

第六章 数理统计的基本概念

前面五章我们讲述了概率论的基本内容,随后的五章将讲述数理统计的基本理论与方法.数理统计是以概率论为理论基础的一个数学分支.它是从实际观测的数据出发研究随机现象的规律性.在科学研究中,数理统计占据一个十分重要的位置,是多种试验数据处理的理论基础.

数理统计的内容很丰富,本书只介绍参数估计、假设检验、方差分析及回归分析的部分内容.

本章中首先讨论总体、随机样本及统计量等基本概念,然后着重介绍几个常用的统计量及抽样分布.

第一节 随机样本

假如我们要研究某厂所生产的一批电子元件的平均寿命,我们可以从这批产品中抽取一部分进行寿命测试,并且根据这部分产品的寿命数据对整批产品的平均寿命作一统计推断.

在数理统计中,我们将研究对象的某项数量指标值的全体称为总体(Population),总体中的每个元素称为个体(Individual).例如,上述的一批电子元件寿命值的全体就组成一个总体,其中每一只电子元件的寿命就是一个个体.要将一个总体的性质了解得十分清楚,初看起来,最理想的办法是对每个个体逐个进行观察,但实际上这样做往往是不现实的.例如,要研究电子元件的寿命,由于寿命试验是破坏性的,一旦我们获得实验的所有结果,这批电子元件也全报废了,我们只能从整批电子元件中抽取一部份显像管做寿命试验,并记录其结果,然后根据这部份数据来推断整批电子元件的寿命情况.由于电子元件的寿命在随机抽样中是随机变量,为了便于数学上处理,我们将总体定义为随机变量.随机变量的分布称为总体分布.

一般地,我们都是从总体中抽取一部分个体进行观察,然后根据所得的数据来推断总体的性质.被抽出的部分个体,叫做总体的一个样本.

所谓从总体抽取一个个体,就是对总体 X 进行一次观察(即进行一次试验),并记录其结果.我们在相同的条件下对总体 X 进行 n 次重复的、独立的观察,将 n 次观察结果按试验的次序记为 X_1, X_2, \dots, X_n .由于 X_1, X_2, \dots, X_n 是对随机变量 X 观察的结果,且各次观察是在相同的条件下独立进行的,于是我们引出以下的样本定义.

定义 6.1 设总体 X 是具有分布函数 F 的随机变量,若 X_1, X_2, \dots, X_n 是与 X 具有同一分布 $F(x)$, 且相互独立的随机变量,则称 X_1, X_2, \dots, X_n 为从总体 X 得到的容量为 n 的简单随机样本(Random sample), 简称为样本.

当 n 次观察一经完成,我们就得到一组实数 x_1, x_2, \dots, x_n .它们依次是随机变量 X_1, X_2, \dots, X_n 的观察值,称为样本值.

对于有限总体,采用放回抽样就能得到简单样本,当总体中个体的总数 N 比要得到的样本的容量 n 大得多时(一般当 $\frac{N}{n} \geq 10$ 时),在实际中可将不放回抽样近似地当作放回抽样来处理.

若 X_1, X_2, \dots, X_n 为总体 X 的一个样本, X 的分布函数为 $F(x)$, 则 X_1, X_2, \dots, X_n 的联合分布函数为

$$F^*(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F(x_i).$$

又若 X 具有概率密度 f , 则 X_1, X_2, \dots, X_n 的联合概率密度为

$$f^*(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i).$$

搜集到的资料如果未经组织和整理, 通常是没有什么价值的, 为了把这些有差异的资料组织成有用的形式, 我们应该编制频数表 (即频数分布表)。

例 6.1 某工厂的劳资部门为了研究该厂工人的收入情况, 首先收集了工人的工资资料, 表 6-1 记录了该厂 30 名工人未经整理的工资数值:

表 6-1

工人序号	工资 (元)	工人序号	工资 (元)	工人序号	工资 (元)
1	530	11	595	21	480
2	420	12	435	22	525
3	550	13	490	23	535
4	455	14	485	24	605
5	545	15	515	25	525
6	455	16	530	26	475
7	550	17	425	27	530
8	535	18	530	28	640
9	495	19	505	29	555
10	470	20	525	30	505

以下, 我们以例 6.1 为例介绍频数分布表的制作方法. 表 6-1 是 30 个工人月工资的原始资料, 这些数据可以记为 x_1, x_2, \dots, x_{30} , 对于这些观测数据, 进行如下操作.

