第 6 章　数理统计的基本概念

6.1　知识要点精讲

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；χ^2分布；t分布；F分布; 分位数；正态总体的常用抽样分布.

一、　总体、个体和样本

1. **总体**

   研究对象的全体称为总体, 它是一个随机变量, 记为X.

   (1)有限总体－所包含的个体的数量是有限的总体.

   (2)无限总体－所包含的个体的数量是无限的总体.

2. **个体**

   组成总体的每一个元素称为个体.

3. **样本**

   从总体X中按某种方式抽取若干个个体所组成的集合是样本. 个体的个数n称为样本容量(或样本大小). 容量为n的样本记为$(X_1, X_2, \cdots, X_n)$.

4. **简单随机样本**

   来自总体X的n个相互独立且与总体同分布的随机变量$X_1, X_2, \cdots, X_n$称为容量为n的简单随机样本, 简称样本.

   (1)样本观察值

   n次观察完成, 它们的观察值$(x_1, x_2, \cdots, x_n)$称为样本$(X_1, X_2, \cdots, X_n)$的观察值, 简称样本值, 也称为总体X的n个独立观察值.

   (2)样本的二重性

   假设$X_1, X_2, \cdots, X_n$是从总体X中抽取的样本.

样本值$x_1, x_2, \cdots, x_n$一数的属性;

随机变量$X_1, X_2, \cdots, X_n$一随机变量的属性.

特别地, 在以后凡是我们脱离开具体的一次观测或试验来谈及样本$X_1, X_2, \cdots, X_n$时, 它们总是被看做随机变量.

5. **随机变量$(X_1, X_2, \cdots, X_n)$的联合分布函数**

设总体X的分布函数为$F(x)$, 则样本$(X_1, X_2, \cdots, X_n)$的联合分布函数为

$$F(x_1, x_2, \cdots, x_n) = \prod_{k=1}^{n} F(x_k).$$

(1)若总体X具有密度函数$f(x)$, 则样本的联合密度函数为

$$f(x_1, x_2, \cdots, x_n) = \prod_{k=1}^{n} f(x_k).$$

(2)若离散总体X具有概率分布律: $P\{X = x_k\} = p_k, k = 1, 2, \cdots$, 则样本的联合分布律为

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n = \prod_{k=1}^{n} P\{X_k = x_k\} = \prod_{k=1}^{n} p_k.$$

6. **经验分布函数**

设$X_1, X_2, \cdots, X_n$是来自总体X的一个样本, 函数

$$F_n(x) = \frac{1}{n}\{X_1, X_2, \cdots, X_n\text{中小于或等于}x\text{的个数}\}(-\infty < x < +\infty),$$

称为样本经验分布函数.

当样本观察值$x_1, x_2, \cdots, x_n$取定时,

$$\overline{F}_n(x) = \frac{1}{n}\{x_1, x_2, \cdots, x_n\text{中小于或等于}x\text{的个数}\}(-\infty < x < +\infty),$$

称为经验分布函数$F_n(x)$的观察值, $F_n(x)$等于样本的n个观察值中不超过x的个数除以样本容量n, 它以概率1一致收敛于分布函数$F(x)$.

7. **分位数(点)**

设随机变量X的分布函数$F(x)$满足

$$P\{X > x_\alpha\} = \alpha,\ 0 < \alpha < 1,$$

的点x_α称为X的上α分位数.

注： 下α分位数: 设随机变量X的分布函数$F(x)$满足

$$P\{X \le x_\alpha\} = \alpha,\ 0 < \alpha < 1,$$

的点x_α称为X的下α分位数.

二、统计量

1. **定义**

   不含总体分布中任何未知参数的样本$X_1, X_2, \cdots, X_n$的任意函数$g(X_1, X_2, \cdots, X_n)$称为统计量. 若$x_1, x_2, \cdots, x_n$是样本$(X_1, X_2, \cdots, X_n)$的一个观察值, 则称$g(x_1, x_2, \cdots, x_n)$是统计量$g(X_1, X_2, \cdots, X_n)$的观察值.

2. **常用统计量**

   设$X_1, X_2, \cdots, X_n$是来自总体X的一个样本, $x_1, x_2, \cdots, x_n$是这一样本的一个观察值.

