概率论与数理统计

概率论与数理统计教研团队[1]

[1] 数学科学学院
哈尔滨工程大学

2024 年　春

大工至善
大学至真

目　录

数理统计是数学的一个分支，分为描述统计和推断统计. 它以概率论为基础，研究大量随机现象的统计规律性.

描述统计的任务是搜集资料，进行整理、分组，编制次数分配表，绘制次数分配曲线，计算各种特征指标，以描述资料分布的集中趋势、离中趋势和次数分布的偏斜度等；推断统计是在描述统计的基础上，根据样本资料归纳出的规律性，对总体进行推断和预测.

数理统计的发展大致可分为古典时期、近代时期和现代时期三个阶段

1. **古典时期 (19 世纪以前)**
   - 描述性的统计学形成和发展阶段，是数理统计的萌芽时期.
2. **近代时期 (19 世纪末至 1945 年)**
   - 数理统计的主要分支建立，是数理统计的形成时期.
3. **现代时期 (1945 年以后)**
   - 数理统计在理论研究和应用方面不断地向纵深发展，并产生一些新的分支和边缘性的新学科，如最优设计和非参数统计推断等.

数理统计中，被研究的随机变量的分布一般是未知的，或者分布类型已知，但其中包含未知的参数，通常需要从所研究的对象全体中抽取一部分进行观测或试验，从而对分布或者未知参数做出统计推断.

数理统计研究的内容

(1) **试验的设计和研究**

- **如何有效地收集、整理、分析所获得的有限的资料.**

(2) **统计推断**

- **通过对收集的数据进行分析研究，从而对所研究的问题，尽可能地做出精确而可靠的结论.**

本章主要介绍统计推断的有关基本概念、基本理论和方法.

总体与样本

总体

1. **总体**

总体

1. **总体**
    - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**

总体

1. **总体**
    - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**

总体

1. **总体**
   - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**
   - **组成总体的每个成员称为个体.**

总体

1. **总体**
   - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**
   - **组成总体的每个成员称为个体.**
3. **总体容量**

总体

1. **总体**
   - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**
   - **组成总体的每个成员称为个体.**
3. **总体容量**
   - **总体中所包含的个体的个数称为总体的容量.**

总体

1. **总体**
    - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**
    - **组成总体的每个成员称为个体.**
3. **总体容量**
    - **总体中所包含的个体的个数称为总体的容量.**
4. **总体类型**

总体

1. **总体**
   - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**
   - **组成总体的每个成员称为个体.**
3. **总体容量**
   - **总体中所包含的个体的个数称为总体的容量.**
4. **总体类型**
   - **有限总体和无限总体.**

总体

1. **总体**
   - **在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.**
2. **个体**
   - **组成总体的每个成员称为个体.**
3. **总体容量**
   - **总体中所包含的个体的个数称为总体的容量.**
4. **总体类型**
   - **有限总体和无限总体.**

总体

1. **总体**
   - 在统计学中，把与所研究问题有关对象的全体构成的集合称为总体.
2. **个体**
   - 组成总体的每个成员称为个体.
3. **总体容量**
   - 总体中所包含的个体的个数称为总体的容量.
4. **总体类型**
   - 有限总体和无限总体.

例如　测量某班级学生的身高.
♣ 全班学生身高的测量值为总体；
♣ 每个学生身高的测量值为一个个体；
♣ 总体容量为全班学生人数，是有限的，此时为有限总体；
♣ 若将问题换成测量全国大学生的身高，样本容量为全国大学生的人数也是有限的，但是容量很大，此时也可以认为它是一个无限总体.

在数理统计研究中，人们往往关心研究对象的某一项（或几项）数量指标，对这一指标进行试验，观察试验结果，考察该数量指标的分布情况，每个具体的数量指标的全体就是总体，从而可以用随机变量及其分布来描述，因此理论上可以把总体与随机变量等同看待.

在数理统计研究中，人们往往关心研究对象的某一项（或几项）数量指标，对这一指标进行试验，观察试验结果，考察该数量指标的分布情况，每个具体的数量指标的全体就是总体，从而可以用随机变量及其分布来描述，因此理论上可以把总体与随机变量等同看待.

例如：研究某批次 LED 灯泡的寿命时，关心的数量指标就是寿命，于是此总体可以用随机变量 X 表示，或用其分布函数 $F(x)$ 表示. 鉴于此，常用随机变量的记号或用其分布函数表示总体，如总体 X 或总体 $F(x)$.

样本

总体分布一般是未知的，或者分布类型已知，但其中包含未知的参数，为了推断总体分布及各种特征，需按一定规则从总体中抽取部分个体进行观察试验，以获得有关总体的信息，这一抽取过程称为“抽样”，所抽取的部分个体称为样本，样本中包含个体的数目称为样本容量.

样本

总体分布一般是未知的，或者分布类型已知，但其中包含未知的参数，为了推断总体分布及各种特征，需按一定规则从总体中抽取部分个体进行观察试验，以获得有关总体的信息，这一抽取过程称为“抽样”，所抽取的部分个体称为样本，样本中包含个体的数目称为样本容量.

对于无限总体或总体容量远大于样本容量的有限总体，将不放回抽样近似看成有放回抽样.

样本

总体分布一般是未知的，或者分布类型已知，但其中包含未知的参数，为了推断总体分布及各种特征，需按一定规则从总体中抽取部分个体进行观察试验，以获得有关总体的信息，这一抽取过程称为“抽样”，所抽取的部分个体称为样本，样本中包含个体的数目称为样本容量.

对于无限总体或总体容量远大于样本容量的有限总体，将不放回抽样近似看成有放回抽样.

从总体 X 中按着某种方式进行抽样，抽取之前无法预知每个个体数量指标的取值，因此可用随机变量 $(X_1, X_2, \cdots, X_n)$ 表示样本容量为 n 的一个样本；抽取后经过试验得到一组确定的值 $(x_1, x_2, \cdots, x_n)$，称为样本容量为 n 的样本值或样本观测值.

样本

总体分布一般是未知的，或者分布类型已知，但其中包含未知的参数，为了推断总体分布及各种特征，需按一定规则从总体中抽取部分个体进行观察试验，以获得有关总体的信息，这一抽取过程称为“抽样”，所抽取的部分个体称为样本，样本中包含个体的数目称为样本容量.

对于无限总体或总体容量远大于样本容量的有限总体，将不放回抽样近似看成有放回抽样.

从总体 X 中按着某种方式进行抽样，抽取之前无法预知每个个体数量指标的取值，因此可用随机变量 $(X_1, X_2, \cdots, X_n)$ 表示样本容量为 n 的一个样本；抽取后经过试验得到一组确定的值 $(x_1, x_2, \cdots, x_n)$，称为样本容量为 n 的样本值或样本观测值.

简单随机样本

(1) 代表性：$X_1, X_2, \cdots, X_n$ 与总体 X 具有相同的分布；

样本

总体分布一般是未知的，或者分布类型已知，但其中包含未知的参数，为了推断总体分布及各种特征，需按一定规则从总体中抽取部分个体进行观察试验，以获得有关总体的信息，这一抽取过程称为“抽样”，所抽取的部分个体称为样本，样本中包含个体的数目称为样本容量.

对于无限总体或总体容量远大于样本容量的有限总体，将不放回抽样近似看成有放回抽样.

从总体 X 中按着某种方式进行抽样，抽取之前无法预知每个个体数量指标的取值，因此可用随机变量 $(X_1, X_2, \cdots, X_n)$ 表示样本容量为 n 的一个样本；抽取后经过试验得到一组确定的值 $(x_1, x_2, \cdots, x_n)$，称为样本容量为 n 的样本值或样本观测值.

简单随机样本

(1) 代表性：$X_1, X_2, \cdots, X_n$ 与总体 X 具有相同的分布；

(2) 独立性：$X_1, X_2, \cdots, X_n$ 相互独立.

例

检测某工厂生产的一批产品的质量，需进行抽样验收，请指出总体、个体及样本.

例

检测某工厂生产的一批产品的质量，需进行抽样验收，请指出总体、个体及样本.

解：设总体 X 表示产品的质量指标，可定义

$$X = \begin{cases} 1, & \text{产品为不合格品,} \\ 0, & \text{产品为合格品.} \end{cases}$$

例

检测某工厂生产的一批产品的质量，需进行抽样验收，请指出总体、个体及样本.

解：设总体 X 表示产品的质量指标，可定义

$$X=\begin{cases}1, & \text{产品为不合格品,}\\ 0, & \text{产品为合格品.}\end{cases}$$

个体与总体同分布，即个体表示为

$$X_i=\begin{cases}1, & \text{第 i 件产品为不合格品,}\\ 0, & \text{第 i 件产品为合格品.}\end{cases}$$

例

检测某工厂生产的一批产品的质量，需进行抽样验收，请指出总体、个体及样本.

解：设总体 X 表示产品的质量指标，可定义

$$X = \begin{cases} 1, & \text{产品为不合格品,} \\ 0, & \text{产品为合格品.} \end{cases}$$

个体与总体同分布，即个体表示为

$$X_i = \begin{cases} 1, & \text{第 i 件产品为不合格品,} \\ 0, & \text{第 i 件产品为合格品.} \end{cases}$$

样本容量为 n 的样本表示为 $(X_1, X_2, \cdots, X_n)$.

样本的联合分布

总体 X 的样本容量为 n 的样本 $(X_1, X_2, \cdots, X_n)$ 是一个 n 维随机变量，称 $X_1, X_2, \cdots, X_n$ 的联合分布为样本的联合分布.

样本的联合分布

总体 X 的样本容量为 n 的样本 $(X_1, X_2, \cdots, X_n)$ 是一个 n 维随机变量，称 $X_1, X_2, \cdots, X_n$ 的联合分布为样本的联合分布.

1. **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**

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

样本的联合分布

总体 X 的样本容量为 n 的样本 $(X_1, X_2, \cdots, X_n)$ 是一个 n 维随机变量，称 $X_1, X_2, \cdots, X_n$ 的联合分布为样本的联合分布.

1. 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

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

2. 设 X 为连续型总体，其概率密度为 $f(x)$，则样本 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度为

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

样本的联合分布

总体 X 的样本容量为 n 的样本 $(X_1, X_2, \cdots, X_n)$ 是一个 n 维随机变量，称 $X_1, X_2, \cdots, X_n$ 的联合分布为样本的联合分布.

1. 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

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

2. 设 X 为连续型总体，其概率密度为 $f(x)$，则样本 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度为

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

3. 设 X 为离散型总体，其分布律为 $P\{X = x_k\} = p_k (k = 1, 2, \cdots)$，则样本 $(X_1, X_2, \cdots, X_n)$ 的联合分布律为

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} = \prod_{i=1}^{n} P\{X = x_i\}$$

例 1

设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个样本，当总体 $X \sim P(\lambda)$ 和 $X \sim N(\mu, \sigma^2)$ 时，求样本的联合分布律或者联合概率密度.

例 1

设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个样本，当总体 $X \sim P(\lambda)$ 和 $X \sim N(\mu, \sigma^2)$ 时，求样本的联合分布律或者联合概率密度.

解：由 $X \sim P(\lambda)$，则其分布律为 $P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, \cdots$

例 1

设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个样本，当总体 $X \sim P(\lambda)$ 和 $X \sim N(\mu, \sigma^2)$ 时，求样本的联合分布律或者联合概率密度.

解：由 $X \sim P(\lambda)$，则其分布律为 $P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}, k = 0, 1, \cdots$

样本 $(X_1, X_2, \cdots, X_n)$ 的联合分布律为

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} = \prod_{i=1}^{n} P\{X = x_i\} = e^{-n\lambda} \frac{\lambda^{\sum\limits_{i=1}^{n} x_i}}{\prod\limits_{i=1}^{n} x_i!}$$

其中 $x_i = 0, 1, \cdots, i = 1, 2, \cdots, n$

例 1

设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个样本，当总体 $X \sim P(\lambda)$ 和 $X \sim N(\mu, \sigma^2)$ 时，求样本的联合分布律或者联合概率密度.

