)$，称使 $L(\theta)$ 达到最大的 $\hat{\theta}=\hat{\theta}(X_1,\cdots,X_n)$ 为未知参数 θ 的最大似然估计量．

(3) 似然方程：为求 $L(\theta)$ 或 $\ln L(\theta)$ 的最大值，利用求极限方法，可以得到方程

$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=\sum_{i=1}^{n}\frac{1}{f(X_i;\theta)}\frac{\mathrm{d}f(X_i;\theta)}{\mathrm{d}\theta}=0,$$

这个方程就称为参数 θ 的似然方程．如果参数 $\theta=(\theta_1,\theta_2)$ 是二维的．则似然方程为

$$\begin{cases}\dfrac{\partial\ln L(\theta)}{\partial\theta_1}=\displaystyle\sum_{i=1}^{n}\frac{1}{f(X_i;\theta)}\frac{\partial f(X_i;\theta)}{\partial\theta_1}=0,\\[2ex]\dfrac{\partial\ln L(\theta)}{\partial\theta_2}=\displaystyle\sum_{i=1}^{n}\frac{1}{f(X_i;\theta)}\frac{\partial f(X_i;\theta)}{\partial\theta_2}=0.\end{cases}$$

在相当广泛的情况下，似然方程的解就是最大似然估计量．在有些情况下，似然函数关于 θ 的导数不存在，这时应采用其他方法，如定义求解最大似然估计量．

最大似然估计具有如下性质：设 $\hat{\theta}$ 是未知参数 θ 的最大似然估计量．对于 θ 的函数 $g(\theta)$．若 $g(\theta)$ 具有单值反函数，则 $g(\hat{\theta})$ 是参数 $g(\theta)$ 的最大似然估计量．例如，均值未知的正态总体 $N(\mu,\sigma^2)$ 的方差 σ^2 的最大似然估计为 $\hat{\sigma}^2=\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$，则总体标准差的最大似然估计为 $\hat{\sigma}=\sqrt{\sum\limits_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n}}$．

(二) 区间估计

未知参数 θ 的区间估计．又称置信区间．

1．置信区间

对于总体 X．其未知参数为 θ．设 $X_1,\cdots,X_n$ 为来自总体 X 的简单随机样本，对于给定的 α $(0<\alpha<1)$，如果两个统计量 $\hat{\theta}_1=\theta_1(X_1,\cdots,X_n)$ 和 $\hat{\theta}_2=\theta_2(X_1,\cdots,X_n)$ 满足

$$P\{\hat{\theta}_1<\theta<\hat{\theta}_2\}=1-\alpha,$$

则称随机区间$(\hat{\theta}_1,\hat{\theta}_2)$为未知参数 θ 的置信度为 $1-\alpha$ 的置信区间．$\hat{\theta}_1$ 和 $\hat{\theta}_2$ 分别称为置信下限与置信上限．$1-\alpha$ 称为置信度或置信水平．

置信度是随机区间$(\hat{\theta}_1,\hat{\theta}_2)$"包含"或"覆盖"未知参数 θ 的值的概率；置信度一般选取非常接近 1 的数，如 0.99，0.95 和 0.90 等．直观上，如果多次使用置信度为 0.95 的置信区间$(\hat{\theta}_1,\hat{\theta}_2)$，则平均有 95% 的区间包含 θ 的值．

2．置信区间的求取方法

设 θ 是总体 X 的未知参数，$\boldsymbol{X}=(X_1,\cdots,X_n)$ 是来自总体 X 的简单随机样本，关于 θ 的置信度为 $1-\alpha$ 的置信区间的构造步骤如下：

(1) 选择一个样本及未知参数 θ 的函数 $T=f(\boldsymbol{X};\theta)$，但是其分布与 θ 无关，且未知参数 θ 可以用样本和 T 表出，即 $\theta=g(\boldsymbol{X};T)$ 是 $T=f(\boldsymbol{X};\theta)$ 的反函数．