   (1)样本平均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

   (2)样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$;

   (3)样本标准差 $S = \frac{1}{n-1}\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}$;

   (4)样本k阶原点矩 $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, k = 1, 2, \cdots$;

   (5)样本k阶中心矩 $B_k = M_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k, k = 1, 2, \cdots$;

   (6)次序统计量

   设样本$(X_1, X_2, \cdots, X_n)$的观察值为$(x_1, x_2, \cdots, x_n)$, 将$x_1, x_2, \cdots, x_n$按从小到大的次序重新排列, 得到$x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$, 记取值为$x_{(i)}$的样本分量为$X_{(i)}$, 则称$X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$为样本$(X_1, X_2, \cdots, X_n)$的次序统计量.

   $X_{(i)}$为第i个次序统计量.

   $X_{(1)}$是最小次序统计量;

   $X_{(n)}$是最大次序统计量.

3. **抽样分布**

   统计量的分布称为抽样分布.

三、三大抽样分布

1. χ^2分布

(1)定义

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

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

所服从的分布称为自由度为n的χ^2分布, 记为$\chi^2 \sim \chi^2(n)$.

(2)密度函数

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

其中Γ函数为$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} \mathrm{d}x$, 且具有性质

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha), \Gamma(n+1) = n!, n \in N$.

(3)分布的“可加性”：若$X \sim \chi^2(n)$, $Y \sim \chi^2(m)$, 且X与Y独立, 则

$$X + Y \sim \chi^2(n+m).$$

(4)数字特征：$E(X) = n, D(X) = 2n$.

(5)χ^2分布上α分位数$\chi^2_\alpha(n)$.

设随机变量$X \sim \chi^2(n)$, 对于任意给定的$\alpha(0 < \alpha < 1)$, 称满足条件

$$P\{X > \chi^2_\alpha(n)\} = \alpha,$$

的点$\chi^2_\alpha(n)$为$\chi^2(n)$的上α分位数, 且当$n > 45$时,

$$\chi^2_\alpha(n) \approx \frac{1}{2}(x_\alpha + \sqrt{2n-1})^2.$$

其中x_α是标准正态分布的上α分位数.

注： (1)χ^2分布是一种非负连续型随机变量的分布, 其密度函数的图形位于第一象限, 峰值随着n的增大向右移动.

(2)$\chi^2(2) = E(1/2)$, 即自由度为2的χ^2分布是参数为1/2的指数分布.

2. t分布

(1)定义

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

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

所服从的分布称为自由度为n的t分布, 记为$T \sim t(n)$.

(2)密度函数

$$f_n(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},\ -\infty < x < +\infty.$$

(3)性质

1)数字特征：$E(T) = 0, D(T) = \frac{n}{n-2}, n > 2.$

2)$\lim_{n\to} +\infty f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} = \phi(x)$. 即当$n$足够大时, $t(n)$近似服从$N(0,1)$.

(4)t分布上α分位数$t_\alpha(n)$.

设随机变量$X \sim t(n)$, 对于任意给定的$\alpha(0 < \alpha < 1)$, 称满足条件

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

的点$t_\alpha(n)$为$t(n)$的上α分位数, 且有

1)$t_{1-\alpha}(n) = -t_\alpha(n)$, $t_{0.5}(n) = 0$.

2)当$n > 45$时, $t_\alpha(n) = x_\alpha$. 其中x_α是标准正态分布的上α分位数.

注：t分布是一种连续型随机变量的分布, 密度函数的图形关于y对称, 形状与标准正态分布相类似.

3. F**分布**(1)定义

设随机变量$X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 且X与Y独立, 则随机变量

$$T = \frac{X/n_1}{\sqrt{Y/n_2}}$$

所服从的分布称为自由度为(n_1, n_2)的t分布, 记为$F \sim F(n_1, n_2)$.

(2)密度函数

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

(3)性质

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

2)若$X \sim t(n)$, 则$X^2 \sim F(1,n)$.

3)数字特征：$E[F(n_1,n_2)] = \dfrac{n_2}{n_2-2}$, $n_2 > 2$, $D[F(n_1,n_2)] = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$, $n_2 > 4$.

4)当$n_1 \geq 3$时, 极大值点$x^* = \dfrac{n_1-2}{n_1} \cdot \dfrac{n_2}{n_2+2}$.

(4)F分布上α分位数$F_\alpha(n_1,n_2)$.