由 $X \sim N(\mu, \sigma^2)$，则其概率密度为 $f(x) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$.

例 1

设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个样本，当总体 $X \sim P(\lambda)$ 和 $X \sim N(\mu, \sigma^2)$ 时，求样本的联合分布律或者联合概率密度.

由 $X \sim N(\mu, \sigma^2)$，则其概率密度为 $f(x) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$.

样本 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度为

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{-\frac{\sum\limits_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

其中 $x_i \in \mathbb{R}, i = 1, 2, \cdots, n$

统计量

统计量

在实际应用中，当从某总体中抽取一个样本 $(X_1, X_2, \cdots, X_n)$ 后，并不能直接用它去对总体的有关性质和特征进行推断，这是因为样本虽然是从总体中获取的代表，含有总体性质的信息，但仍较分散. 为了推断总体，必须把分散在样本中的信息集中起来，针对不同的研究目的，构造不同的样本函数，这种函数在统计学中称为统计量.

统计量

在实际应用中，当从某总体中抽取一个样本 $(X_1, X_2, \cdots, X_n)$ 后，并不能直接用它去对总体的有关性质和特征进行推断，这是因为样本虽然是从总体中获取的代表，含有总体性质的信息，但仍较分散. 为了推断总体，必须把分散在样本中的信息集中起来，针对不同的研究目的，构造不同的样本函数，这种函数在统计学中称为统计量.

定义 1
设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的容量为 n 的一个样本，如果由此样本构造一个函数 $T = T(X_1, X_2, \cdots, X_n)$，不依赖于任何未知参数，则称函数 T 是一个统计量.

统计量

在实际应用中，当从某总体中抽取一个样本 $(X_1, X_2, \cdots, X_n)$ 后，并不能直接用它去对总体的有关性质和特征进行推断，这是因为样本虽然是从总体中获取的代表，含有总体性质的信息，但仍较分散. 为了推断总体，必须把分散在样本中的信息集中起来，针对不同的研究目的，构造不同的样本函数，这种函数在统计学中称为统计量.

定义 1

设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的容量为 n 的一个样本，如果由此样本构造一个函数 $T = T(X_1, X_2, \cdots, X_n)$，不依赖于任何未知参数，则称函数 T 是一个统计量.

通常，又称 $T = T(X_1, X_2, \cdots, X_n)$ 为样本统计量. 当获得样本的一组观测值 $x_1, x_2, \cdots, x_n$ 后，代入 T，计算出数值 $T(x_1, x_2, \cdots, x_n)$，获得了一个具体的统计量值.

例

设 $X_1, X_2, \cdots, X_n$ 是从正态总体 $N(\mu, \sigma^2)$ 中抽取的容量为 n 的一个样本，其中 μ, σ^2 未知，

$$\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\text{是统计量}$$

$$\frac{1}{n}(X_1 + X_2 + \cdots + X_n) - \mu\text{不是统计量}$$

常见的统计量——样本矩

定义 2

设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的容量为 n 的一个样本.

(1) **样本均值** $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$，反映总体 X 数学期望的信息，是最常用的统计量.

定义 2

设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的容量为 n 的一个样本.

(1) **样本均值** $\overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$，反映总体 X 数学期望的信息，是最常用的统计量.

(2) **样本方差** $S^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2$，反映的是总体 X 方差的信息，称 S 为样本标准差，样本方差和样本标准差是最常用的统计量.

常见的统计量——样本矩

定义 2

设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的容量为 n 的一个样本.

(1) **样本均值** $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，反映总体 X 数学期望的信息，是最常用的统计量.

(2) **样本方差** $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$，反映的是总体 X 方差的信息，称 S 为样本标准差，样本方差和样本标准差是最常用的统计量.

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

常见的统计量——样本矩

定义 2
设 $X_1, X_2, \cdots, X_n$ 是从总体 X 中抽取的容量为 n 的一个样本.

(1) **样本均值** $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$，反映总体 X 数学期望的信息，是最常用的统计量.

(2) **样本方差** $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$，反映的是总体 X 方差的信息，称 S 为样本标准差，样本方差和样本标准差是最常用的统计量.

(3) **样本 k 阶原点矩** $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k,\ k = 1, 2, \cdots$

(4) **样本 k 阶中心矩** $B_k = \dfrac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^k,\ k = 1, 2, \cdots$

显然，样本 1 阶原点矩 A_1 就是样本均值 $\overline{X}$；样本 2 阶中心矩 B_2 与样本方差 S^2 满足

$$B_2 = \frac{n-1}{n} S^2$$

设 $x_1, x_2, \cdots, x_n$ 是样本 $X_1, X_2, \cdots, X_n$ 的一组观测值，统计量的值

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \, , s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

也称为样本均值和样本方差.

定理 3
设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本，若总体 X 的数学期望 $E(X) = \mu$ 和方差 $D(X) = \sigma^2$ 存在，且 k 阶原点矩也存在，并设 $E(X^k) = \mu_k$，则样本矩具有如下性质

(1) $E(\overline{X}) = \mu, D(\overline{X}) = \dfrac{\sigma^2}{n}$;

(2) $E(A_k) = \mu_k$;

(3) $E(S^2) = \sigma^2$.

其中 $\overline{X}, S^2, A_k$ 分别为样本均值、样本方差、样本的 k 阶原点矩.

证明：(1) $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

证明：(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

$$E\left(\overline{X}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\mu$$

证明：(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

$$E\left(\overline{X}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\mu$$

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X)=\frac{\sigma^2}{n}$$

证明：(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

$$E\left(\overline{X}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\mu$$

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X)=\frac{\sigma^2}{n}$$

(2) $E(A_k)=\mu_k$

证明：(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

$$E\left(\overline{X}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\mu$$

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X)=\frac{\sigma^2}{n}$$

(2) $E(A_k)=\mu_k$

$$E\left(A_k\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^k\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=\frac{1}{n}\sum_{i=1}^{n}E(X^k)=\mu_k$$

证明：(1) $\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$

$$E\left(\overline{X}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X)=\mu$$

$$D\left(\overline{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X)=\frac{\sigma^2}{n}$$

(2) $E(A_k)=\mu_k$

$$E\left(A_k\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^k\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=\frac{1}{n}\sum_{i=1}^{n}E(X^k)=\mu_k$$

注： $X_1^k,X_2^k\cdots,X_n^k$ 独立同分布，且 k 阶原点矩存在，由辛钦大数定律有

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k\xrightarrow{P}\mu_k(n\to\infty)$$

(3) $E(S^2)=\sigma^2$，其中 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

(3) $E(S^2)=\sigma^2$，其中 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

因为

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\sum_{i=1}^{n}X_i^2-n\overline{X}^2$$

(3) $E(S^2)=\sigma^2$，其中 $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

因为

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\sum_{i=1}^{n}X_i^2-n\overline{X}^2$$

又因为

$$E(X_i^2)=D(X_i)+\left[E(X_i)\right]^2=\sigma^2+\mu^2, E(\overline{X}^2)=D(\overline{X})+\left[E(\overline{X})\right]^2=\frac{\sigma^2}{n}+\mu^2.$$

(3) $E(S^2)=\sigma^2$，其中 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

因为

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\sum_{i=1}^{n}X_i^2-n\overline{X}^2$$

又因为

$$E(X_i^2)=D(X_i)+\left[E(X_i)\right]^2=\sigma^2+\mu^2, E(\overline{X}^2)=D(\overline{X})+\left[E(\overline{X})\right]^2=\frac{\sigma^2}{n}+\mu^2.$$

所以有

$$\begin{aligned}E\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)&=E\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)=n(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\\&=(n-1)\sigma^2\end{aligned}$$

(3) $E(S^2)=\sigma^2$，其中 $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

因为

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\sum_{i=1}^{n}X_i^2-n\overline{X}^2$$

又因为

$$E(X_i^2)=D(X_i)+\left[E(X_i)\right]^2=\sigma^2+\mu^2, E(\overline{X}^2)=D(\overline{X})+\left[E(\overline{X})\right]^2=\frac{\sigma^2}{n}+\mu^2.$$

所以有

$$\begin{aligned}E\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)&=E\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)=n(\sigma^2+\mu^2)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\\&=(n-1)\sigma^2\end{aligned}$$

两边同除以 $n-1$ 得 $E(S^2)=\sigma^2$.

例 1

设总体 X 服从参数为 λ 的泊松分布, $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本, 样本均值为 $\overline{X}$, 样本方差为 S^2, 求: (1)$E\left(\overline{X}\right)$; (2)$D\left(\overline{X}\right)$; (3)$E\left(S^2\right)$; (4)$E\left(\overline{X}^2\right)$.

例 1

设总体 X 服从参数为 λ 的泊松分布，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，样本均值为 $\overline{X}$，样本方差为 S^2，求：(1)$E\left(\overline{X}\right)$；(2)$D\left(\overline{X}\right)$；(3)$E\left(S^2\right)$；(4)$E\left(\overline{X}^2\right)$.

解：总体为泊松分布，所以总体均值和方差为 $E(X)=\lambda, D(X)=\lambda$.

例 1

设总体 X 服从参数为 λ 的泊松分布，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，样本均值为 $\overline{X}$，样本方差为 S^2，求：(1)$E\left(\overline{X}\right)$；(2)$D\left(\overline{X}\right)$；(3)$E\left(S^2\right)$；(4)$E\left(\overline{X}^2\right)$.

解：总体为泊松分布，所以总体均值和方差为 $E(X)=\lambda, D(X)=\lambda$.

(1) $E\left(\overline{X}\right)=E(X)=\lambda$;

例 1

设总体 X 服从参数为 λ 的泊松分布，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，样本均值为 $\overline{X}$，样本方差为 S^2，求：(1)$E\left(\overline{X}\right)$；(2)$D\left(\overline{X}\right)$；(3)$E\left(S^2\right)$；(4)$E\left(\overline{X}^2\right)$.

解：总体为泊松分布，所以总体均值和方差为 $E(X) = \lambda, D(X) = \lambda$.

(1) $E\left(\overline{X}\right) = E(X) = \lambda$;

(2) $D\left(\overline{X}\right) = \dfrac{1}{n} D(X) = \dfrac{\lambda}{n}$;

例 1

设总体 X 服从参数为 λ 的泊松分布，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，样本均值为 $\overline{X}$，样本方差为 S^2，求：(1)$E\left(\overline{X}\right)$；(2)$D\left(\overline{X}\right)$；(3)$E\left(S^2\right)$；(4)$E\left(\overline{X}^2\right)$.

解：总体为泊松分布，所以总体均值和方差为 $E(X)=\lambda, D(X)=\lambda$.

(1) $E\left(\overline{X}\right)=E(X)=\lambda$;

(2) $D\left(\overline{X}\right)=\dfrac{1}{n}D(X)=\dfrac{\lambda}{n}$;

(3) $E\left(S^2\right)=D(X)=\lambda$;

例 1

设总体 X 服从参数为 λ 的泊松分布，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，样本均值为 $\overline{X}$，样本方差为 S^2，求：(1)$E\left(\overline{X}\right)$；(2)$D\left(\overline{X}\right)$；(3)$E\left(S^2\right)$；(4)$E\left(\overline{X}^2\right)$.

解：总体为泊松分布，所以总体均值和方差为 $E(X) = \lambda, D(X) = \lambda$.

(1) $E\left(\overline{X}\right) = E(X) = \lambda$;

(2) $D\left(\overline{X}\right) = \dfrac{1}{n} D(X) = \dfrac{\lambda}{n}$;

(3) $E\left(S^2\right) = D(X) = \lambda$;

(4) $E\left(\overline{X}^2\right) = D\left(\overline{X}\right) + \left[E\left(\overline{X}\right)\right]^2 = \dfrac{\lambda}{n} + \lambda^2$.

