似服从参数为 $\mu = a_2$，$\sigma^2 = \frac{a_4 - a_2^2}{n}$ 的正态分布．

**例5.8** 一生产线生产的产品成箱包装，每箱的重量是随机的．假设每箱平均重 50 千克，标准差为 5 千克．若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多可以装多少箱，才能保障不超载的概率大于 0.977．（$\Phi(2) = 0.977$，其中 $\Phi(x)$ 是标准正态分布函数．）

［提示］ 中心极限定理是针对 $n$ 个独立同分布的随机变量和而言的，根据题意构造这样的一个独立随机变量和，最后求出 $n$．

［解］ 设 $X_i\ (i = 1,2,\cdots,n)$ 是装运的第 $i$ 箱的重量（单位：千克），$n$ 是所求箱数．由条件可以把 $X_1$，$X_2,\cdots,X_n$ 视为独立同分布随机变量，则 $n$ 箱的总重量

$$T_n = X_1 + X_2 + \cdots + X_n$$

是独立同分布随机变量之和．

由条件知 $E(X_i) = 50$，$\sqrt{D(X_i)} = 5$；$E(T_n) = 50n$，$\sqrt{D(T_n)} = 5\sqrt{n}$（单位：千克）．

根据列维－林德伯格中心极限定理，$T_n$ 近似服从正态分布 $N(50n, 25n)$．

箱数 $n$ 决定于条件

$$P\{T_n \leqslant 5\ 000\} = P\left\{\frac{T_n - 50n}{5\sqrt{n}} \leqslant \frac{5\ 000 - 50n}{5\sqrt{n}}\right\}$$

$$\approx \Phi\left(\frac{1\ 000 - 10n}{\sqrt{n}}\right) > 0.977 = \Phi(2).$$

由此可见

$$\frac{1\ 000 - 10n}{\sqrt{n}} > 2,$$

从而 $n < 98.019\ 9$，即最多可以装 98 箱．

## 六、数理统计的基本概念

### ·考试内容与要求·

**考试内容**

总体　个体　简单随机样本　统计量　样本均值　样本方差和样本矩　$\chi^2$ 分布　$t$ 分布　$F$ 分布　分位数　正态总体的常用抽样分布

**考试要求**

1．理解总体、简单随机样本、统计量、样本均值、样本方差及样本矩的概念．其中样本方差定义为：

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

2．了解 $\chi^2$ 分布、$t$ 分布和 $F$ 分布的概念及性质，了解分位数的概念并会查表计算．

3．了解正态总体的常用抽样分布．

### ·考试内容解析·

**（一）总体和样本**

1．总体

在数理统计中所研究对象的某项数量指标 $X$ 取值的全体称为总体，$X$ 是一个随机变量．$X$ 的分布函数和数字特征分别称为总体的分布函数和数字特征．总体中的每个元素称为个体．每个个体是一个实数．总体中个体的数量称为总体的容量．容量为有限的总体称为有限总体，容量为无限的总体称为无限总体．

2．简单随机样本

设总体 $X$ 的分布函数为 $F(x)$，如果随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，并且都具有分布函数 $F(x)$，则称 $X_1,X_2,\cdots,X_n$ 为来自总体 $X$ 的简单随机样本，简称为样本，$n$ 称为样本容量．它们的观测值 $x_1,x_2,\cdots,x_n$

称为样本观测值，简称为样本值，也称为总体 $X$ 的 $n$ 个独立观测值.