(2) 对于给定的置信度 $1-\alpha$，根据 T 的分布选取两个常数 a 与 b，使得

$$P\{a<T<b\}=1-\alpha.$$

(3) 利用 $\theta=g(\boldsymbol{X};T)$ 与 $T=f(\boldsymbol{X};\theta)$ 间的反函数关系．有

$$1-\alpha=P\{a<T<b\}=P\{\hat{\theta}_1<\theta<\hat{\theta}_2\},$$

则$(\hat{\theta}_1,\hat{\theta}_2)$即为 θ 的置信度为 $1-\alpha$ 的置信区间．这种方法常称为枢轴量法．

上述选取置信区间的方法可能会导致置信区间不唯一，于是对于对称分布(如正态分布，t 公布)及一些常用的非对称分布(如 χ^2 分布和 F 分布)，我们通常按如下原则选取 a 与 b：

$$P\{T\leqslant a\}=P\{T\geqslant b\}=\frac{\alpha}{2}.$$

3. 单个正态总体参数的置信区间

设总体 $X\sim N(\mu,\sigma^2)$，$X_1,\cdots,X_n$ 为来自总体 X 的简单随机样本，$\bar{X}$ 是样本均值，S^2 是样本方差.

(1) 当 σ^2 已知时，μ 的置信度为 $1-\alpha$ 的置信区间为

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

(2) 当 σ^2 未知时，μ 的置信度为 $1-\alpha$ 的置信区间为

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

(3) 当 μ 已知时，σ^2 的置信度为 $1-\alpha$ 的置信区间为

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

(4) 当 μ 未知时，σ^2 的置信度为 $1-\alpha$ 的置信区间为

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

4. 两个正态总体参数的置信区间

设总体 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，$(X_1,\cdots,X_m)$和$(Y_1,\cdots,Y_n)$分别是来自总体 X 和 Y 的简单随机样本，$\bar{X},S_x^2,\bar{Y},S_y^2$ 分别是相应的样本均值和方差，并且两组样本是相互独立的.

(1) 当 σ_1^2 和 σ_2^2 已知时，$\mu_1-\mu_2$ 的置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\bar{Y}-u_{\frac{\alpha}{2}}\cdot\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}},\ \bar{X}-\bar{Y}+u_{\frac{\alpha}{2}}\cdot\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\right).$$

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

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

其中

$$S_w^2=\frac{(m-1)S_x^2+(n-1)S_y^2}{m+n-2}.$$

(3) 当 μ_1，μ_2 已知时，σ_1^2/σ_2^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left[F^{-1}_{\frac{\alpha}{2}}(m,n)\cdot\frac{n\sum\limits_{i=1}^{m}(X_i-\mu_1)^2}{m\sum\limits_{i=1}^{n}(Y_i-\mu_2)^2},F_{\frac{\alpha}{2}}(n,m)\cdot\frac{n\sum\limits_{i=1}^{m}(X_i-\mu_1)^2}{m\sum\limits_{i=1}^{n}(Y_i-\mu_2)^2}\right].$$

(4) 当 μ_1 和 μ_2 未知时，σ_1^2/σ_2^2 的置信度为 $1-\alpha$ 的置信区间为

$$\left(F^{-1}_{\frac{\alpha}{2}}(m-1,n-1)\frac{S_x^2}{S_y^2},\ F_{\frac{\alpha}{2}}(n-1,m-1)\frac{S_x^2}{S_y^2}\right).$$

·例题详解·

例 3.7.1 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] 本题主要考查正态分布的分位数、置信区间.

[解] 记 $\bar{X}$ 为样本均值，则置信度为 $\alpha=0.95$ 的置信区间为

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

由于 $u_{\frac{\alpha}{2}}=1.96$，故所求置信区间为

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

[典型错误] 用错正态分布的分位数.

例 3.7.2 设总体 X 的概率密度为

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

而 $X_1,X_2,\cdots,X_n$ 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] $\frac{1}{n}\sum\limits_{i=1}^{n}X_i-1$ 或 $\bar{X}-1$.