例 2

设总体 X 服从区间 $(0,\theta)$ 上的均匀分布，其中 $\theta > 0$ 是未知的参数，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，下列样本的函数中哪些是统计量，哪些不是.

$$T_1 = \frac{X_1 + X_2 + \cdots + X_6}{6}, \qquad T_2 = X_6 - \theta,$$

$$T_3 = X_6 - E(X_1), \qquad T_4 = \max(X_1, X_2, \cdots, X_6).$$

例 2

设总体 X 服从区间 $(0,\theta)$ 上的均匀分布，其中 $\theta > 0$ 是未知的参数，$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的一个容量为 n 的样本，下列样本的函数中哪些是统计量，哪些不是.

$$T_1 = \frac{X_1 + X_2 + \cdots + X_6}{6}, \qquad T_2 = X_6 - \theta,$$

$$T_3 = X_6 - E(X_1), \qquad T_4 = \max(X_1, X_2, \cdots, X_6).$$

解：T_1, T_4 是统计量，T_2, T_3 不是统计量.

抽样分布

统计量的分布称为抽样分布. 在使用统计量进行推断时，通常需要知道统计量的分布. 有很多推断是基于正态分布假设的，以标准正态分布为基石而构造的三个著名统计量在实际中有广泛的应用，它们被称为统计中的“三大抽样分布”.

χ^2 分布

定义 4

设 $X_1, X_2, \cdots, X_n$ 为来自标准正态总体 $N(0,1)$ 的样本，则统计量

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

服从自由度为 n 的 χ^2 分布，简记为 $\chi^2 \sim \chi^2(n)$ 分布.

χ^2 分布

定义 4

设 $X_1, X_2, \cdots, X_n$ 为来自标准正态总体 $N(0,1)$ 的样本，则统计量

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

服从自由度为 n 的 χ^2 分布，简记为 $\chi^2 \sim \chi^2(n)$ 分布.

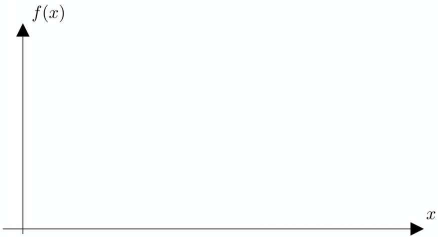
$\chi^2(n)$ 分布的概率密度函数

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

式中 $\Gamma(r) = \int_0^{+\infty} e^{-x} x^{r-1} \mathrm{d}x$ 为 Γ 函数, 其中 $r > 0$.

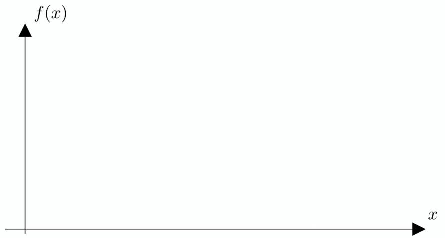
χ^2 分布的性质

χ^2 分布的概率密度图像如图所示



χ^2 分布的性质

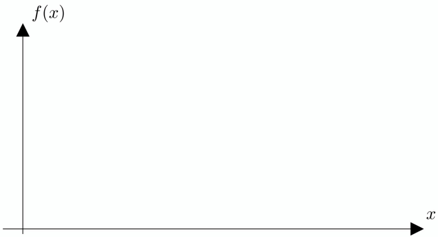
χ^2 分布的概率密度图像如图所示



χ^2 分布的性质

χ^2 分布的性质

χ^2 分布的概率密度图像如图所示

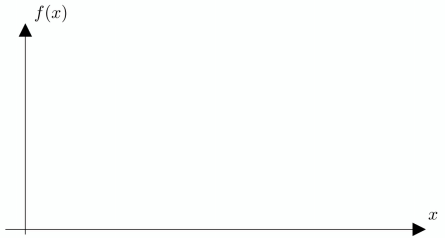


χ^2 分布的性质

(1) 设 $\chi^2 \sim \chi^2(n)$, 则 $E(\chi^2) = n, D(\chi^2) = 2n$;

χ^2 分布的性质

χ^2 分布的概率密度图像如图所示



χ^2 分布的性质

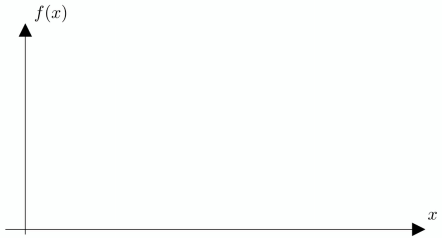
(1) 设 $\chi^2 \sim \chi^2(n)$, 则 $E(\chi^2) = n, D(\chi^2) = 2n$;

(2) 设 $Y_1 \sim \chi^2(n_1), Y_2 \sim \chi^2(n_2)$, 且 Y_1, Y_2 相互独立，则有

$$Y_1 + Y_2 \sim \chi^2(n_1 + n_2)$$

χ^2 分布的性质

χ^2 分布的概率密度图像如图所示



χ^2 分布的性质

(1) **设** $\chi^2 \sim \chi^2(n)$, **则** $E(\chi^2) = n, D(\chi^2) = 2n$;

(2) **设** $Y_1 \sim \chi^2(n_1), Y_2 \sim \chi^2(n_2)$, **且** Y_1, Y_2 **相互独立，则有**

$$Y_1 + Y_2 \sim \chi^2(n_1 + n_2)$$

(3) **设** $\chi^2 \sim \chi^2(n)$, **当** $n \to \infty, \chi^2$ **分布近似于** $N(n, 2n)$.

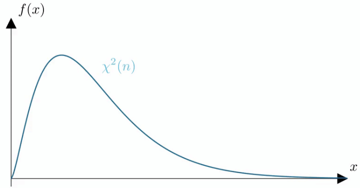
χ^2 分布的分位点

定义 5

设 $\chi^2 \sim \chi^2(n)$，对于给定的 $0 < \alpha < 1$，若实数 b 满足

$$P\left\{\chi^2 > b\right\} = \alpha$$

则称数 b 为 χ^2 分布的上 α 分位点，记作 $\chi^2_\alpha(n)$.



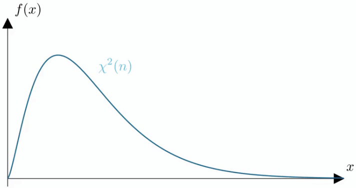
χ^2 分布的分位点

定义 5

设 $\chi^2 \sim \chi^2(n)$，对于给定的 $0 < \alpha < 1$，若实数 b 满足

$$P\left\{\chi^2 > b\right\} = \alpha$$

则称数 b 为 χ^2 分布的上 α 分位点，记作 $\chi^2_\alpha(n)$.



χ^2 分布的分位点

定义 5

设 $\chi^2 \sim \chi^2(n)$，对于给定的 $0 < \alpha < 1$，若实数 b 满足

$$P\left\{\chi^2 > b\right\} = \alpha$$

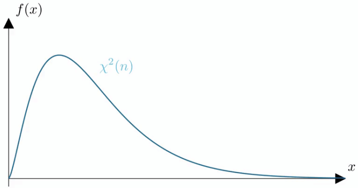
则称数 b 为 χ^2 分布的上 α 分位点，记作 $\chi^2_\alpha(n)$.

$\chi^2_\alpha(n)$ 可通过 χ^2 分布表查到，例如

$$\chi^2_{0.1}(25) = 34.382.$$

当 $n > 45$ 时，可利用近似公式

$$\chi^2_\alpha(n) \approx \frac{1}{2}\left(z_\alpha + \sqrt{2n-1}\right)^2$$



例 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 的一个容量为 n 的样本，求 $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$ 服从什么分布?

例 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 的一个容量为 n 的样本，求 $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$ 服从什么分布?

解：因为总体 $X\sim N(\mu,\sigma^2)$，所以 $X_i\sim N(\mu,\sigma^2)$，则

$$\frac{X_i-\mu}{\sigma}\sim N(0,1), i=1,2,\cdots,n$$

例 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 的一个容量为 n 的样本，求 $\frac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 服从什么分布?

解：因为总体 $X\sim N(\mu,\sigma^2)$，所以 $X_i\sim N(\mu,\sigma^2)$，则

$$\frac{X_i-\mu}{\sigma}\sim N(0,1),i=1,2,\cdots,n$$

又因为 $\frac{X_i-\mu}{\sigma},i=1,2,\cdots,n$ 相互独立，则

例 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 的一个容量为 n 的样本，求 $\frac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 服从什么分布?

解：因为总体 $X\sim N(\mu,\sigma^2)$，所以 $X_i\sim N(\mu,\sigma^2)$，则

$$\frac{X_i-\mu}{\sigma}\sim N(0,1),i=1,2,\cdots,n$$

又因为 $\frac{X_i-\mu}{\sigma},i=1,2,\cdots,n$ 相互独立，则

$$\left(\frac{X_1-\mu}{\sigma}\right)^2+\left(\frac{X_2-\mu}{\sigma}\right)^2+\cdots+\left(\frac{X_n-\mu}{\sigma}\right)^2=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

例 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 的一个容量为 n 的样本，求 $\frac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 服从什么分布?

解：因为总体 $X\sim N(\mu,\sigma^2)$，所以 $X_i\sim N(\mu,\sigma^2)$，则

$$\frac{X_i-\mu}{\sigma}\sim N(0,1), i=1,2,\cdots,n$$

又因为 $\frac{X_i-\mu}{\sigma}, i=1,2,\cdots,n$ 相互独立，则

$$\left(\frac{X_1-\mu}{\sigma}\right)^2+\left(\frac{X_2-\mu}{\sigma}\right)^2+\cdots+\left(\frac{X_n-\mu}{\sigma}\right)^2=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

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

例 2

设 $X_1, X_2, \cdots, X_{10}$ 是来自总体 $X \sim N(0, 0.3^2)$ 的一个样本，求 $P\left\{\sum\limits_{i=1}^{10} X_i^2 > 1.44\right\}$，已知 $\chi^2_{0.1}(10) = 15.987$.

例 2

设 $X_1, X_2, \cdots, X_{10}$ 是来自总体 $X \sim N(0, 0.3^2)$ 的一个样本，求 $P\left\{\sum_{i=1}^{10} X_i^2 > 1.44\right\}$，已知 $\chi^2_{0.1}(10) = 15.987.$

解：因为总体 $X \sim N(0, 0.3^2)$，所以 $X_i \sim N(0, 0.3^2)$，则

$$\frac{X_i}{0.3} \sim N(0,1), i = 1, 2, \cdots, 10$$

例 2

设 $X_1, X_2, \cdots, X_{10}$ 是来自总体 $X \sim N(0, 0.3^2)$ 的一个样本，求 $P\left\{\sum_{i=1}^{10} X_i^2 > 1.44\right\}$，已知 $\chi^2_{0.1}(10) = 15.987$.

解：因为总体 $X \sim N(0, 0.3^2)$，所以 $X_i \sim N(0, 0.3^2)$，则

$$\frac{X_i}{0.3} \sim N(0,1), i = 1, 2, \cdots, 10$$

$$\left(\frac{X_1}{0.3}\right)^2 + \left(\frac{X_2}{0.3}\right)^2 + \cdots + \left(\frac{X_{10} - \mu}{0.3}\right)^2 = \frac{1}{0.09}\sum_{i=1}^{10} X_i^2 \sim \chi^2(10)$$

例 2

设 $X_1, X_2, \cdots, X_{10}$ 是来自总体 $X \sim N(0, 0.3^2)$ 的一个样本，求 $P\left\{\sum_{i=1}^{10} X_i^2 > 1.44\right\}$，已知 $\chi^2_{0.1}(10) = 15.987$.