如果总体 $X$ 的分布函数为 $F(x)$，$X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的样本，则随机变量 $X_1,X_2,\cdots,X_n$ 的联合分布函数为

$$F(t_1,t_2,\cdots,t_n)=\prod_{i=1}^{n}F(t_i),\quad t_i\in\mathbf{R}\ (i=1,2,\cdots,n).$$

如果总体 $X$ 的概率密度为 $f(x)$，则 $X_1,X_2,\cdots,X_n$ 的联合概率密度为

$$f(t_1,t_2,\cdots,t_n)=\prod_{i=1}^{n}f(t_i),\quad t_i\in\mathbf{R}\ (i=1,2,\cdots,n).$$

3. 统计量及抽样分布

设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的样本，$g(t_1,t_2,\cdots,t_n)$ 是一个不含未知数的 $n$ 元函数，则称随机变量 $X_1,X_2,\cdots,X_n$ 的函数 $T=g(X_1,X_2,\cdots,X_n)$ 为一个统计量. 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，则称数值 $g(x_1,x_2,\cdots,x_n)$ 为统计量 $g(X_1,X_2,\cdots,X_n)$ 的观测值.

统计量是样本的函数，是一个随机变量. 统计量的分布称为抽样分布.

（二）样本矩

设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的样本，$x_1,x_2,\cdots,x_n$ 是相应的样本值.

1. 样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，观测值为 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$.

2. 样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 观测值为

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right).$$

样本标准差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$，观测值为 $s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$.

如果总体 $X$ 具有数学期望 $E(X)=\mu$ 和方差 $D(X)=\sigma^2$，则

$$E(\bar{X})=E(X)=\mu,\quad D(\bar{X})=\frac{D(X)}{n}=\frac{\sigma^2}{n},\quad E(S^2)=D(X)=\sigma^2.$$

3. 样本 $k$ 阶原点矩 $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$，观测值为 $a_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k$，$k=1,2,\cdots$.

4. 样本 $k$ 阶中心矩 $B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})k$，观测值为 $b_k=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^k$，$k=1,2,\cdots$.

如果总体 $X$ 的 $k$ 阶原点矩 $E(X^k)=\mu_k\ (k=1,2,\cdots)$ 存在，则当 $n\to\infty$ 时，有

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k\xrightarrow{P}\mu_k,\quad k=1,2,\cdots,$$

如果 $g$ 是连续函数，则有

$$g(A_1,A_2,\cdots,A_k)\xrightarrow{P}g(\mu_1,\mu_2,\cdots,\mu_k).$$

（三）$\chi^2$ 分布

1. 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，都服从标准正态分布 $N(0,1)$，则随机变量

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

的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

称随机变量 $\chi^2$ 服从自由度为 $n$ 的 $\chi^2$ 分布，记作 $\chi^2\sim\chi^2(n)$.

2. $\chi^2$ 分布的性质

(1) 设 $\chi_1^2 \sim \chi^2(n_1)$，$\chi_2^2 \sim \chi^2(n_2)$，并且 $\chi_1^2$ 和 $\chi_2^2$ 相互独立，则

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

(2) 如果 $\chi^2 \sim \chi^2(n)$，则有

$$E(\chi^2) = n,\quad D(\chi^2) = 2n.$$

3. 上 $\alpha$ 分位点 $\chi_\alpha^2(n)$

设 $\chi^2 \sim \chi^2(n)$，对于给定的 $\alpha\ (0<\alpha<1)$，称满足条件

$$P\{\chi^2 > \chi_\alpha^2(n)\} = \alpha$$

的点 $\chi_\alpha^2(n)$ 为 $\chi^2(n)$ 分布的上 $\alpha$ 分位点.

**(四) $t$ 分布**

1. 设随机变量 $X$ 和 $Y$ 相互独立，且 $X \sim N(0,1)$，$Y \sim \chi^2(n)$，则随机变量

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

的概率密度为

$$f(x) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\dfrac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},\quad x\in \mathbf{R}.$$

称随机变量 $t$ 服从自由度为 $n$ 的 $t$ 分布，记作 $t \sim t(n)$.

2. $t(n)$ 分布的概率密度 $f(x)$ 是偶函数，且有

$$\lim_{n\to\infty} f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},$$

即当 $n$ 充分大时，$t(n)$ 分布近似于 $N(0,1)$ 分布.

3. 上 $\alpha$ 分位点 $t_\alpha(n)$

设 $t \sim t(n)$，对于给定的 $\alpha\ (0<\alpha<1)$，称满足条件

$$P\{t > t_\alpha(n)\} = \alpha$$

的点 $t_\alpha(n)$ 为 $t(n)$ 分布的上 $\alpha$ 分位点.