第一步 确定最大值 x_{\max} 和最小值 x_{\min} , 根据表 6-1, 有

$$x_{\max}=640, x_{\min}=420.$$

第二步 分组, 即确定每一收入组的界限和组数, 在实际工作中, 第一组下限一般取一个小于 x_{\min} 的数, 例如, 我们取 400, 最后一组上限取一个大于 x_{\max} 的数, 例如取 650, 然后从 400 元到 650 元分成相等的若干段, 比如分成 5 段, 每一段就对应于一个收入组. 表 6-1 资料的频数分布表如表 6-2 所示.

表 6-2

组限	频数	累积频数
400~450	3	3
450~500	8	11
500~550	13	24
550~600	4	28
600~650	2	30

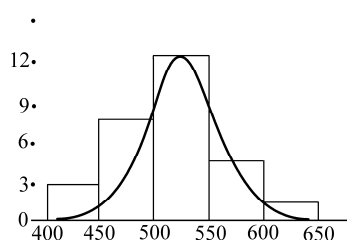


图 6-1

为了研究频数分布，我们可用图示法表示.

直方图 直方图是垂直条形图，条与条之间无间隔，用横轴上的点表示组限，纵轴上的单位数表示频数.与一个组对应的频数，用以组距为底的矩形（长条）的高度表示，表 6-2 资料的直方图如图 6-1 所示.

按上述方法我们对抽取数据加以整理，编制频数分布表，作直方图，画出频率分布曲线，于是就可以直观地看到数据分布的情况，在什么范围，较大、较小的各有多少，在哪些地方分布得比较集中，以及分布图形是否对称等等，所以，样本的频率分布是总体概率分布的近似.

样本是总体的反映，但是样本所含的信息不能直接用于解决我们所要研究的问题，而需要把样本所含的信息进行数学上的加工，使其浓缩起来，从而解决我们的问题.针对不同的问题构造样本的适当函数，利用这些样本的函数进行统计推断.

定义 6.2 设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本， $g(X_1, X_2, \dots, X_n)$ 是 X_1, X_2, \dots, X_n 的函数，若 g 中不含任何未知参数，则称 $g(X_1, X_2, \dots, X_n)$ 是一个**统计量** (Statistic).

设 x_1, x_2, \dots, x_n 是相应于样本 X_1, X_2, \dots, X_n 的样本值，则称 $g(x_1, x_2, \dots, x_n)$ 是 $g(X_1, X_2, \dots, X_n)$ 的**观察值**.

下面我们定义一些常用的统计量.设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本， x_1, x_2, \dots, x_n 是这一样本的观察值.定义

样本平均值

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i ;$$

样本方差

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left[\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right];$$

样本标准差

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2};$$

样本 k 阶（原点）矩

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k, \quad k=1, 2, \dots;$$

样本 k 阶中心矩

$$B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, \quad k=1, 2, \dots.$$

它们的观察值分别为

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i ;$$

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

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2};$$

$$a_k = \frac{1}{n} \sum_{i=1}^n x_i^k, \quad k=1, 2, \dots;$$

$$b_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k, \quad k=1, 2, \dots.$$

这些观察值仍分别称为样本均值、样本方差、样本标准差、样本 k 阶矩、样本 k 阶中心矩.

第二节 抽样分布

统计量是样本的函数, 它是一个随机变量. 统计量的分布称为抽样分布. 在使用统计量进行统计推断时常需知道它的分布. 当总体的分布函数已知时, 抽样分布是确定的, 然而要求出统计量的精确分布, 一般来说是困难的. 本节介绍来自正态总体的几个常用的统计量的分布.

1. χ^2 分布

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

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

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

$\chi^2(n)$ 分布的概率密度函数为

$$f(y) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} y^{n/2-1} e^{-y/2}, & y > 0, \\ 0, & \text{其他.} \end{cases}$$

$f(y)$ 的图形如图 6-2 所示.

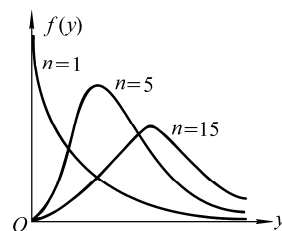


图 6-2

χ^2 分布具有以下性质:

(1) 如果 $\chi_1^2 \sim \chi^2(n_1)$, $\chi_2^2 \sim \chi^2(n_2)$, 且它们相互独立, 则有

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

这一性质称为 χ^2 分布的可加性.