解：因为总体 $X \sim N(0, 0.3^2)$，所以 $X_i \sim N(0, 0.3^2)$，则

$$\frac{X_i}{0.3} \sim N(0,1), i = 1, 2, \cdots, 10$$

$$\left(\frac{X_1}{0.3}\right)^2 + \left(\frac{X_2}{0.3}\right)^2 + \cdots + \left(\frac{X_{10}-\mu}{0.3}\right)^2 = \frac{1}{0.09}\sum_{i=1}^{10} X_i^2 \sim \chi^2(10)$$

故 $P\left\{\sum_{i=1}^{10} X_i^2 > 1.44\right\} = P\left\{\frac{1}{0.09}\sum_{i=1}^{10} X_i^2 > \frac{1.44}{0.09}\right\} = 0.1$,

定义 6

设 $X \sim N(0,1), Y \sim \chi^2(n)$，且 X, Y 相互独立，则称统计量

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

服从自由度为 n 的 t 分布，简记为 $T \sim t(n)$ 分布.

t 分布

定义 6
设 $X \sim N(0,1), Y \sim \chi^2(n)$，且 X, Y 相互独立，则称统计量

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

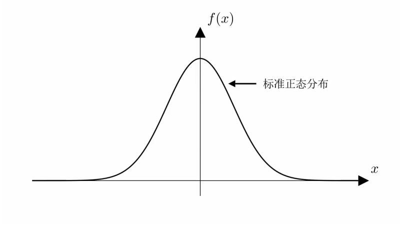
服从自由度为 n 的 t 分布，简记为 $T \sim t(n)$ 分布.

t 分布的概率密度函数

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

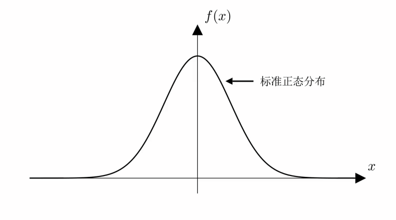
t 分布的性质

t 分布的概率密度图像如图所示



t 分布的性质

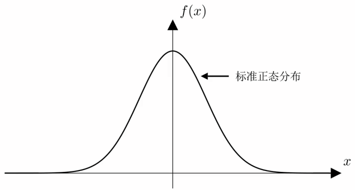
t 分布的概率密度图像如图所示



t 分布的性质

t 分布的性质

t 分布的概率密度图像如图所示



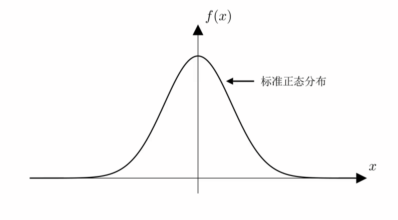
t 分布的性质

(1) t 分布的概率密度函数 $f(x)$ 是偶函数，即对于任意的 $x_0 > 0$，有

$$P\{T < -x_0\} = P\{T > x_0\}$$

t 分布的性质

t 分布的概率密度图像如图所示



t 分布的性质

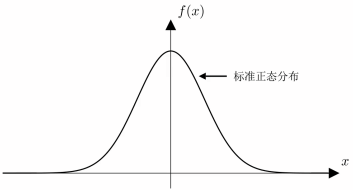
(1) t 分布的概率密度函数 $f(x)$ 是偶函数，即对于任意的 $x_0 > 0$，有

$$P\{T < -x_0\} = P\{T > x_0\}$$

(2) 若 $T \sim t(n), n > 1$，则 $E(T) = 0$;

t 分布的性质

t 分布的概率密度图像如图所示



t 分布的性质

(1) t **分布的概率密度函数** $f(x)$ **是偶函数，即对于任意的** $x_0 > 0$，**有**

$$P\{T < -x_0\} = P\{T > x_0\}$$

(2) **若** $T \sim t(n), n > 1$，**则** $E(T) = 0$;

(3) **当** n **足够大时，**t **分布近似于** $N(0,1)$, **即** $\lim_{n\to\infty} f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

t 分布的分位点

定义 7

设 $T \sim t(n)$，对于给定的 $0<\alpha<1$，若实数 b 满足

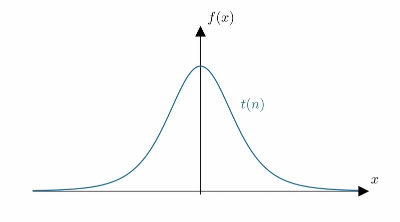
$$P\{T>b\}=\alpha$$

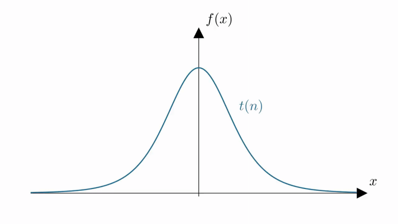
则称数 b 为 T 分布的上 α 分位点，记作 $t_{\alpha}(n)$. 若实数 c 满足

$$P\{|T|>c\}=\alpha$$

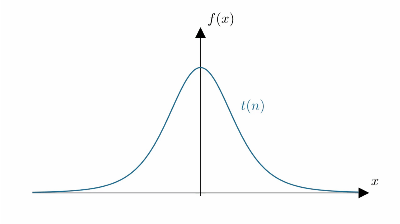
则称数 c 为 t 分布的双侧 α 分位点，记作 $t_{\frac{\alpha}{2}}(n)$.

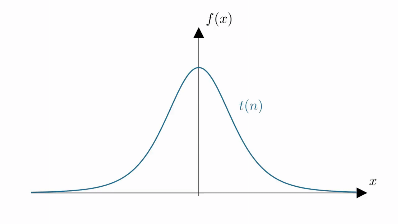
t 分布的上 α 分位点示意图





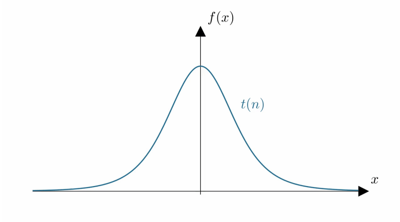
t 分布的上 α 分位点示意图

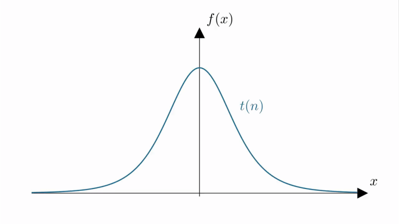




已知 n, α 通过查 t 分布表可求得 $t_\alpha(n)$. 当 $t > 0$ 时，可利用标准正态分布的上 α 分位点近似 t 分布的上 α 分位点，即 $t_\alpha \approx z_\alpha$.

t 分布的上 α 分位点示意图





已知 n, α 通过查 t 分布表可求得 $t_\alpha(n)$. 当 $t > 0$ 时，可利用标准正态分布的上 α 分位点近似 t 分布的上 α 分位点，即 $t_\alpha \approx z_\alpha$.

由 t 分布概率密度对称性可知

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

例 3

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，问 $Y = \dfrac{\sqrt{n-1}(X_1-\mu)}{\sqrt{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}}$ 服从什么分布?

例 3

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，问 $Y = \dfrac{\sqrt{n-1}(X_1-\mu)}{\sqrt{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}}$ 服从什么分布?

解：由 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的样本，则

$$\frac{X_1-\mu}{\sigma} \sim N(0,1), \left(\frac{X_2-\mu}{\sigma}\right)^2+\cdots+\left(\frac{X_n-\mu}{\sigma}\right)^2 \sim \chi^2(n-1)$$

例 3

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，问 $Y = \dfrac{\sqrt{n-1}(X_1-\mu)}{\sqrt{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}}$ 服从什么分布?

解：由 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的样本，则

$$\frac{X_1-\mu}{\sigma} \sim N(0,1), \left(\frac{X_2-\mu}{\sigma}\right)^2 + \cdots + \left(\frac{X_n-\mu}{\sigma}\right)^2 \sim \chi^2(n-1)$$

且 $\frac{X_1-\mu}{\sigma}$ 与 $\frac{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}{\sigma^2}$ 相互独立，因此

例 3

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，问 $Y = \dfrac{\sqrt{n-1}(X_1-\mu)}{\sqrt{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}}$ 服从什么分布?

解：由 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的样本，则

$$\frac{X_1-\mu}{\sigma} \sim N(0,1), \left(\frac{X_2-\mu}{\sigma}\right)^2+\cdots+\left(\frac{X_n-\mu}{\sigma}\right)^2 \sim \chi^2(n-1)$$

且 $\frac{X_1-\mu}{\sigma}$ 与 $\frac{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}{\sigma^2}$ 相互独立，因此

$$\frac{\frac{X_1-\mu}{\sigma}}{\sqrt{\frac{\frac{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}{\sigma^2}}{n-1}}} = \frac{\sqrt{n-1}(X_1-\mu)}{\sqrt{(X_2-\mu)^2+\cdots+(X_n-\mu)^2}} \sim t(n-1)$$

F 分布

定义 8

设 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，且 X, Y 相互独立，则称统计量

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

服从自由度为 n_1, n_2 的 F 分布，简记为 $F \sim F(n_1, n_2)$ 分布.

F 分布

定义 8
设 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，且 X, Y 相互独立，则称统计量

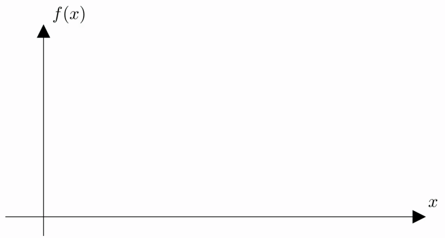
$$F = \frac{\frac{X}{n_1}}{\frac{Y}{n_2}}$$

服从自由度为 n_1, n_2 的 F 分布，简记为 $F \sim F(n_1, n_2)$ 分布.

F 分布的概率密度函数为

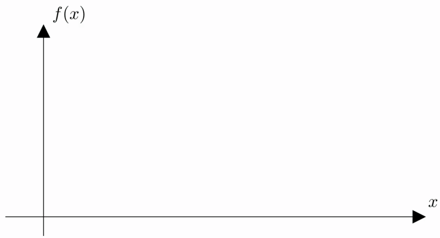
$$f(x) = \begin{cases} \dfrac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} x^{\frac{n_1}{2}-1}\left(1+\frac{n_1}{n_2}x\right)^{-\frac{n_1+n_2}{2}}, & x > 0, \\ 0, & \text{其他}. \end{cases}$$

F 分布的概率密度图像如图所示



F 分布的性质

F 分布的概率密度图像如图所示



F 分布的性质

F 分布的性质

F 分布的概率密度图像如图所示

F 分布的性质

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

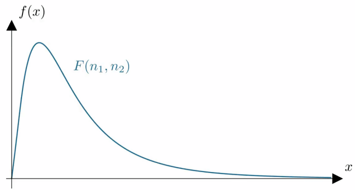
F 分布的分位点

定义 9

设 $F \sim F(n_1, n_2)$，对于给定的 $0 < \alpha < 1$，若实数 b 满足

$$P\{F > b\} = \alpha$$

则称数 b 为 F 分布的上 α 分位点，记作 $F_\alpha(n_1, n_2)$.



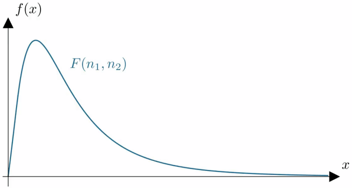
F 分布的分位点

定义 9

设 $F \sim F(n_1, n_2)$，对于给定的 $0 < \alpha < 1$，若实数 b 满足

$$P\{F > b\} = \alpha$$

则称数 b 为 F 分布的上 α 分位点，记作 $F_\alpha(n_1, n_2)$.



F 分布的分位点

定义 9
设 $F \sim F(n_1, n_2)$，对于给定的 $0 < \alpha < 1$，若实数 b 满足

$$P\{F > b\} = \alpha$$

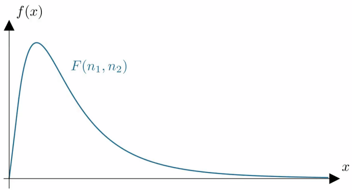
则称数 b 为 F 分布的上 α 分位点，记作 $F_\alpha(n_1, n_2)$.