由于 $t(n)$ 分布的概率密度是偶函数，因此

$$t_{1-\alpha}(n) = -t_\alpha(n).$$

**(五) $F$ 分布**

1. 设随机变量 $X$ 和 $Y$ 相互独立，且 $X \sim \chi^2(n_1)$，$Y \sim \chi^2(n_2)$，则随机变量

$$F = \frac{\dfrac{X}{n_1}}{\dfrac{Y}{n_2}}$$

的概率密度为

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)} n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}} \cdot \dfrac{x^{\frac{n_1}{2}-1}}{(n_1 x + n_2)^{\frac{n_1+n_2}{2}}}, & x>0; \\ 0, & x\leqslant 0. \end{cases}$$

称随机变量 $F$ 服从自由度为 $(n_1,n_2)$ 的 $F$ 分布，记作 $F \sim F(n_1,n_2)$，其中 $n_1$ 和 $n_2$ 分别称为第一自由度和第二自由度.

2. 如果 $F \sim F(n_1,n_2)$，则 $\dfrac{1}{F} \sim F(n_2,n_1)$.

3. 上 $\alpha$ 分位点 $F_\alpha(n_1,n_2)$

设 $F \sim F(n_1,n_2)$，对于给定的 $\alpha$ $(0<\alpha<1)$，称满足条件

$$P\{F > F_\alpha(n_1,n_2)\} = \alpha$$

的点 $F_\alpha(n_1,n_2)$为 $F(n_1,n_2)$分布的上 $\alpha$ 分位点，且有

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

**（六）正态总体的几个常用统计量的分布**

1．设总体 $X$ 服从正态分布 $N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的样本，样本均值为 $\overline{X}$，样本方差为 $S^2$，则有

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

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

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

(4) $\chi^2 = \dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$.

2．设总体 $X \sim N(\mu_1,\sigma_1^2)$，$X_1,X_2,\cdots,X_{n_1}$是来自总体 $X$ 的样本，样本均值为 $\overline{X}$，样本方差为 $S_1^2$；总体 $Y \sim N(\mu_2,\sigma_2^2)$，$Y_1,Y_2,\cdots,Y_{n_2}$是来自总体 $Y$ 的样本，样本均值为 $\overline{Y}$，样本方差为 $S_2^2$，且来自两个总体的样本相互独立，则有

(1) $u = \dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$.

(2) 如果 $\sigma_1^2=\sigma_2^2=\sigma^2$，则有

$$t = \frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2),$$

其中

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

(3) $F = \dfrac{n_2\sigma_2^2}{n_1\sigma_1^2}\cdot\dfrac{\sum\limits_{i=1}^{n_1}(X_i-\mu_1)^2}{\sum\limits_{j=1}^{n_2}(Y_j-\mu_2)^2} \sim F(n_1,n_2)$.

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

**（七）$X_{(n)}=\max(X_1,X_2,\cdots,X_n)$和 $X_{(1)}=\min(X_1,X_2,\cdots,X_n)$的分布**

设总体 $X$ 的分布函数为 $F(x)$，$X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的样本，则统计量 $X_{(n)}=\max(X_1,X_2,\cdots,X_n)$和$X_{(1)}=\min(X_1,X_2,\cdots,X_n)$的分布函数分别为

$$F_{\max}(x) = P\{\max(X_1,X_2,\cdots,X_n)\leqslant x\} = [F(x)]^n,$$
$$F_{\min}(x) = P\{\min(X_1,X_2,\cdots,X_n)\leqslant x\} = 1-[1-F(x)]^n.$$

## · 例 题 详 解 ·

**例 6.1** 设随机变量 $X$ 和 $Y$ 相互独立且都服从正态分布 $N(0,3^2)$，而$X_1,\cdots,X_9$ 和 $Y_1,\cdots,Y_9$ 分别是来自总体 $X$ 和 $Y$ 的简单随机样本，则统计量 $U=\dfrac{X_1+\cdots+X_9}{\sqrt{Y_1^2+\cdots+Y_9^2}}$服从________分布，参数为________.