[提示] 本题主要考查矩估计.

[解] 由于 $EX=\int_{-\infty}^{+\infty}xf(x;\theta)\,dx=\int_{\theta}^{+\infty}xe^{-(x-\theta)}dx=\theta+1$，故由矩估计方法，满足 $EX=\bar{X}$ 的 θ 即为 θ 的矩估计量，因此 $\hat{\theta}=\bar{X}-1$.

例 3.7.3 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 $\bar{x}=20$ (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] 本题主要考查方差未知时正态总体均值的置信区间．对于 n 个来自 $N(\mu,\sigma^2)$ 的简单随机样本 $X_1,\cdots,X_n$，当 σ^2 未知时，μ 的置信度为 $1-\alpha$ 的置信区间为

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

[解] 由于 μ 与 σ^2 均未知，故知

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

于是 μ 的置信度为 $1-\alpha$ 的置信区间为

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

把 $\bar{x}=20$，$S=1$，$\alpha=1-0.9=0.1$ 代入上式，得

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

[典型错误] 误以为 σ^2 已知而选(A).

例 3.7.4 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] 本题主要考查矩估计、最大似然估计．本题前两问均可以按照常规的方法求解，而第(Ⅲ)问必须用最大似然估计的定义来求解，这是由于其支撑集与未知参数有关.

[解] 当 $\alpha=1$ 时，X 的概率密度为

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

（Ⅰ）由于

$$EX=\int_{-\infty}^{+\infty}xf(x;\beta)\mathrm{d}x=\int_{1}^{+\infty}x\frac{\beta}{x^{\beta+1}}\mathrm{d}x=\frac{\beta}{\beta-1}.$$

令$\dfrac{\beta}{\beta-1}=\overline{X}$，解得 $\beta=\dfrac{\overline{X}}{\overline{X}-1}$，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

$$L(\beta)=\prod_{i=1}^{n}f(x_i;\beta)=\begin{cases}\dfrac{\beta^n}{(x_1x_2\cdots x_n)^{\beta+1}}, & x_i>1\ (i=1,2,\cdots,n),\\ 0, & \text{其他}.\end{cases}$$

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

$$L(\alpha)=\prod_{i=1}^{n}f(x_i;\alpha)=\begin{cases}\dfrac{2^n\alpha^{2n}}{(x_1x_2\cdots x_n)^3}, & x_i>\alpha\ (i=1,2,\cdots,n),\\ 0, & \text{其他}.\end{cases}$$

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 3.7.5 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 θ $\left(0<\theta<\frac{1}{2}\right)$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] 本题主要考查矩估计和最大似然估计.

[解]
$$EX=0\times\theta^2+1\times2\theta(1-\theta)+2\times\theta^2+3\times(1-2\theta)=3-4\theta,$$
$$\bar{x}=\frac{1}{8}(3+1+3+0+3+1+2+3)=2.$$

令 $EX=\bar{x}$，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 的极大似然估计值为$\hat{\theta}=\frac{1}{12}(7-\sqrt{13})$.

[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 EX 计算错.

③ 不会写 $L(\theta)$，有的甚至同底幂相乘也乘错.

例 3.7.6 设总体 X 的概率密度为

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

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

(Ⅰ) 求总体 X 的分布函数 $F(x)$；

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

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

[提示] 本题主要考查由密度求分布函数、极小值统计量的分布和无偏估计.

[解] (Ⅰ) 分布函数

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

(Ⅱ) 统计量 $\hat{\theta}$ 的分布函数

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

（Ⅲ）$\hat\theta$ 的概率密度为

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

因为

$$
E\hat\theta = \int_{-\infty}^{+\infty} x f_{\hat\theta}(x) \mathrm{d}x = \int_0^{+\infty} 2nx e^{-2n(x-\theta)} \mathrm{d}x = \theta + \frac{1}{2n} \neq 0,
$$

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 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_\theta^x 2e^{-2(x-\theta)}\mathrm{d}\theta$，或甚至写成 $F(x) = \int_0^x 2e^{-2(x-\theta)}\mathrm{d}\theta$. 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 $\int_\theta^x 2e^{-2(x-\theta)}\mathrm{d}\theta$ 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 EX 的公式写成 $\int_{-\infty}^{+\infty} f(t)\mathrm{d}t = 1$，应是 $EX = \int_{-\infty}^{+\infty} x f(x)\mathrm{d}x$.