$F_\alpha(n_1, n_2)$ 可通过 F 分布表查到，如

$$F_{0.05}(10, 5) = 4.74.$$

由 F 分布的性质可知

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



求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_{\alpha}(n_2, n_1)$

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_\alpha(n_2, n_1)$

证明：设 $F \sim F(n_1, n_2)$，则由性质可知 $\frac{1}{F} \sim F(n_2, n_1)$，由分位点定义有

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_\alpha(n_2,n_1)$

证明：设 $F \sim F(n_1,n_2)$，则由性质可知 $\frac{1}{F} \sim F(n_2,n_1)$，由分位点定义有

$$1-\alpha = P\{F > F_{1-\alpha}(n_1,n_2)\}$$

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_\alpha(n_2, n_1)$

证明：设 $F \sim F(n_1, n_2)$，则由性质可知 $\frac{1}{F} \sim F(n_2, n_1)$，由分位点定义有

$$\begin{aligned} 1-\alpha &= P\{F > F_{1-\alpha}(n_1, n_2)\} \\ &= P\left\{\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} \end{aligned}$$

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_\alpha(n_2, n_1)$

证明：设 $F \sim F(n_1, n_2)$，则由性质可知 $\frac{1}{F} \sim F(n_2, n_1)$，由分位点定义有

$$
\begin{aligned}
1-\alpha &= P\{F > F_{1-\alpha}(n_1, n_2)\} \\
&= P\left\{\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} \\
&= 1 - P\left\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\}
\end{aligned}
$$

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_\alpha(n_2, n_1)$

证明：设 $F \sim F(n_1, n_2)$，则由性质可知 $\frac{1}{F} \sim F(n_2, n_1)$，由分位点定义有

$$\begin{aligned}1-\alpha &= P\{F > F_{1-\alpha}(n_1, n_2)\} \\ &= P\left\{\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} \\ &= 1 - P\left\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\}\end{aligned}$$

因此 $P\left\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1,n_2)}\right\} = \alpha$，另一方面，由分位点的定义有

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)}=F_\alpha(n_2,n_1)$

证明：设 $F\sim F(n_1,n_2)$，则由性质可知 $\frac{1}{F}\sim F(n_2,n_1)$，由分位点定义有

$$\begin{aligned}1-\alpha&=P\{F>F_{1-\alpha}(n_1,n_2)\}\\&=P\left\{\frac{1}{F}<\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}\\&=1-P\left\{\frac{1}{F}>\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}\end{aligned}$$

因此 $P\left\{\frac{1}{F}>\frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}=\alpha$，另一方面，由分位点的定义有

$$P\left\{\frac{1}{F}>F_\alpha(n_2,n_1)\right\}=\alpha$$

求证 F 分布满足 $\frac{1}{F_{1-\alpha}(n_1,n_2)} = F_\alpha(n_2, n_1)$

证明：设 $F \sim F(n_1, n_2)$，则由性质可知 $\frac{1}{F} \sim F(n_2, n_1)$，由分位点定义有

$$\begin{aligned} 1-\alpha &= P\{F > F_{1-\alpha}(n_1, n_2)\} \\ &= P\left\{\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} \\ &= 1 - P\left\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1, n_2)}\right\} \end{aligned}$$

因此 $P\left\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1,n_2)}\right\} = \alpha$，另一方面，由分位点的定义有

$$P\left\{\frac{1}{F} > F_\alpha(n_2, n_1)\right\} = \alpha$$

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

例 4

已知 $X \sim t(n)$，证明 $\frac{1}{X^2} \sim F(n,1)$.

例 4

已知 $X \sim t(n)$，证明 $\frac{1}{X^2} \sim F(n,1)$.

证明：由 $X \sim t(n)$，则存在相互独立的随机变量 U, V，使得

$$X = \frac{U}{\sqrt{\frac{V}{n}}}$$

其中 $U \sim N(0,1), V \sim \chi^2(n)$，于是

例 4

已知 $X \sim t(n)$，证明 $\frac{1}{X^2} \sim F(n,1)$.

证明：由 $X \sim t(n)$，则存在相互独立的随机变量 U, V，使得

$$X = \frac{U}{\sqrt{\frac{V}{n}}}$$

其中 $U \sim N(0,1), V \sim \chi^2(n)$，于是

$$\frac{1}{X^2} = \frac{\frac{V}{n}}{\frac{U^2}{1}} \sim F(n,1)$$

例 5

设随机变量 $X \sim F(n,n)$，已知 $P\{X \geq \alpha\} = 0.05$，求 $P\left\{X > \frac{1}{\alpha}\right\}$.

例 5

设随机变量 $X \sim F(n,n)$，已知 $P\{X \geq \alpha\} = 0.05$，求 $P\left\{X > \frac{1}{\alpha}\right\}$.

解： 由 $X \sim F(n,n)$，则 $\frac{1}{X} \sim F(n,n)$，即 X 和 $\frac{1}{X}$ 同分布，则

$$P\{X \geq \alpha\} = P\left\{\frac{1}{X} \geq \alpha\right\} = 0.05$$

因此

例 5

设随机变量 $X \sim F(n,n)$**，已知** $P\{X \geq \alpha\} = 0.05$**，求** $P\left\{X > \frac{1}{\alpha}\right\}$.

解：由 $X \sim F(n,n)$**，则** $\frac{1}{X} \sim F(n,n)$**，即** X **和** $\frac{1}{X}$ **同分布，则**

$$P\{X \geq \alpha\} = P\left\{\frac{1}{X} \geq \alpha\right\} = 0.05$$

因此

$$P\left\{X > \frac{1}{\alpha}\right\} = P\left\{\frac{1}{X} < \alpha\right\} = 1 - P\left\{\frac{1}{X} \geq \alpha\right\} = 1 - 0.05 = 0.95$$

例 6

设 X_1, X_2 为总体 $X \sim N(0, \sigma^2)$，求证 $\dfrac{(X_1 + X_2)^2}{(X_1 - X_2)^2} \sim F(1, 1)$.

例 6

设 X_1, X_2 为总体 $X \sim N(0, \sigma^2)$，求证 $\dfrac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$.

证明：由于 X_1, X_2 独立同分布于 $N(0, \sigma^2)$，则

$$\frac{X_1+X_2}{\sqrt{2}\sigma} \sim N(0,1), \frac{X_1-X_2}{\sqrt{2}\sigma} \sim N(0,1)$$

例 6

设 X_1, X_2 为总体 $X \sim N(0, \sigma^2)$，求证 $\dfrac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$.

证明：由于 X_1, X_2 独立同分布于 $N(0, \sigma^2)$，则

$$\frac{X_1+X_2}{\sqrt{2}\sigma} \sim N(0,1), \frac{X_1-X_2}{\sqrt{2}\sigma} \sim N(0,1)$$

因为 X_1+X_2 与 X_1-X_2 的任意非零线性组合服从一维正态分布，所以 (X_1+X_2, X_1-X_2) 服从二维正态分布，又因为

例 6

设 X_1, X_2 为总体 $X \sim N(0,\sigma^2)$，求证 $\dfrac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$.

证明：由于 X_1, X_2 **独立同分布于** $N(0,\sigma^2)$**，则**

$$\frac{X_1+X_2}{\sqrt{2}\sigma} \sim N(0,1), \frac{X_1-X_2}{\sqrt{2}\sigma} \sim N(0,1)$$

因为 X_1+X_2 **与** X_1-X_2 **的任意非零线性组合服从一维正态分布，所以** (X_1+X_2, X_1-X_2) **服从二维正态分布，又因为**

$$\begin{aligned}\operatorname{Cov}(X_1+X_2, X_1-X_2) &= \operatorname{Cov}(X_1,X_1) - \operatorname{Cov}(X_1,X_2) \\ &\quad + \operatorname{Cov}(X_2,X_1) - \operatorname{Cov}(X_2,X_2)\end{aligned}$$

例 6

设 X_1, X_2 为总体 $X \sim N(0, \sigma^2)$，求证 $\dfrac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$.

证明：由于 X_1, X_2 独立同分布于 $N(0, \sigma^2)$，则

$$\frac{X_1+X_2}{\sqrt{2}\sigma} \sim N(0,1), \frac{X_1-X_2}{\sqrt{2}\sigma} \sim N(0,1)$$

因为 X_1+X_2 与 X_1-X_2 的任意非零线性组合服从一维正态分布，所以 (X_1+X_2, X_1-X_2) 服从二维正态分布，又因为

$$\begin{aligned}\mathrm{Cov}(X_1+X_2, X_1-X_2) &= \mathrm{Cov}(X_1, X_1) - \mathrm{Cov}(X_1, X_2) \\ &\quad + \mathrm{Cov}(X_2, X_1) - \mathrm{Cov}(X_2, X_2) \\ &= \mathrm{Cov}(X_1, X_1) - \mathrm{Cov}(X_2, X_2) = 0\end{aligned}$$

所以 X_1+X_2, X_1-X_2 相互独立，从而 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2, \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2$ 相互独立.

所以 X_1+X_2, X_1-X_2 相互独立，从而 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2, \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2$ 相互独立.

又因为 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1), \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1)$，则

所以 X_1+X_2, X_1-X_2 相互独立，从而 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2, \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2$ 相互独立.

又因为 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1), \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1)$，则

$$\frac{\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2}{\left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2} \sim F(1,1)$$

所以 X_1+X_2, X_1-X_2 相互独立，从而 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2, \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2$ 相互独立.

又因为 $\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1), \left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1)$，则

$$\frac{\left(\frac{X_1+X_2}{\sqrt{2}\sigma}\right)^2}{\left(\frac{X_1-X_2}{\sqrt{2}\sigma}\right)^2} \sim F(1,1)$$

即

$$\frac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$$

正态总体样本均值与样本方差的分布

定理 10

设总体 X 服从正态分布 $X \sim N(\mu, \sigma^2)$，$X_1, X_2, \cdots, X_n$ 是来自总体 X 的一个样本，$\overline{X}$ 是样本均值，S^2 是样本方差，则

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

(2) $\overline{X}$ **与** S^2 **相互独立;**

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

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

证明：仅证明 (1)(4)，(2)(3) 证明略

证明：仅证明 (1)(4)，(2)(3) 证明略

(1) $\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$，由于 $X_i\sim N(\mu,\sigma^2)(i=1,2,\cdots,n)$，且相互独立，故 $\overline{X}$ 服从正态分布.

证明：仅证明 (1)(4)，(2)(3) 证明略

(1) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，由于 $X_i\sim N(\mu,\sigma^2)(i=1,2,\cdots,n)$，且相互独立，故 $\overline{X}$ 服从正态分布.

由于 $E(\overline{X})=\mu, D(\overline{X})=\frac{\sigma^2}{n}$，所以

证明：仅证明 (1)(4)，(2)(3) 证明略

(1) $\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$，由于 $X_i\sim N(\mu,\sigma^2)(i=1,2,\cdots,n)$，且相互独立，故 $\overline{X}$ 服从正态分布.

由于 $E(\overline{X})=\mu, D(\overline{X})=\frac{\sigma^2}{n}$，所以

$$\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

证明：仅证明 (1)(4)，(2)(3) 证明略

(1) $\overline{X}=\frac{1}{n}\sum\limits_{i=1}^{n}X_i$，由于 $X_i\sim N(\mu,\sigma^2)(i=1,2,\cdots,n)$，且相互独立，故 $\overline{X}$ 服从正态分布.