［答案］ $t$，9.

［提示］ 本题考查 $t$ 分布的相关概念与参数计算.

［解］ 令 $X'_i=\frac{X_i}{3}$，$Y'_i=\frac{Y_i}{3}$，　$i=1,2,\cdots,9$.

则
$$X'_i\sim N(0,1),\ Y'_i\sim N(0,1),\quad i=1,2,\cdots,9.$$
$$X'=X'_1+\cdots+X'_9\sim N(0,3^2).$$
$$Y'=Y_1'^2+\cdots+Y_9'^2\sim\chi^2(9).$$

因此
$$U=\frac{X_1+\cdots+X_9}{\sqrt{Y_1^2+\cdots+Y_9^2}}=\frac{X'_1+\cdots+X'_9}{\sqrt{Y_1'^2+\cdots+Y_9'^2}}=\frac{X'}{\sqrt{Y'}}=\frac{X'/3}{\sqrt{\frac{Y'}{9}}},$$

由于
$$X'/3\sim N(0,1),\ Y'\sim\chi^2(9).$$

故
$$U\sim t(9).$$

例6.2　设 $X_1$，$X_2$，$X_3$，$X_4$ 是来自正态总体 $N(0,2^2)$的简单随机样本，

$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$．则当 $a=$________，$b=$________时，统计量 $X$ 服从$\chi^2$ 分布，其自由度为________.

［答案］ $\frac{1}{20}$，$\frac{1}{100}$，2.

［提示］ 本题考查 $\chi^2$ 分布的概念与性质.

［解］ 根据 $\chi^2$ 分布的定义，若 $Y_1,Y_2,\cdots,Y_m$ 服从标准正态分布，则 $Y=Y_1^2+Y_2^2+\cdots+Y_m^2$ 服从自由度为 $m$ 的 $\chi^2$ 分布．对于本题，若 $X$ 服从 $\chi^2$ 分布，则 $m=2$，且须$\sqrt{a}(X_1-2X_2)\sim N(0,1)$；$\sqrt{b}(3X_3-4X_4)\sim N(0,1)$.

于是
$$D[\sqrt{a}(X_1-2X_2)]=(a+4a)D(X_1)=5a\times 2^2=1，即 a=\frac{1}{20}.$$
$$D[\sqrt{b}(3X_3-4X_4)]=(9b+16b)D(X_1)=25b\times 2^2=1，即 b=\frac{1}{100}.$$

例6.3　设总体 $X$ 服从正态分布$N(0,2^2)$．而$X_1,X_2,\cdots,X_{15}$是来自总体 $X$ 的简单随机样本，则随机变量
$$Y=\frac{X_1^2+\cdots+X_{10}^2}{2(X_{11}^2+\cdots+X_{15}^2)}$$
服从________分布，参数为________.

［答案］ $F$，(10,5).

［提示］ 本题考查 $F$ 分布的概念与性质.

［解］ 由于 $X_i\sim N(0,2^2)$．则$\frac{1}{2}X_i\sim N(0,1)$.

故
$$U=\frac{1}{4}X_1^2+\cdots+\frac{1}{4}X_{10}^2\sim\chi^2(10).$$
$$V=\frac{1}{4}X_{11}^2+\cdots+\frac{1}{4}X_{15}^2\sim\chi^2(5)$$
$$Y=\frac{U/10}{V/5}\sim F(10,5).$$

例6.4　设总体 $X$ 服从正态分布 $N(\mu_1,\sigma^2)$，总体 $Y$ 服从正态分布 $N(\mu_2,\sigma^2)$，$X_1,X_2,\cdots,X_{n_1}$ 和 $Y_1,Y_2,\cdots,Y_{n_2}$ 分别是来自总体 $X$ 和 $Y$ 的简单随机样本，则
$$E\left[\frac{\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{j=1}^{n_2}(Y_j-\overline{Y})^2}{n_1+n_2-2}\right]=\_\_\_\_\_\_\_\_.$$

［答案］ $\sigma^2$.