这些都说明基础不扎实.

例 3.7.7 设总体 X 的概率密度为

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

其中 $\theta > -1$ 是未知参数，$X_1, X_2, \cdots, X_n$ 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] 本题主要考查矩估计和最大似然估计.

[解] 总体 X 的数学期望为

$$
EX = \int_{-\infty}^{+\infty} x f(x)\mathrm{d}x = \int_0^1 (\theta+1)x^{\theta+1}\mathrm{d}x = \frac{\theta+1}{\theta+2} .
$$

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

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

解得未知参数 θ 的矩估计量为

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

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

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

当$0<x_i<1\ (i=1,2,\cdots,n)$时，$L>0$，且

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

令

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

解得θ的极大似然估计值为

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

从而得θ的极大似然估计量为

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

[典型错误]

① 不会求EX.

② 矩估计和极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 3.7.8 设总体X的概率密度为

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

$X_1,X_2,\cdots,X_n$是取自总体X的简单随机样本.

（Ⅰ）求θ的矩估计量$\hat{\theta}$；

（Ⅱ）求$\hat{\theta}$的方差$D\hat{\theta}$.

[提示] 本题主要考查矩估计及其期望与方差. 求θ的矩估计量$\hat{\theta}$，只要令$EX=\overline{X}$即可. 求$\hat{\theta}$的方差需计算DX. 本题实质上是计算已给密度函数的随机变量的均值和方差问题.

[解] （Ⅰ）$EX=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^{\theta}\frac{6x^2}{\theta^3}(\theta-x)\mathrm{d}x=\frac{\theta}{2}$.

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

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

（Ⅱ）由于

$$EX^2=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_0^{\theta}\frac{6x^3}{\theta^3}(\theta-x)\mathrm{d}x=\frac{6\theta^2}{20}.$$

$$DX=EX^2-(EX)^2=\frac{6\theta^2}{20}-\left(\frac{\theta}{2}\right)^2=\frac{\theta^2}{20}.$$

所以$\hat{\theta}=2\overline{X}$的方差为

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

[典型错误] 不会求EX，EX^2.

例 3.7.9 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] 本题主要考查最大似然估计. 由于此分布的支撑集与未知参数θ有关，故要按照定义来取θ的最大似然估计.

［解］ 似然函数为

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

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

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］ 不知用定义求取某些参数的最大似然估计．

例 3.7.10 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 $N(\mu,1)$．

（Ⅰ）求 X 的数学期望 EX（记 EX 为 b）；

（Ⅱ）求 μ 的置信度为 0.95 的置信区间；

（Ⅲ）利用上述结果求 b 的置信度为 0.95 的置信区间．

［提示］ 本题主要考查随机变量函数的分布、数学期望、置信区间．已知 Y 的分布，$X=e^Y$．因此 $b=EX=E(e^Y)$，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 $\bar{Y}$ 的抽样分布即可求出 μ 的置信区间，根据所求出的 $b\,(EX)$与 μ 之间的关系即可求出 b 的置信区间．

［解］（Ⅰ）Y 的概率密度为

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

于是，有

$$\begin{aligned}b&=EX=E(e^Y)\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^y e^{-\frac{(y-\mu)^2}{2}}\mathrm{d}y\quad（令\ t=y-\mu）\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{t+\mu}e^{-\frac{1}{2}t^2}\mathrm{d}t\\&=e^{\mu+\frac{1}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(t-1)^2}\mathrm{d}t\\&=e^{\mu+\frac{1}{2}}.\end{aligned}$$