(2) 如果 $\chi^2 \sim \chi^2(n)$, 则有

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

证 只证(2)因为 $X_i \sim N(0, 1)$, 所以

$$E(X_i^2) = D(X_i) = 1, E(X_i^4) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^4 e^{-\frac{x^2}{2}} dx = 3,$$

$$D(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = 3 - 1 = 2, i=1, 2, \dots, n.$$

于是

$$E(\chi^2) = E\left(\sum_{i=1}^n X_i^2\right) = \sum_{i=1}^n E(X_i^2) = n,$$

$$D(\chi^2) = D\left(\sum_{i=1}^n X_i^2\right) = \sum_{i=1}^n D(X_i^2) = 2n.$$

对于给定的正数 α , $0 < \alpha < 1$, 称满足条件

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

的点 $\chi_{\alpha}^2(n)$ 为 $\chi^2(n)$ 分布的上 α 分位点 (Percentile of α), 如图 6-3 所示, 对于不同的 α , n , 上 α 分位点的值已制成表格, 可以查用 (见附表 5), 例如对于 $\alpha=0.05, n=16$, 查附表得 $\chi_{0.05}^2(16)=26.296$. 但该表只详列到 $n=45$ 为止.

当 $n > 45$ 时, 近似地有 $\chi_{\alpha}^2(n) \approx \frac{1}{2}(z_{\alpha} + \sqrt{2n-1})^2$, 其中 z_{α} 是标准正态分布的上 α 分位点. 例如

$$\chi_{0.05}^2(50) \approx \frac{1}{2}(1.645 + \sqrt{99})^2 = 67.221.$$

2. t 分布

设 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 并且 X, Y 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 (t-distribution), 记为 $t \sim t(n)$.

$t(n)$ 分布的概率密度函数为

$$h(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}, \quad -\infty < t < \infty.$$

证明略.

图 6-4 中画出了当 $n=1, 10$ 时 $h(t)$ 的图形. $h(t)$ 的图形关于 $t=0$ 对称, 当 n 充分大时其图形类似于标准正态变量概率密度的图形. 但对于较小的 n , t 分布与 $N(0, 1)$ 分布相差很大 (见附表 2 与附表 4).

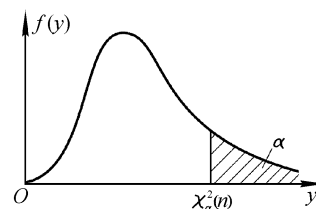


图 6-3

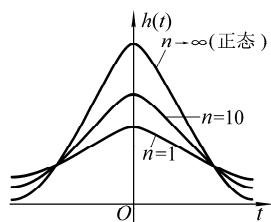


图 6-4

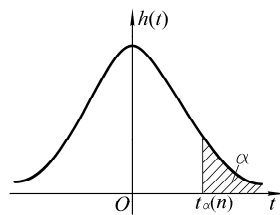


图 6-5

对于给定的 a , $0 < a < 1$, 称满足条件

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

的点 $t_{\alpha}(n)$ 为 $t(n)$ 分布的上 a 分位点 (见图 6-5)。

由 t 分布的上 a 分位点的定义及 $h(t)$ 图形的对称性知

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

t 分布的上 a 分位点可从附表 4 查得. 在 $n > 45$ 时, 就用正态分布近似:

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

3. F 分布

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

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

服从自由度为 (n_1, n_2) 的 F 分布 (F -distribution), 记 $F \sim F(n_1, n_2)$ 。

$F(n_1, n_2)$ 分布的概率密度为

$$\psi(y) = \begin{cases} \frac{\Gamma[(n_1 + n_2)/2] (n_1/n_2)^{n_1/2} y^{(n_1/2)-1}}{\Gamma(n_1/2) \Gamma(n_2/2) [1 + (n_1 y/n_2)]^{(n_1+n_2)/2}}, & y > 0, \\ 0, & \text{其他.} \end{cases}$$

证明略.

$\psi(y)$ 的图形如图 6-6 所示.

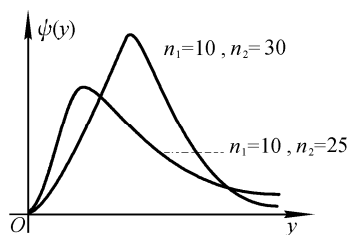


图 6-6

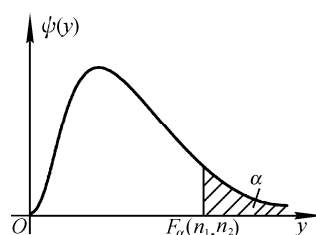


图 6-7

F 分布经常被用来对两个样本方差进行比较. 它是方差分析的一个基本分布, 也被用于回归分析中的显著性检验.