［提示］ 本题考查的是样本方差为 $\sigma^2$ 的无偏估计的性质以及数学期望的线性性质.

［解］ 记

$$S_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2,\ S_2^2=\frac{1}{n_2-1}\sum_{j=1}^{n_2}(Y_j-\overline{Y})^2,$$

则
$$E(S_1^2)=E(S_2^2)=\sigma^2,$$

故
$$\text{原式}=E\left[\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\right]$$
$$=\frac{(n_1-1)E(S_1^2)+(n_2-1)E(S_2^2)}{n_1+n_2-2}$$
$$=\sigma^2.$$

[典型错误] 许多考生不能把本题和样本方差的无偏性联系起来，而直接求数学期望，从而造成计算错误或无从下手.

例6.5 设随机变量 $X\sim t(n)(n>1)$，$Y=\frac{1}{X^2}$，则

(A) $Y\sim\chi^2(n)$. (B) $Y\sim\chi^2(n-1)$. (C) $Y\sim F(n,1)$. (D) $Y\sim F(1,n)$.

[答案] (C).

[提示] 本题考查 $t(n)$分布与 $\chi^2(n)$分布的定义，以及 $F(n_1,n_2)$的定义.

[解] 设 $W\sim N(0,1)$，$Z\sim\chi^2(n)$，且 $W$、$Z$ 独立. 则 $X=\frac{W}{\sqrt{Z/n}}\sim t(n)$. 又 $W\sim N(0,1)$，所以 $W^2\sim\chi^2(1)$. 所以$\frac{1}{X^2}=\frac{Z/n}{W^2/1}\xrightarrow{\text{按定义}}F(n,1)$.

[典型错误] 因为有相当一些考生只是在辅导班上学了点数理统计，理解不深，所以无法做此题，乱填一气.

例6.6 设随机变量 $X$ 服从正态分布 $N(0,1)$，对给定的 $\alpha$ $(0<\alpha<1)$. 数 $u_\alpha$ 满足 $P\{X>u_\alpha\}=\alpha$. 若 $P\{|X|<x\}=\alpha$，则 $x$ 等于

(A) $u_{\frac{\alpha}{2}}$. (B) $u_{1-\frac{\alpha}{2}}$. (C) $u_{\frac{1-\alpha}{2}}$. (D) $u_{1-\alpha}$.

[答案] (C).

[提示] 本题考查概率运算及标准正态分布概率密度的性质，题中的 $u_\alpha$ 实际上就是标准正态分布的上侧分位数.

[解] 由标准正态分布概率密度的对称性，可知
$$P\{X\leqslant -x\}=P\{X\geqslant x\},$$

则
$$\alpha=P\{|X|<x\}=1-P\{X\geqslant x\}$$
$$=1-[P\{|X|\geqslant x\}+P\{X\leqslant -x\}]$$
$$=1-2P\{X>x\},$$

故
$$P\{X\geqslant x\}=P\{X\geqslant x\}=\frac{1-\alpha}{2},$$

再由 $u_\alpha$ 的定义知
$$x=u_{\frac{1-\alpha}{2}}.$$

[典型错误] 没能正确理解 $u_\alpha$ 的定义，从而无从下手解答本题.

例6.7 设$X_1,X_2,\cdots,X_n$ $(n\geqslant2)$为来自总体 $N(0,1)$的简单随机样本，$\overline{X}$ 为样本均值. $S^2$ 为样本方差，则

(A) $n\overline{X}\sim N(0,1)$. (B) $nS^2\sim\chi^2(n)$.

(C) $\frac{(n-1)\overline{X}}{S}\sim t(n-1)$. (D) $\frac{(n-1)X_1^2}{\sum\limits_{i=2}^{n}X_i^2}\sim F(1,n-1)$.

[答案] (D).

[提示] 本题涉及的知识点有标准正态分布、$\chi^2$ 分布、$t$ 分布与$F$ 分布及它们之间的关系.