设随机变量$X \sim F(n_1,n_2)$, 对于任意给定的$\alpha(0<\alpha<1)$, 称满足条件

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

的点$F_\alpha(n_1,n_2)$为$F(n_1,n_2)$的上α分位数, 且有

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

注：F分布是一种连续型随机变量的分布, 其密度函数包含两个参数.

四、正态总体的常用统计量的分布

1. 单个正态总体的抽样分布

设$X_1,X_2,\cdots,X_n$是来自正态总体$N(\mu,\sigma^2)$的样本, $\overline{X}$, S^2分别是样本均值与样本方差, 则

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

(2) $\overline{X}$与S^2相互独立, 并且$E(\overline{X}) = \mu$, $E(S^2) = \sigma^2$,

(3) $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi^2(n-1)$.

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

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

2. 两个正态总体的抽样分布

设$X_1,X_2,\cdots,X_{n_1}$是来自正态总体$N(\mu_1,\sigma_1^2)$的样本, $Y_1,Y_2,\cdots,Y_{n_2}$是来自正态总体$N(\mu_2,\sigma_2^2)$的样本, 并且这两个样本相互独立, 设

$$\overline{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i, \qquad S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2,$$

$$\overline{Y}=\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i,\qquad S_2^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2.$$

分别是两个样本的样本均值与样本方差, 则

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

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

(3)$\dfrac{(n_1-1)S_1^2}{\sigma_1^2}+\dfrac{(n_2-1)S_2^2}{\sigma_2^2}\sim\chi^2(n_1+n_2-2).$

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

(5)当$\sigma_1^2=\sigma_2^2=\sigma_2$时, $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2),$

其中$S_w=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}.$

五、最大、最小次序统计量的分布

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$X_{(1)}=\min(X_1,X_2,\cdots,X_n)$和$X_{(n)}=\max(X_1,X_2,\cdots,X_n)$的分布函数分别为

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

6.2　补充注释与重要结论

1. 统计量是随机变量, 本身不含总体分布中任何未知参数, 但它的分布可能含总体分布中的未知参数.

2. 若$(X_1, X_2, \cdots, X_n)$是来自总体X的样本, 则$X_1, X_2, \cdots, X_n$相互独立且与总体同分布.
3. 无论总体是否为正态分布, 样本$(X_1, X_2, \cdots, X_n)$取自总体X, 且$E(X) = \mu$, $D(X) = \sigma^2$, 则
   1)$E(\overline{X}) = \mu$, $D(\overline{X}) = \dfrac{\sigma^2}{n}$.
   2)$E(S^2) = \sigma^2$, $E(S^{*2}) = \frac{n-1}{n}\sigma^2$, 其中$S^{*2} = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2$.
4. 三大抽样分布的有关性质.
   1)若$X \sim \chi^2(n)$, 则$E(X) = n$, $D(X) = 2n$.
   2)若$X \sim \chi^2(n)$, $Y \sim \chi^2(m)$, 且X与Y独立, 则$X + Y \sim \chi^2(n + m)$.
   3)对于$T \sim t(n)$, 有$E(T) = 0$, $D(T) = \frac{n}{n-2}$, $n > 2$, $T^2 \sim F(1, n)$.
   4)对于$F \sim F(m, n)$, 有$\frac{1}{F} \sim F(n, m)$, $F_{\alpha/2}(m, n) = \dfrac{1}{F_{1-\alpha/2}(n, m)}$.
5. 若总体$X \sim N(\mu, \sigma^2)$, 则$D(S^2) = \dfrac{2\sigma^4}{n-1}$.
6. 若总体X的k阶原点矩$E(X^k) = \mu_k$存在, $X_1, X_2, \cdots, X_n$是取自总体X的样本, 则
   1)$E(X_1^k) = E(X_2^k) = \cdots = E(X_n^k) = \mu_k$.
   2)$A_j = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^j \xrightarrow{P} \mu_j$, $j = 1, 2, \cdots, k$.
   3)若g为连续函数, 则$g(A_1, A_2, \cdots, A_k) \xrightarrow{P} g(\mu_1, \mu_2, \cdots, \mu_k)$.
7. 设$X_1, X_2, \cdots, X_{n_1}$是来自正态总体$N(\mu_1, \sigma^2)$的样本, $Y_1, Y_2, \cdots, Y_{n_2}$是来自正态总体$N(\mu_2, \sigma^2)$的样本, 并且这两个样本相互独立, 设
$$\overline{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i,\quad S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_i - \overline{X})^2,$$
$$\overline{Y} = \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i,\quad S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(Y_i - \overline{Y})^2.$$
   分别是两个样本的样本均值与样本方差, $S_w = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$, 则$E(S_w) = \sigma^2$.