由于 $E(\overline{X})=\mu, D(\overline{X})=\frac{\sigma^2}{n}$，所以

$$\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

标准化后得

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

(4) 因为 $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1)$, $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，并且 $\overline{X}$ 与 S^2 相互独立，

所以有

(4) 因为 $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1), \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，并且 $\overline{X}$ 与 S^2 相互独立，

所以有

$$\frac{\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}}{\sqrt{\dfrac{\dfrac{(n-1)S^2}{\sigma^2}}{n-1}}} \sim t(n-1)$$

(4) 因为 $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1), \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，并且 $\overline{X}$ 与 S^2 相互独立，

所以有

$$\frac{\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}}{\sqrt{\dfrac{\dfrac{(n-1)S^2}{\sigma^2}}{n-1}}} \sim t(n-1)$$

即

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

注： 推断过程中，统计量 $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1)$ 与 $\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$ 的区别

(1) 总体方差已知时，用 $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$;

(2) 总体方差未知时，用 $\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}}$;

(3) 当样本容量 n 较大时，差别不大.

注： 推断过程中，统计量 $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \sim \chi^2(n-1)$ 与 $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$ 的区别

(1) 总体均值已知时，用 $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$；

(2) 总体均值未知时，用 $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$；

(3) 当样本容量 n 较大时，差别不大.

例 7

在总体 $N(30,4)$ 中随机抽取一个容量为 16 的简单随机样本，求样本均值落在区间 $(29,31)$ 之间的概率.

例 7

在总体 $N(30,4)$ 中随机抽取一个容量为 16 的简单随机样本，求样本均值落在区间 $(29,31)$ 之间的概率.

解：由定理可知 $\dfrac{\overline{X}-30}{\dfrac{2}{\sqrt{16}}} \sim N(0,1)$，**化简得** $\dfrac{\overline{X}-30}{0.5} \sim N(0,1)$，**则**

例 7

在总体 $N(30,4)$ 中随机抽取一个容量为 16 的简单随机样本，求样本均值落在区间 $(29,31)$ 之间的概率.

解：由定理可知 $\dfrac{\overline{X}-30}{\dfrac{2}{\sqrt{16}}} \sim N(0,1)$，**化简得** $\dfrac{\overline{X}-30}{0.5} \sim N(0,1)$，**则**

$$
\begin{aligned}
P\{29<\overline{X}<31\} &= P\left\{\frac{29-30}{0.5}<\frac{\overline{X}-30}{0.5}<\frac{31-30}{0.5}\right\} \\
&= \Phi(2)-\Phi(-2) \\
&= 2\Phi(2)-1 = 0.9544
\end{aligned}
$$

例 8

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma^2)$ 的样本，其样本均值为 $\overline{X} = \dfrac{1}{n_1}\sum\limits_{i=1}^{n_1} X_i$，设 $Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma^2)$ 的样本，其样本均值为 $\overline{Y} = \dfrac{1}{n_2}\sum\limits_{i=1}^{n_2} Y_i$，求

(1) $E\left[\sum\limits_{i=1}^{n_1}\left(X_i - \overline{X}\right)^2 + \sum\limits_{i=1}^{n_2}\left(Y_i - \overline{Y}\right)^2\right]$

(2) $D\left[\sum\limits_{i=1}^{n_1}\left(X_i - \overline{X}\right)^2 + \sum\limits_{i=1}^{n_2}\left(Y_i - \overline{Y}\right)^2\right]$

例 8

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma^2)$ 的样本，其样本均值为 $\overline{X} = \dfrac{1}{n_1}\sum\limits_{i=1}^{n_1} X_i$，设 $Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma^2)$ 的样本，其样本均值为 $\overline{Y} = \dfrac{1}{n_2}\sum\limits_{i=1}^{n_2} Y_i$，求

(1) $E\left[\sum\limits_{i=1}^{n_1}\left(X_i - \overline{X}\right)^2 + \sum\limits_{i=1}^{n_2}\left(Y_i - \overline{Y}\right)^2\right]$

(2) $D\left[\sum\limits_{i=1}^{n_1}\left(X_i - \overline{X}\right)^2 + \sum\limits_{i=1}^{n_2}\left(Y_i - \overline{Y}\right)^2\right]$

解：(1) 由 $E(S_1^2) = E(S_2^2) = \sigma^2$，且

$$(n_1 - 1)S_1^2 = \sum_{i=1}^{n_1}\left(X_i - \overline{X}\right)^2, (n_2 - 1)S_2^2 = \sum_{i=1}^{n_2}\left(Y_i - \overline{Y}\right)^2$$

$$
\begin{aligned}
E\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right] &= E\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right] \\
&= (n_1-1)E\left(S_1^2\right)+(n_2-1)E(S_2^2) \\
&= (n_1+n_2-2)\sigma^2
\end{aligned}
$$

$$
\begin{aligned}
E\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right] &= E\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right] \\
&= (n_1-1)E\left(S_1^2\right)+(n_2-1)E(S_2^2) \\
&= (n_1+n_2-2)\sigma^2
\end{aligned}
$$

(2)　由定理可知 $\frac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1)$**，且两个样本独立，所以有**

$$E\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right]=E\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right]$$
$$=(n_1-1)E\left(S_1^2\right)+(n_2-1)E(S_2^2)$$
$$=(n_1+n_2-2)\sigma^2$$

(2) 由定理可知 $\frac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1)$**，且两个样本独立，所以有**

$$D\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right]=D\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right]$$

$$E\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right]=E\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right]$$
$$=(n_1-1)E\left(S_1^2\right)+(n_2-1)E(S_2^2)$$
$$=(n_1+n_2-2)\sigma^2$$

(2) 由定理可知 $\frac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1)$，且两个样本独立，所以有

$$D\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right]=D\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right]$$
$$=\sigma^4D\left[\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\right]$$

$$E\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right]=E\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right]$$
$$=(n_1-1)E\left(S_1^2\right)+(n_2-1)E(S_2^2)$$
$$=(n_1+n_2-2)\sigma^2$$

(2) 由定理可知 $\frac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1)$**，且两个样本独立，所以有**

$$D\left[\sum_{i=1}^{n_1}\left(X_i-\overline{X}\right)^2+\sum_{i=1}^{n_2}\left(Y_i-\overline{Y}\right)^2\right]=D\left[(n_1-1)S_1^2+(n_2-1)S_2^2\right]$$
$$=\sigma^4 D\left[\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\right]$$
$$=2(n_1+n_2-2)\sigma^4$$

定理 11

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本，$\overline{X}$ 为其样本均值，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的样本，$\overline{Y}$ 为其样本均值，且两个样本相互独立，则有

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

定理 11

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本，$\overline{X}$ 为其样本均值，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的样本，$\overline{Y}$ 为其样本均值，且两个样本相互独立，则有

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

证明：因为 $\overline{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right), \overline{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$，且样本相互独立，所以 $\overline{X}$ 与 $\overline{Y}$ 独立，故

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

定理 12

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma^2)$ 的样本，$\overline{X}, S_1^2$ 为其样本均值和样本方差，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma^2)$ 的样本，$\overline{Y}, S_2^2$ 为其样本均值和样本方差，且两个样本相互独立，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 $S_W = \sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$.

定理 12

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma^2)$ 的样本，$\overline{X}, S_1^2$ 为其样本均值和样本方差，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma^2)$ 的样本，$\overline{Y}, S_2^2$ 为其样本均值和样本方差，且两个样本相互独立，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 $S_W = \sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$.

证明：因为 $\dfrac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi^2(n_1 - 1), \dfrac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_2 - 1)$，且独立

从而有

$$\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_1+n_2-2)$$

从而有

$$\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_1+n_2-2)$$

又因为

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}}\sim N(0,1)$$

从而有

$$\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\sim \chi^2(n_1+n_2-2)$$

又因为

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}}\sim N(0,1)$$

显然 $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}}$ **与** $\dfrac{(n_1-1)S_1^2}{\sigma^2}+\dfrac{(n_2-1)S_2^2}{\sigma^2}$ **相互独立**

由 t 分布的定义可得

$$\frac{\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1-1)S_1^2}{\sigma^2}+\dfrac{(n_2-1)S_2^2}{\sigma^2}}{n_1+n_2-2}}} \sim t(n_1+n_2-2)$$

由 t 分布的定义可得

$$\frac{\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1-1)S_1^2}{\sigma^2}+\dfrac{(n_2-1)S_2^2}{\sigma^2}}{n_1+n_2-2}}} \sim t(n_1+n_2-2)$$

整理可得

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

定理 13

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本，S_1^2 为其样本方差，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的样本，S_2^2 为其样本方差，且两个样本相互独立，则有

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

定理 13

设 $X_1, X_2, \cdots, X_{n_1}$ 是来自总体 $X \sim N(\mu_1, \sigma_1^2)$ 的样本，S_1^2 为其样本方差，$Y_1, Y_2, \cdots, Y_{n_2}$ 是来自总体 $Y \sim N(\mu_2, \sigma_2^2)$ 的样本，S_2^2 为其样本方差，且两个样本相互独立，则有

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

证明：因为 $\dfrac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1), \dfrac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1)$，且相互独立

由 F 分布的定义可得

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}}{n_1-1} \Big/ \frac{\frac{(n_2-1)S_2^2}{\sigma_2^2}}{n_2-1} \sim F(n_1-1, n_2-1)$$

由 F 分布的定义可得

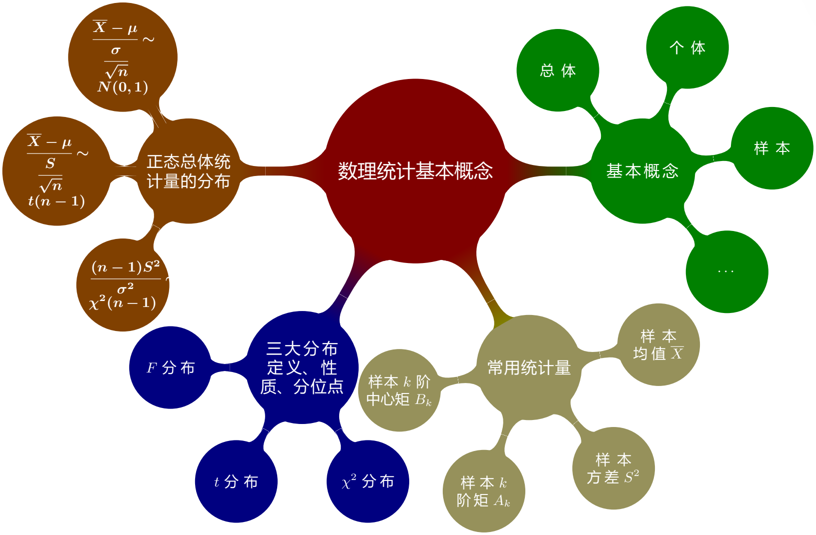
$$\frac{\dfrac{\dfrac{(n_1-1)S_1^2}{\sigma_1^2}}{n_1-1}}{\dfrac{\dfrac{(n_2-1)S_2^2}{\sigma_2^2}}{n_2-1}} \sim F(n_1-1, n_2-1)$$

整理可得

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

数理统计的基本概念习题

知识点总结



数理统计基本概念
基本概念
总 体
个 体
样 本
...
正态总体统计量的分布
$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$
$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
三大分布定义、性质、分位点
F 分布
t 分布
χ^2 分布
常用统计量
样本均值 $\overline{X}$
样本方差 S^2
样本 k 阶矩 A_k
样本 k 阶中心矩 B_k

例 1

设总体 $X \sim N(0,4)$ 与总体 $Y \sim N(0,9)$ 相互独立，$\overline{X} = \dfrac{1}{10}\sum_{i=1}^{10} X_i$ 与 $\overline{Y} = \dfrac{1}{15}\sum_{i=1}^{15} Y_i$ 分别是来自总体 X 与 Y 的样本均值，则 $\overline{X} - \overline{Y}$ 服从 __________

例 1

设总体 $X \sim N(0,4)$ **与总体** $Y \sim N(0,9)$ **相互独立，** $\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$ **与** $\overline{Y} = \frac{1}{15}\sum_{i=1}^{15} Y_i$ **分别是来自总体** X **与** Y **的样本均值，则** $\overline{X} - \overline{Y}$ **服从** ____________

解：由 $X \sim N(0,4), Y \sim N(0,9)$**，可知** $\overline{X} \sim N\left(0, \frac{4}{10}\right), \overline{Y} \sim N\left(0, \frac{9}{15}\right)$.