［解］ 由于 $X_i \sim N(0,1)$，$i=1,2,\cdots,n\ (n\geqslant 2)$，故 $X_1^2 \sim \chi^2(1)$，$\sum\limits_{i=2}^{n} X_i^2 \sim \chi^2(n-1)$，且 $X_1^2$ 与 $\sum\limits_{i=2}^{n} X_i^2$ 相互独立，于是 $\dfrac{X_1^2}{\sum\limits_{i=2}^{n} X_i^2/(n-1)} \sim F(1,n-1)$，即 $\dfrac{(n-1)X_1^2}{\sum\limits_{i=2}^{n} X_i^2} \sim F(1,n-1)$，故应选(D).

又 $n\overline{X} \sim N(0,n)$，$(n-1)S^2 \sim \chi^2(n-1)$，$\dfrac{X}{S/\sqrt{n}} \sim t(n-1)$，故(A)，(B)，(C)均不正确.

例6.8　设 $X_1,X_2,\cdots,X_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，$\overline{X}$ 是样本均值，记

$$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2,\qquad S_2^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2,$$

$$S_3^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2,\qquad S_4^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2,$$

则服从自由度为 $n-1$ 的 $t$ 分布的随机变量是

(A) $t=\dfrac{\overline{X}-\mu}{S_1/\sqrt{n-1}}$.　　(B) $t=\dfrac{\overline{X}-\mu}{S_2/\sqrt{n-1}}$.　　(C) $t=\dfrac{\overline{X}-\mu}{S_3/\sqrt{n}}$.　　(D) $t=\dfrac{\overline{X}-\mu}{S_4/\sqrt{n}}$.

［答案］ (B).

［提示］ 本题考查 $t$ 分布的概念以及相关性质.

［解］ 因为

$$U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1),\quad V=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1),$$

$U$ 与 $V$ 相互独立，所以

$$\frac{U}{\sqrt{\dfrac{V}{n-1}}}=\frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\Big/(n-1)}}=\frac{\overline{X}-\mu}{\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}\Big/\sqrt{n-1}}$$

$$=\frac{\overline{X}-\mu}{S_2/\sqrt{n-1}}\sim t(n-1).$$

故选(B).

例6.9　设随机变量 $X$ 和 $Y$ 都服从标准正态分布，则

(A) $X+Y$ 服从正态分布.　　(B) $X^2+Y^2$ 服从 $\chi^2$ 分布.

(C) $X^2$ 和 $Y^2$ 都服从 $\chi^2$ 分布.　　(D) $X^2/Y^2$ 服从 $F$ 分布.

［答案］ (C).

［提示］ 本题主要考查正态分布的性质以及 $\chi^2$ 分布和 $F$ 分布的定义.

［解］ 当随机变量 $X$ 和 $Y$ 都服从正态分布，且二者相互独立时，(A)、(B)、(C)、(D)四选项均成立. 当未给出 $X$、$Y$ 相互独立这一条件，(A)、(B)、(D)均不一定成立.

例6.10　设总体 $X$ 服从正态分布 $N(\mu,\sigma^2)(\sigma>0)$，从该总体中抽取简单随机样本 $X_1,X_2,\cdots,X_{2n}(n\geqslant 2)$，其样本均值为 $\overline{X}=\dfrac{1}{2n}\sum\limits_{i=1}^{2n}X_i$，求统计量 $Y=\sum\limits_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$ 的数学期望 $E(Y)$.

［提示］ 本题考查要点：①样本的独立性和同分布性质；②数学期望的计算.

如果将 $(X_1+X_{n+1})$，$(X_2+X_{n+2})$，…，$(X_n+X_{2n})$ 视作正态总体 $N(2\mu,2\sigma^2)$ 的 $n$ 个简单随机样本，容易计算出 $E(Y)$. 也可以将 $(X_i+X_{n+i}-2\overline{X})^2$ 分成 $[(X_i-\overline{X})+(X_{n+i}-\overline{X})]^2$，分别考虑 $X_1,X_2,\cdots,X_n$ 以及 $X_{n+1},X_{n+2},\cdots,X_{2n}$ 两个样本的性质. 注意 $X_1,X_2,\cdots,X_{2n}$ 是相互独立的.