6.3 典型题型与例题分析

题型一、求样本容量n, 或与样本均值$\overline{X}$和样本方差S^2有关的概率

> 解题提示
>
> 若总体$X \sim N(\mu, \sigma^2)$, 则从$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, $P\left\{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le x\right\} = \Phi(x)$着手分析.

【例6.1】设总体X服从正态分布$N(72, 100)$, 为使样本均值大于70的概率不小于90%, 样本容量至少应取多大?

【例6.2】设总体$X \sim N(\mu, 2^2)$, $X_1, X_2, \cdots, X_n$是来自总体的一个样本, $\overline{X}$为样本均值, 问样本容量n至少应取多大才能使(1)$E(|\overline{X}-\mu|) \le 0.1$, (2)$E(\overline{X}-\mu)^2 \le 0.1$.

【例6.3】设总体$X \sim N(\mu, \sigma^2)$, $X_1, X_2, \cdots, X_{16}$是来自总体的一个样本, $\overline{X}$为样本均值, S^2为样本方差, 求k使$P\{\overline{X} > \mu + kS\} = 0.95$.

【例6.4】在天平上重复量一重为a的物品, 假设各次称量结果相互独立且服从正态分布$N(a, 0.2^2)$, 若以$\overline{X}_n$表示n次得量结果的算术平均值, 则为使

$$P\{|\overline{X}_n - a| < 0.1\} \ge 0.95,$$

n的最小值应不小于自然数_______.

【例6.5】从正态总体$X \sim N(\mu, 0.5^2)$中抽取样本$X_1, X_2, \cdots, X_{10}$,

(1)已知$\mu = 0$, 求概率$P\{\sum\limits_{i=1}^{10} X_i^2\} \ge 4$;

(2)μ未知, 求概率$P\{\sum\limits_{i=1}^{10} (X_i - \overline{X})^2\} \ge 2.85$.

题型二、求统计量的数字特征

> 解题提示
>
> 统计量是随机变量, 从而可求其数字特征, 通常要综合利用统计量的性质和数字特征的性质进行求解.

【例6.6】设总体$X\sim N(\mu,\sigma^2)$, $X_1,X_2,\cdots,X_n$是来自总体的一个样本, $\overline{X}$为样本均值, S^2为样本方差, 则$E(S^4)=$ ______.

【例6.7】设总体X服从正态分布$N(\mu,\sigma^2)(\sigma>0)$, 从该样本中抽取简单随机样本$X_1,X_2,\cdots,X_{2n}(n\geq 2)$, 其样本均值$\overline{X}=\frac{1}{2n}\sum\limits_{i=1}^{2n}X_i$, 求统计量$Y=\sum\limits_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$的数学期望.

【例6.8】设总体X服从分布$\chi^2(n)$, $X_1,X_2,\cdots,X_n$为从该总体中抽取简单随机样本, 其样本均值$\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$, 求$E\dfrac{(X_1+X_2+\cdots+X_m)}{n\overline{X}}$.

【例6.9】设$X_1,X_2,\cdots,X_n$是总体为$N(\mu,\sigma^2)$的简单随机样本, 记

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

(1)证明T是μ^2的无偏估计量, 即证明$E(T)=\mu^2$.

(2)当$\mu=0,\sigma=1$时, 求$D(T)$.

题型三、求统计量的分布

> 解题提示
>
> 利用χ^2分布, t分布和F分布的典型结构以及正态总体下常用统计量的性质求解.

【例6.10】设总体$X \sim N(\mu,\sigma^2)$, $X_1,X_2,\cdots,X_n$是来自总体X的一个简单随机样本, $\overline{X}$为样本均值, 记

$$S_1^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2,\quad S_2^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2,$$
$$S_3^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2,\quad S_4^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

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)$T=\dfrac{\overline{X}-\mu}{S_3/\sqrt{n-1}}$.　　(D)$T=\dfrac{\overline{X}-\mu}{S_4/\sqrt{n-1}}$.