对于给定的 a , $0 < a < 1$, 称满足条件

$$P\{F > F_{\alpha}(n_1, n_2)\} = \int_{F_{\alpha}(n_1, n_2)}^{+\infty} \psi(y) dy = a$$

的点 $F_{\alpha}(n_1, n_2)$ 为 $F(n_1, n_2)$ 分布的上 a 分位点 (图 6-7). F 分布的上 a 分位点可查表 (见附表 6)。

F 分布的上 α 分位点有如下的性质:

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

这个性质常用来求 F 分布表中没有包括的数值.例如, 由附表 6 查得 $F_{0.05}(9,12)=2.80$, 则可利用上述性质求得

$$F_{0.95}(12,9)=1/F_{0.05}(9,12)=\frac{1}{2.80} \approx 0.357.$$

4. 正态总体的样本均值与样本方差的分布

设正态总体的均值为 μ , 方差为 σ^2 , X_1, X_2, \dots, X_n 是来自正态总体 X 的一个简单样本, 则总有

$$E(\bar{X}) = \mu,$$

$$D(\bar{X}) = \sigma^2/n,$$

$$\bar{X} \sim N(\mu, \sigma^2/n).$$

对于正态总体 $N(\mu, \sigma^2)$ 的样本方差 S^2 , 我们有以下的性质.

定理 6.1 设 X_1, X_2, \dots, X_n 是总体 $N(\mu, \sigma^2)$ 的样本, \bar{X} , S^2 分别是样本均值和样本方差, 则有

$$(1) \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1);$$

$$(2) \bar{X} \text{ 与 } S^2 \text{ 独立.}$$

证明略.

定理 6.2 设 X_1, X_2, \dots, X_n 是总体 $N(\mu, \sigma^2)$ 的样本, \bar{X} , S^2 分别是样本均值和样本方差, 则有

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

证 因为

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

且两者独立, 由 t 分布的定义知

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} / \sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}} \sim t(n-1).$$

化简上式左边, 即得

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

定理 6.3 设 X_1, X_2, \dots, X_{n_1} 与 Y_1, Y_2, \dots, Y_{n_2} 分别是来自具有相同方差的两正态

总体 $N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$ 的样本, 且这两个样本相互独立. 设 $\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i$,

$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i$ 分别是这两个样本的均值. $S_1^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$, $S_2^2 = \frac{1}{n_2-1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$ 分

别是这两个样本的样本方差, 则有:

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

其中

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

证明略.

本节所介绍的 3 个分布以及 3 个定理, 在下面各章中都起着重要的作用. 应注意, 它们都是在总体为正态总体这一基本假定下得到的.

例 6.2 设总体 X 服从正态分布 $N(62, 100)$, 为使样本均值大于 60 的概率不小于 0.95, 问样本容量 n 至少应取多大?

解 设需要样本容量为 n , 则

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n} \sim N(0,1),$$

$$P(\bar{X} > 60) = P\left\{\frac{\bar{X} - 62}{10} \cdot \sqrt{n} > \frac{60 - 62}{10} \cdot \sqrt{n}\right\}.$$

查标准正态分布表, 得 $\Phi(1.64) \approx 0.95$. 所以 $0.2\sqrt{n} \geq 1.64, n \geq 67.24$.

故样本容量至少应取 68.

小 结

在数理统计中往往研究有关对象的某一项数量指标, 对这一数量指标进行试验和观察, 将试验的全部可能的观察值称为总体, 每个观察值称为个体. 总体中的每一个个体是某一随机变量 X 的值, 因此一个总体对应于一个随机变量 X , 我们笼统称为总体 X . 随机变量 X 服从什么分布就称总体服从什么分布.

若 X_1, X_2, \dots, X_n 是相同条件下, 对总体 X 进行 n 次重复独立的观察所得到的 n 个结

果, 称随机变量 X_1, X_2, \dots, X_n 为来自总体 X 的简单随机样本, 它具有两条性质.

1. X_1, X_2, \dots, X_n 都与总体具有相同的分布;

2. X_1, X_2, \dots, X_n 相互独立.

我们就是利用来自样本的信息推断总体, 得到有关总体分布的种种结论.