［解法1］ 考虑 $(X_1+X_{n+1})$，$(X_2+X_{n+2})$，…，$(X_n+X_{2n})$，将其视为取自总体 $N(2\mu,2\sigma^2)$ 的简单随

机样本，则其样本均值为$\frac{1}{n}\sum_{i=1}^{n}(X_i+X_{n+i})=\frac{1}{n}\sum_{i=1}^{2n}X_i=2\overline{X}$，样本方差为$\frac{1}{n-1}Y$.

由于 $E\left(\frac{1}{n-1}Y\right)=2\sigma^2$，所以 $E(Y)=(n-1)(2\sigma^2)=2(n-1)\sigma^2$.

［解法 2］ 记

$$\overline{X}'=\frac{1}{n}\sum_{i=1}^{n}X_i,\ \overline{X}''=\frac{1}{n}\sum_{i=1}^{n}X_{n+i},$$

显然有 $2\overline{X}=\overline{X}'+\overline{X}''$. 因此

$$\begin{aligned}E(Y)&=E\left[\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2\right]\\&=E\left\{\sum_{i=1}^{n}[(X_i-\overline{X}')+(X_{n+i}-\overline{X}'')]^2\right\}\\&=E\left\{\sum_{i=1}^{n}[(X_i-\overline{X}')^2+2(X_i-\overline{X}')(X_{n+i}-\overline{X}'')+(X_{n+i}-\overline{X}'')]^2\right\}\\&=E\left[\sum_{i=1}^{n}(X_i-\overline{X}')^2\right]+0+E\left[\sum_{i=1}^{n}(X_{n+i}-\overline{X}'')^2\right]\\&=(n-1)\sigma^2+(n-1)\sigma^2\\&=2(n-1)\sigma^2.\end{aligned}$$

［典型错误］ ①相当多的考生将 $\overline{X}=\frac{1}{2n}\sum_{i=1}^{2n}X_i$ 写成$\frac{1}{n}\sum_{i=1}^{n}X_i$；②$E(X_i\overline{X})$认为等于 $E(X_i)E(\overline{X})$，其实 $X_i$ 与 $\overline{X}$ 并不独立，所以二者并不相等；③$D(X_i+\overline{X})$认为等于 $D(X_i)+D(\overline{X})$. 错误性质同②；$D(X_i-2\overline{X})=D(X_i)-4D(\overline{X})$. 这里比②多一层错误.

例6.11 设$X_1,X_2,\cdots,X_9$是来自正态总体 $X$ 的简单随机样本，

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

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

证明统计量 $Z$ 服从自由度为 2 的 $t$ 分布.

［提示］ 要证明统计量 $Z$ 服从自由度为 2 的 $t$ 分布，必须证明 $Z=\frac{U}{\sqrt{\chi^2/2}}$，其中 $U\sim N(0,1)$，$\chi^2$ 服从自由度为 2 的 $\chi^2$ 分布.

［证明］ 记 $D(X)=\sigma^2$（未知）. 易见

$$E(Y_1)=E(Y_2),\ D(Y_1)=\sigma^2/6,\ D(Y_2)=\sigma^2/3.$$

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 $\chi^2$ 分布.

由于 $Y_1$ 与 $Y_2$，$Y_1$ 与 $S^2$ 独立，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.

**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 $\theta$；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 $\theta$，用它的观测值$\hat{\theta}(x_1,x_2,\cdots,x_n)$作为未知参数 $\theta$ 的近似值. 称 $\hat{\theta}(X_1,X_2,\cdots,X_n)$为未知参数 $\theta$ 的估计量，$\hat{\theta}(x_1,x_2,\cdots,x_n)$为 $\theta$ 的估计值. 在不至于混淆的情况下将 $\theta$ 的估计量和估计值统称为$\theta$ 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