【例6.11】设随机变量$X \sim t(n)(n>1)$, 则$Y=\frac{1}{X^2}$

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

【例6.12】设总体$X \sim N(0,1)$, 从此总体中抽取一个容量为6的样本$(X_1,X_2,X_3,X_4,X_5,X_6)$, 设$Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$, 试确定常数$C$, 使得随机变量$CY$服从$\chi^2$分布.

【例6.13】设总体$X \sim N(\mu,\sigma^2)$, $X_1,X_2,\cdots,X_n$是来自总体X的一个简单随机样本, $\overline{X}$为样本均值, S^2为样本方差, X_{n+1}是对X的又一次独立观察. 试证明统计量

$$T=\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}$$

服从t分布, 自由度为$n-1$.

【例6.14】设$X_1,X_2,\cdots,X_5$是相互独立的随机变量, 且每一个都服从标准正态分布, 求常数c, 使$\dfrac{c(X_1+X_2)}{\sqrt{X_3^2+X_4^2+X_5^2}}$服从$t$分布.

习题精选六

一、填空题

1. 设X_1, X_2, X_3, X_4是来自正态总体$N(0, 2^2)$的样本，则统计量

$$Y = \frac{1}{20}(X_1 - 2X_2)^2 + \frac{1}{100}(3X_3 - 4X_4)^2$$

服从_______分布，参数为_______.

2. 设(X_1, X_2, X_3, X_4)是来自总体$N(0, 1)$的一个样本，则统计量$\dfrac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}}$服从_______分布，参数为_______.

3. 设(X_1, X_2)是来自总体$N(0, \sigma^2)$的一个样本，则统计量$\dfrac{(X_1 + X_2)^2}{(X_1 - X_2)^2}$服从_____分布，参数为_______.

4. 设随机变量$X \sim F(n, n)$, 则概率$P(X < 1)$=_______.

二、选择题

1. 设随机变量$X \sim N(1, 4)$，$(X_1, X_2, \cdots, X_{100})$是来自总体$X$的一个样本，$\overline{X}$是样本均值，已知$Y = a\overline{X} + b \sim N(0, 1)$, 则【　　】

(A) $a = -5, b = 5$.　　(B) $a = 5, b = 5$.

(C)$a = \frac{1}{5}, b = -\frac{1}{5}$.　　(D)$a = -\frac{1}{5}, b = \frac{1}{5}$.

2. 设总体$X \sim N(\mu, \sigma^2)$, $(X_1, X_2, \cdots, X_n)$是来自总体X的一个样本，$\overline{X}$是样本均值，则【　　】

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

(C)$\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2(n-1)$.　　(D)$\dfrac{n-1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2(n-1)$.

3. 设X_1, X_2是来自总体$N(\mu, \sigma^2)$的样本，则$X_1 + X_2$与$X_1 - X_2$必【　　】

(A)不相关.　　(B)线性相关.　　(C)相关但非线性相关.　　(D)不独立.

4. 设$(X_1, X_2, \cdots, X_n)$是来自总体$N(\mu, \sigma^2)$的一个样本，统计量$Y = n\left(\dfrac{\overline{X} - \mu}{S}\right)^2$, 则有【　　】

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

5. 设随机变量$X \sim N(0, 1)$和$Y \sim N(0, 2)$, 并且相互独立，则【　　】

(A)$\frac{1}{3}X^2+\frac{2}{3}Y^2$服从$\chi^2$分布.　　(B)$\frac{1}{3}(X+Y)^2$服从$\chi^2$分布.

(C)$\frac{1}{2}X^2+\frac{1}{2}Y^2$服从$\chi^2$分布.　　(D)$\frac{1}{2}(X+Y)^2$服从$\chi^2$分布.

三、解答题

1. 设总体$X\sim B(1,p)$, $X_1,X_2,\cdots,X_n$是来自总体X的样本，
   (1)求$(X_1,X_2,\cdots,X_n)$的分布律；
   (2)求$\sum\limits_{i=1}^{n}X_i$的分布律；
   (3)求$E(\overline{X}),D(\overline{X}),E(S^2)$.
2. 设$X_1,X_2,\cdots,X_9$是来自正态总体$N(0,3^2)$的样本，求系数a,b,c使统计量$Y=a(X_1+X_2)^2+b(X_3+X_4+X_5)^2+c(X_6+X_7+X_8+X_9)^2$服从$\chi^2$分布，并求其自由度.