例 1

设总体 $X \sim N(0,4)$ 与总体 $Y \sim N(0,9)$ 相互独立，$\overline{X} = \dfrac{1}{10}\sum\limits_{i=1}^{10} X_i$ 与 $\overline{Y} = \dfrac{1}{15}\sum\limits_{i=1}^{15} Y_i$ 分别是来自总体 X 与 Y 的样本均值，则 $\overline{X} - \overline{Y}$ 服从 __________

解：由 $X \sim N(0,4), Y \sim N(0,9)$，可知 $\overline{X} \sim N\left(0, \frac{4}{10}\right), \overline{Y} \sim N\left(0, \frac{9}{15}\right)$.
因为总体 X 与总体 Y 相互独立，所以 $\overline{X}$ 与 $\overline{Y}$ 相互独立，因此

例 1

设总体 $X \sim N(0,4)$ 与总体 $Y \sim N(0,9)$ 相互独立，$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$ 与 $\overline{Y} = \frac{1}{15}\sum_{i=1}^{15} Y_i$ 分别是来自总体 X 与 Y 的样本均值，则 $\overline{X} - \overline{Y}$ 服从 ________

解： 由 $X \sim N(0,4), Y \sim N(0,9)$，可知 $\overline{X} \sim N\left(0, \frac{4}{10}\right), \overline{Y} \sim N\left(0, \frac{9}{15}\right)$.
因为总体 X 与总体 Y 相互独立，所以 $\overline{X}$ 与 $\overline{Y}$ 相互独立，因此

$$\overline{X} - \overline{Y} \sim N\left(0, \frac{4}{10} + \frac{9}{19}\right)$$

即 $\overline{X} - \overline{Y} \sim N(0,1)$.

例 1

设总体 $X \sim N(0,4)$ **与总体** $Y \sim N(0,9)$ **相互独立，** $\overline{X} = \dfrac{1}{10}\sum_{i=1}^{10} X_i$ **与** $\overline{Y} = \dfrac{1}{15}\sum_{i=1}^{15} Y_i$ **分别是来自总体** X **与** Y **的样本均值，则** $\overline{X} - \overline{Y}$ **服从** $\underline{\quad N(0,1) \quad}$

解：由 $X \sim N(0,4), Y \sim N(0,9)$**，可知** $\overline{X} \sim N\left(0, \frac{4}{10}\right), \overline{Y} \sim N\left(0, \frac{9}{15}\right)$.
因为总体 X **与总体** Y **相互独立，所以** $\overline{X}$ **与** $\overline{Y}$ **相互独立，因此**

$$\overline{X} - \overline{Y} \sim N\left(0, \frac{4}{10} + \frac{9}{19}\right)$$

即 $\overline{X} - \overline{Y} \sim N(0,1)$.

例 2

设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, \sigma^2)$ 的样本，$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$，则当 $a =$ ______，$b =$ ______ 时，统计量 X 服从 χ^2 分布，其自由度为 ______

例 2

设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, \sigma^2)$ 的样本，$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$，则当 $a =$ ______，$b =$ ______ 时，统计量 X 服从 χ^2 分布，其自由度为 ______

解：因为 $X_1 - 2X_2 \sim N(0, 5\sigma^2), 3X_3 - 4X_4 \sim N(0, 25\sigma^2)$，并且相互独立，由 χ^2 分布定义有

例 2

设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0,\sigma^2)$ 的样本，$X = a(X_1-2X_2)^2 + b(3X_3-4X_4)^2$，则当 $a =$ ________，$b =$ ________ 时，统计量 X 服从 χ^2 分布，其自由度为 ______

解：因为 $X_1-2X_2 \sim N(0,5\sigma^2), 3X_3-4X_4 \sim N(0,25\sigma^2)$，并且相互独立，由 χ^2 分布定义有

$$\left(\frac{X_1-2X_2}{\sqrt{5}\sigma}\right)^2 + \left(\frac{3X_3-4X_4}{5\sigma}\right)^2 \sim \chi^2(2)$$

例 2

设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0,\sigma^2)$ 的样本，$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$，则当 $a =$ ______，$b =$ ______ 时，统计量 X 服从 χ^2 分布，其自由度为 ______

解：因为 $X_1 - 2X_2 \sim N(0, 5\sigma^2), 3X_3 - 4X_4 \sim N(0, 25\sigma^2)$，并且相互独立，由 χ^2 分布定义有

$$\left(\frac{X_1 - 2X_2}{\sqrt{5}\sigma}\right)^2 + \left(\frac{3X_3 - 4X_4}{5\sigma}\right)^2 \sim \chi^2(2)$$

整理得

$$\frac{1}{5\sigma^2}(X_1 - 2X_2)^2 + \frac{1}{25\sigma^2}(3X_3 - 4X_4)^2 \sim \chi^2(2)$$

例 2

设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0,\sigma^2)$ 的样本，$X = a(X_1-2X_2)^2 + b(3X_3-4X_4)^2$，则当 $a = \underline{\frac{1}{5\sigma^2}}$，$b = \underline{\frac{1}{25\sigma^2}}$ 时，统计量 X 服从 χ^2 分布，其自由度为 $\underline{2}$

解：因为 $X_1 - 2X_2 \sim N(0, 5\sigma^2), 3X_3 - 4X_4 \sim N(0, 25\sigma^2)$，并且相互独立，由 χ^2 分布定义有

$$\left(\frac{X_1-2X_2}{\sqrt{5}\sigma}\right)^2 + \left(\frac{3X_3-4X_4}{5\sigma}\right)^2 \sim \chi^2(2)$$

整理得

$$\frac{1}{5\sigma^2}(X_1-2X_2)^2 + \frac{1}{25\sigma^2}(3X_3-4X_4)^2 \sim \chi^2(2)$$

例 3

设 X_1, X_2, X_3 是来自正态总体 $N(0, \sigma^2)$ 的一个样本，则统计量 $\dfrac{X_1 - X_2}{\sqrt{2}\,|X_3|}$ 服从的分布为 ______.

(A) $F(1,1)$;　　(B) $F(2,1)$;　　(C) $t(1)$;　　(D) $t(2)$.

例 3

设 X_1, X_2, X_3 是来自正态总体 $N(0,\sigma^2)$ 的一个样本，则统计量 $\dfrac{X_1 - X_2}{\sqrt{2}\,|X_3|}$ 服从的分布为 ______.

(A) $F(1,1)$； (B) $F(2,1)$； (C) $t(1)$； (D) $t(2)$.

解：由 $X_1 - X_2 \sim N(0, 2\sigma^2)$ 可知 $\dfrac{X_1 - X_2}{\sqrt{2\sigma^2}}$.

例 3

设 X_1, X_2, X_3 是来自正态总体 $N(0,\sigma^2)$ 的一个样本，则统计量 $\dfrac{X_1 - X_2}{\sqrt{2}\,|X_3|}$ 服从的分布为______.

(A) $F(1,1)$; (B) $F(2,1)$; (C) $t(1)$; (D) $t(2)$.

解：由 $X_1 - X_2 \sim N(0, 2\sigma^2)$ 可知 $\dfrac{X_1 - X_2}{\sqrt{2\sigma^2}}$.

由 $X_3 \sim N(0,\sigma^2)$ 可知 $\dfrac{X_3}{\sigma} \sim N(0,1)$，进而 $\left(\dfrac{X_3}{\sigma}\right)^2 \sim \chi^2(1)$,

例 3

设 X_1, X_2, X_3 是来自正态总体 $N(0,\sigma^2)$ 的一个样本，则统计量 $\dfrac{X_1 - X_2}{\sqrt{2}\,|X_3|}$ 服从的分布为 ______.

(A) $F(1,1)$; (B) $F(2,1)$; (C) $t(1)$; (D) $t(2)$.

解：由 $X_1 - X_2 \sim N(0, 2\sigma^2)$ 可知 $\dfrac{X_1 - X_2}{\sqrt{2\sigma^2}}$.

由 $X_3 \sim N(0,\sigma^2)$ 可知 $\dfrac{X_3}{\sigma} \sim N(0,1)$，进而 $\left(\dfrac{X_3}{\sigma}\right)^2 \sim \chi^2(1)$，并且

X_1, X_2, X_3 相互独立，由 t 分布的定义有 $\dfrac{\frac{X_1-X_2}{\sqrt{2\sigma^2}}}{\sqrt{\frac{\left(\frac{X_3}{\sigma}\right)^2}{1}}} = \dfrac{X_1-X_2}{\sqrt{2}|X_3|} \sim t(1)$

例 3

设 X_1, X_2, X_3 是来自正态总体 $N(0,\sigma^2)$ 的一个样本，则统计量 $\dfrac{X_1 - X_2}{\sqrt{2}\,|X_3|}$ 服从的分布为 $\underline{\quad C \quad}$.

(A) $F(1,1)$； (B) $F(2,1)$； (C) $t(1)$； (D) $t(2)$.

解：由 $X_1 - X_2 \sim N(0, 2\sigma^2)$ 可知 $\dfrac{X_1 - X_2}{\sqrt{2\sigma^2}}$.

由 $X_3 \sim N(0,\sigma^2)$ 可知 $\dfrac{X_3}{\sigma} \sim N(0,1)$，进而 $\left(\dfrac{X_3}{\sigma}\right)^2 \sim \chi^2(1)$，并且

X_1, X_2, X_3 相互独立，由 t 分布的定义有 $\dfrac{\frac{X_1 - X_2}{\sqrt{2\sigma^2}}}{\sqrt{\frac{\left(\frac{X_3}{\sigma}\right)^2}{1}}} = \frac{X_1 - X_2}{\sqrt{2}|X_3|} \sim t(1)$

例 4

设 $X_1, X_2, \cdots, X_n (n \geq 2)$ 是来自正态总体 $N(\mu, 1)$ 的一个样本，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，则下列结论不正确的是______.

(A) $\sum_{i=1}^{n} (X_i - \mu)^2$ 服从 χ^2 分布；　(B) $2(X_n - X_1)^2$ 服从 χ^2 分布；

(C) $\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ 服从 χ^2 分布；　(D) $n\left(\overline{X} - \mu\right)^2$ 服从 χ^2 分布.

例 4

设 $X_1, X_2, \cdots, X_n (n \geq 2)$ 是来自正态总体 $N(\mu, 1)$ 的一个样本，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，则下列结论不正确的是 ______.

(A) $\sum_{i=1}^{n} (X_i - \mu)^2$ 服从 χ^2 分布；　(B) $2(X_n - X_1)^2$ 服从 χ^2 分布；

(C) $\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ 服从 χ^2 分布；　(D) $n\left(\overline{X} - \mu\right)^2$ 服从 χ^2 分布.

解：(A) $\frac{X_i - \mu}{1} \sim N(0, 1)$，故 $\sum_{i=1}^{n} (X_i - \mu)^2 \sim \chi^2(n)$

例 4

设 $X_1, X_2, \cdots, X_n (n \geq 2)$ 是来自正态总体 $N(\mu, 1)$ 的一个样本，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，则下列结论不正确的是 ______.

(A) $\sum_{i=1}^{n}(X_i - \mu)^2$ 服从 χ^2 分布；　(B) $2(X_n - X_1)^2$ 服从 χ^2 分布；

(C) $\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ 服从 χ^2 分布；　(D) $n\left(\overline{X} - \mu\right)^2$ 服从 χ^2 分布.