完全由样本 X_1, X_2, \dots, X_n 所确定的函数 $g=g(X_1, X_2, \dots, X_n)$ 称为统计量, 统计量是一个随机变量. 它是统计推断的一个重要工具. 在数理统计中的地位相当重要, 相当于随机变量在概率论中的地位.

$$\text{样本均值} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\text{和样本方差} \quad S^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2$$

是两个最重要的统计量, 统计量的分布称为抽样分布, 读者需要掌握统计学中三大抽样分布:

χ^2 分布, t 分布, F 分布.

读者学习后续内容还需要掌握以下重要结果:

1. 设总体 X 的一个样本为 X_1, X_2, \dots, X_n . 且 X 的均值和方差存在.

记 $\mu=EX, \sigma^2=DX$. 则

$$E(\bar{X}) = \mu, \quad D(\bar{X}) = \sigma^2/n, \quad ES^2 =$$

2. 设总体 $X \sim N(\mu, \sigma^2)$, X_1, X_2, \dots, X_n 是 X 的一个样本, 则

(1) $\bar{X} \sim N(\mu, \sigma^2/n)$;

(2) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$;

(3) \bar{X} 和 S^2 相互独立;

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

3. 定理 6.3 的结果.

重要术语及主题

总体 样本 统计量

χ^2 分布、 t 分布、 F 分布的定义及它们的密度函数图形上的 α 分位点.

习 题 六

1. 设总体 $X \sim N(60, 15^2)$, 从总体 X 中抽取一个容量为 100 的样本, 求样本均值与总体均值之差的绝对值大于 3 的概率.

2. 从正态总体 $N(4.2, 5^2)$ 中抽取容量为 n 的样本, 若要求其样本均值位于区间 $(2.2, 6.2)$ 内的概率不小于 0.95, 则样本容量 n 至少取多大?

3. 设某厂生产的灯泡的使用寿命 $X \sim N(1000, \sigma^2)$ (单位: h), 随机抽取一容量为 9 的样本, 并测得样本均值及样本方差. 但是由于工作上的失误, 事后失去了此试验的结果, 只记得样本方差为 $S^2=1002$, 试求 $P(\bar{X} > 1062)$.

4. 从一正态总体中抽取容量为 10 的样本, 假定有 2% 的样本均值与总体均值之差的绝对值在 4 以上, 求总体的标准差.

5. 设总体 $X \sim N(\mu, 16)$, X_1, X_2, \dots, X_{10} 是来自总体 X 的一个容量为 10 的简单随机样本, S^2 为其样本方差, 且 $P(S^2 > a) = 0.1$, 求 a 之值.

6. 设总体 X 服从标准正态分布, X_1, X_2, \dots, X_n 是来自总体 X 的一个简单随机样本, 试问统计量

$$Y = \frac{\left(\frac{n}{5} - 1\right) \sum_{i=1}^5 X_i^2}{\sum_{i=6}^n X_i^2}, \quad n > 5$$

服从何种分布?

7. 求总体 $X \sim N(20, 3)$ 的容量分别为 10, 15 的两个独立随机样本平均值差的绝对值大于 0.3 的概率.

8. 设总体 $X \sim N(0, \sigma^2)$, $X_1, \dots, X_{10}, \dots, X_{15}$ 为总体的一个样本. 则 $Y = \frac{X_1^2 + X_2^2 + \dots + X_{10}^2}{2(X_{11}^2 + X_{12}^2 + \dots + X_{15}^2)}$

服从_____分布, 参数为_____.

(2001 研考)

9. 设总体 $X \sim N(\mu_1, \sigma^2)$, 总体 $Y \sim N(\mu_2, \sigma^2)$, X_1, X_2, \dots, X_{n_1} 和 Y_1, Y_2, \dots, Y_{n_2} 分别来自总体 X 和 Y 的简单随机样本, 则

$$E \left[\frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{n_1 + n_2 - 2} \right] = \underline{\hspace{2cm}}. \quad (2004 \text{ 研考})$$

10. 设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \dots, X_n, n \geq 2$ 是总体 X 的一个样本, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$,

令 $Y = \sum_{i=1}^n (X_i + X_{n+i} - 2\bar{X})^2$, 求 $E(Y)$. (2001 研考)

11. 设总体 X 的概率密度为 $f(x) = \frac{1}{2} e^{-|x|}$ ($-\infty < x < +\infty$), X_1, X_2, \dots, X_n 为总体 X 的简单随机样本, 其样本方差为 S^2 , 求 $E(S^2)$. (2006 研考)