（Ⅱ）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 $\bar{Y}\sim N\left(\mu,\dfrac{1}{4}\right)$，可得参数 μ 的置信度为 0.95 的置信区间为

$$\left(\bar{Y}-1.96\times\frac{1}{\sqrt{4}},\ \bar{Y}+1.96\times\frac{1}{\sqrt{4}}\right)=(\bar{Y}-0.98,\bar{Y}+0.98).\tag{*}$$

其中 $\bar{Y}$ 表示总体 Y 的样本均值．于是，将

$$\bar{Y}=\frac{1}{4}(\ln 0.5+\ln 0.8+\ln 1.25+\ln 2)=\frac{1}{4}\ln 1=0$$

代入区间（*）的端点，得参数 μ 的置信度为 0.95 的置信区间为$(-0.98,0.98)$．

（Ⅲ）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$(e^{-0.48},e^{1.48})$．

［典型错误］ 不知如何求随机变量函数的数学期望、不会利用 μ 的置信区间求 b 的置信区间．

例 3.7.11 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 $\bar{X}$．记 $Y_i=$

$X_i-\bar{X}$，$i=1,2,\cdots,n$.

(Ⅰ) 求 Y_i 的方差 DY_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] 本题主要考查随机变量的数字特征、无偏估计．在计算 DY_i 及 $\mathrm{cov}(Y_1,Y_n)$ 时，用 $X_1,\cdots,X_n$ 的线性组合表示 Y_i 之后再利用 X_i 间的独立性，易于计算.

[解] (Ⅰ)
$$DY_i=D(X_i-\bar{X})=D\left[\left(1-\frac{1}{n}\right)X_i-\frac{1}{n}\sum_{k\neq i}X_k\right]=\frac{n-1}{n}\sigma^2,i=1,2,\cdots,n.$$

(Ⅱ)
$$\begin{aligned}\mathrm{cov}(Y_1,Y_n)&=E(Y_1-EY_1)(Y_n-EY_n)\\&=E(X_1-\bar{X})(X_n-\bar{X})\\&=E(X_1X_n)+E\bar{X}^2-E(X_1\bar{X})-E(X_n\bar{X})\\&=EX_1EX_n+D\bar{X}-\frac{1}{n}EX_1^2-\frac{1}{n}\sum_{i=2}^{n}E(X_1X_i)\\&\quad-\frac{1}{n}EX_n^2-\frac{1}{n}\sum_{i=1}^{n-1}E(X_iX_n)\\&=-\frac{1}{n}\sigma^2.\end{aligned}$$

(Ⅲ)
$$\begin{aligned}E[c(Y_1+Y_n)^2]&=cD(Y_1+Y_n)\\&=c[DY_1+DY_n+2\mathrm{cov}(Y_1,Y_n)]\\&=c\left(\frac{n-1}{n}+\frac{n-1}{n}-\frac{2}{n}\right)\sigma^2\\&=\frac{2(n-2)}{n}c\sigma^2\\&=\sigma^2,\end{aligned}$$

故
$$c=\frac{n}{2(n-2)}.$$

[典型错误] 误认为 $DY_i=DX_i+D\bar{X}$，$D(Y_1+Y_n)=DY_1+DY_n$.

八、假设检验(数学三)

估计和检验是数理统计中的两大类问题．估计是依据样本推断总体未知参数或分布，检验是依据样本来推断关于总体参数或分布的某种假设是否成立．本节的重点是理解“假设”的概念、类型及显著性检验的思想，掌握显著性检验的基本步骤，会构造简单假设的显著性检验．另外也要理解检验的两类错误，并会计算某些简单情形的两类错误的概率．对于单个及两个正态总体的均值和方差的假设会进行显著性检验.

·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“假设”的概念和基本类型；理解显著性检验的基本思想．掌握假设检验的基本步骤；会构造简单假设的显著性检验.

2．理解假设检验可能产生的两类错误，对于较简单的情形，会计算两类错误的概率.

3．掌握单个及两个正态总体的均值和方差的假设检验.