(B)　因为 $X_n - X_1 \sim N(0, 2)$，所以 $\frac{X_n - X_1}{\sqrt{2}} \sim N(0, 1)$，故

$$\left(\frac{X_n - X_1}{\sqrt{2}}\right)^2 = \frac{(X_n - X_1)^2}{2} \sim \chi^2(1)$$

例 4

设 $X_1, X_2, \cdots, X_n (n \geq 2)$ 是来自正态总体 $N(\mu, 1)$ 的一个样本，记 $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$，则下列结论不正确的是 ______.

(A) $\sum\limits_{i=1}^{n} (X_i - \mu)^2$ 服从 χ^2 分布； (B) $2\left(X_n - X_1\right)^2$ 服从 χ^2 分布；

(C) $\sum\limits_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ 服从 χ^2 分布； (D) $n\left(\overline{X} - \mu\right)^2$ 服从 χ^2 分布.

(C) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，当 $\sigma^2 = 1$ 时，代入可得

$$\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \sim \chi^2(n-1)$$

例 4

设 $X_1, X_2, \cdots, X_n (n \geq 2)$ 是来自正态总体 $N(\mu, 1)$ 的一个样本，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，则下列结论不正确的是 ______.

(A) $\sum_{i=1}^{n}(X_i - \mu)^2$ 服从 χ^2 分布；　(B) $2(X_n - X_1)^2$ 服从 χ^2 分布；

(C) $\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ 服从 χ^2 分布；　(D) $n\left(\overline{X} - \mu\right)^2$ 服从 χ^2 分布.

(D)　因为 $\overline{X} \sim N\left(\mu, \frac{1}{n}\right)$，从而 $\frac{\overline{X}-\mu}{\sqrt{\frac{1}{n}}} \sim N(0,1)$，因此

$$\left(\frac{\overline{X} - \mu}{\sqrt{\frac{1}{n}}}\right)^2 = n\left(\overline{X} - \mu\right)^2 \sim \chi^2(1)$$

例 4

设 $X_1, X_2, \cdots, X_n (n \geq 2)$ 是来自正态总体 $N(\mu, 1)$ 的一个样本，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，则下列结论不正确的是 $\underline{\quad B \quad}$.

(A) $\sum_{i=1}^{n} (X_i - \mu)^2$ 服从 χ^2 分布； (B) $2\left(X_n - X_1\right)^2$ 服从 χ^2 分布；

(C) $\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ 服从 χ^2 分布； (D) $n\left(\overline{X} - \mu\right)^2$ 服从 χ^2 分布.

(D) 因为 $\overline{X} \sim N\left(\mu, \frac{1}{n}\right)$，从而 $\frac{\overline{X}-\mu}{\sqrt{\frac{1}{n}}} \sim N(0,1)$，因此

$$\left(\frac{\overline{X} - \mu}{\sqrt{\frac{1}{n}}}\right)^2 = n\left(\overline{X} - \mu\right)^2 \sim \chi^2(1)$$

例 5

设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，证明统计量

$$T = \frac{\sqrt{n(n-1)}(\overline{X}-\mu)}{\sqrt{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2}}$$

服从自由度为 $n-1$ 的 t 分布.

例 5

设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，证明统计量 $T = \dfrac{\sqrt{n(n-1)}(\overline{X}-\mu)}{\sqrt{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}}$ 服从自由度为 $n-1$ 的 t 分布.

证明：因为 $\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$，将 $S^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ 代入可得

例 5

设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，证明统计量 $T = \dfrac{\sqrt{n(n-1)}(\overline{X}-\mu)}{\sqrt{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}}$ 服从自由度为 $n-1$ 的 t 分布.

证明：因为 $\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$，将 $S^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$ 代入可得

$$\frac{\overline{X}-\mu}{\sqrt{\dfrac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{n(n-1)}}} = \frac{\sqrt{n(n-1)}(\overline{X}-\mu)}{\sqrt{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}} \sim t(n-1)$$

例 6

设 $X_1, X_2, \cdots, X_9$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，且 $Y_1 = \frac{1}{6}\sum_{i=1}^{6} X_i, Y_2 = \frac{1}{3}\sum_{i=7}^{9} X_i, S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2, Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S}$,

证明统计量 Z 服从自由度为 2 的 t 分布.

例 6

设 $X_1, X_2, \cdots, X_9$ **是来自正态总体** $N(\mu, \sigma^2)$ **的一个样本，且** $Y_1 = \frac{1}{6}\sum_{i=1}^{6} X_i, Y_2 = \frac{1}{3}\sum_{i=7}^{9} X_i, S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2, Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S}$，**证明统计量** Z **服从自由度为 2 的** t **分布.**

证明：由题意可知 Y_1 **与** Y_2 **相互独立，且有**

$$Y_1 \sim N\left(\mu, \frac{\sigma^2}{6}\right), Y_2 \sim N\left(\mu, \frac{\sigma^2}{3}\right)$$

例 6

设 $X_1, X_2, \cdots, X_9$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，且 $Y_1 = \frac{1}{6}\sum_{i=1}^{6} X_i, Y_2 = \frac{1}{3}\sum_{i=7}^{9} X_i, S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2, Z = \frac{\sqrt{2}(Y_1 - Y_2)}{S}$，证明统计量 Z 服从自由度为 2 的 t 分布.

证明：由题意可知 Y_1 与 Y_2 相互独立，且有

$$Y_1 \sim N\left(\mu, \frac{\sigma^2}{6}\right), Y_2 \sim N\left(\mu, \frac{\sigma^2}{3}\right)$$

所以 $Y_1 - Y_2 \sim N\left(0, \frac{\sigma^2}{2}\right)$，标准化得

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

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

因为 $S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2$，所以有

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

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

因为 $S^2 = \frac{1}{2}\sum_{i=7}^{9}(X_i - Y_2)^2$，所以有

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

因为 Y_1, Y_2, S^2 相互独立，由 t 分布的定义有

$$\frac{\frac{\sqrt{2}(Y_1-Y_2)}{\sigma}}{\sqrt{\frac{\frac{2S^2}{\sigma^2}}{2}}} = \frac{\sqrt{2}(Y_1 - Y_2)}{S} \sim t(2)$$

例 7

设 $X_1, X_2, \cdots, X_n (n > 2)$ 是来自总体 $N(0,1)$ 的一个样本，$\overline{X}$ 为样本均值，记 $Y_i = X_i - \overline{X}$. 求

(1)Y_i 的方差 $D(Y_i)(i = 1, 2, \cdots, n)$;(2)$Y_1$ 与 Y_n 的协方差 $\mathrm{Cov}(Y_1, Y_n)$.

例 7

设 $X_1, X_2, \cdots, X_n (n > 2)$ 是来自总体 $N(0,1)$ 的一个样本，$\overline{X}$ 为样本均值，记 $Y_i = X_i - \overline{X}$. 求

(1)Y_i 的方差 $D(Y_i)(i = 1, 2, \cdots, n)$;(2)$Y_1$ 与 Y_n 的协方差 $\mathrm{Cov}(Y_1, Y_n)$.

解：(1) $D(Y_i) = D(X_i - \overline{X}) = D(X_i) + D(\overline{X}) - 2\,\mathrm{Cov}(X_i, \overline{X})$

例 7

设 $X_1, X_2, \cdots, X_n (n > 2)$ 是来自总体 $N(0,1)$ 的一个样本，$\overline{X}$ 为样本均值，记 $Y_i = X_i - \overline{X}$. 求
(1)Y_i 的方差 $D(Y_i)(i = 1, 2, \cdots, n)$;(2)$Y_1$ 与 Y_n 的协方差 $\mathrm{Cov}(Y_1, Y_n)$.

解：(1) $D(Y_i) = D(X_i - \overline{X}) = D(X_i) + D(\overline{X}) - 2\,\mathrm{Cov}(X_i, \overline{X})$

例 7

设 $X_1, X_2, \cdots, X_n (n > 2)$ 是来自总体 $N(0,1)$ 的一个样本，$\overline{X}$ 为样本均值，记 $Y_i = X_i - \overline{X}$. 求

(1)Y_i 的方差 $D(Y_i)(i = 1, 2, \cdots, n)$;(2)$Y_1$ 与 Y_n 的协方差 $\mathrm{Cov}(Y_1, Y_n)$.

解：(1) $D(Y_i) = D(X_i - \overline{X}) = D(X_i) + D(\overline{X}) - 2\,\mathrm{Cov}(X_i, \overline{X})$

$$\mathrm{Cov}\left(X_i, \overline{X}\right) = \mathrm{Cov}\left(X_i, \frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right)$$

$$= \frac{1}{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j) = \frac{D(X_i)}{n} = \frac{1}{n}$$

例 7

设 $X_1, X_2, \cdots, X_n (n > 2)$ 是来自总体 $N(0,1)$ 的一个样本，$\overline{X}$ 为样本均值，记 $Y_i = X_i - \overline{X}$. 求

(1)Y_i 的方差 $D(Y_i)(i = 1, 2, \cdots, n)$;(2)$Y_1$ 与 Y_n 的协方差 $\mathrm{Cov}(Y_1, Y_n)$.

解：(1) $D(Y_i) = D(X_i - \overline{X}) = D(X_i) + D(\overline{X}) - 2\,\mathrm{Cov}(X_i, \overline{X})$

$$\mathrm{Cov}\left(X_i, \overline{X}\right) = \mathrm{Cov}\left(X_i, \frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right)$$

$$= \frac{1}{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j) = \frac{D(X_i)}{n} = \frac{1}{n}$$

故 $D(Y_i) = D(X_i) + D(\overline{X}) - 2\,\mathrm{Cov}(X_i, \overline{X}) = 1 + \frac{1}{n} - 2 \times \frac{1}{n} = 1 - \frac{1}{n}$

(2)　由题意可知 $Y_1 = X_1 - \overline{X}, Y_n = X_n - \overline{X}$，所以

(2)　由题意可知 $Y_1 = \boldsymbol{X_1} - \overline{\boldsymbol{X}}, Y_n = \boldsymbol{X_n} - \overline{\boldsymbol{X}}$**，所以**

$$\operatorname{Cov}(\boldsymbol{Y_1}, \boldsymbol{Y_n}) = \operatorname{Cov}\left(\boldsymbol{X_1} - \overline{\boldsymbol{X}}, \boldsymbol{X_n} - \overline{\boldsymbol{X}}\right)$$

(2)　由题意可知 $Y_1 = X_1 - \overline{X}, Y_n = X_n - \overline{X}$**，所以**

$$\begin{aligned}\operatorname{Cov}(\boldsymbol{Y}_1, \boldsymbol{Y}_n) &= \operatorname{Cov}\left(\boldsymbol{X}_1 - \overline{\boldsymbol{X}}, \boldsymbol{X}_n - \overline{\boldsymbol{X}}\right) \\ &= \operatorname{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_n) - \operatorname{Cov}(\boldsymbol{X}_1, \overline{\boldsymbol{X}}) - \operatorname{Cov}(\overline{\boldsymbol{X}}, \boldsymbol{X}_n) + \operatorname{Cov}(\overline{\boldsymbol{X}}, \overline{\boldsymbol{X}})\end{aligned}$$

(2)　由题意可知 $Y_1 = X_1 - \overline{X}, Y_n = X_n - \overline{X}$**，所以**

$$
\begin{aligned}
\operatorname{Cov}(Y_1, Y_n) &= \operatorname{Cov}\left(X_1 - \overline{X}, X_n - \overline{X}\right) \\
&= \operatorname{Cov}(X_1, X_n) - \operatorname{Cov}(X_1, \overline{X}) - \operatorname{Cov}(\overline{X}, X_n) + \operatorname{Cov}(\overline{X}, \overline{X}) \\
&= 0 - \frac{1}{n} - \frac{1}{n} + \frac{1}{n}
\end{aligned}
$$

(2)　由题意可知 $\boldsymbol{Y_1 = X_1 - \overline{X}, Y_n = X_n - \overline{X}}$**，所以**

$$
\begin{aligned}
\operatorname{Cov}(\boldsymbol{Y_1}, \boldsymbol{Y_n}) &= \operatorname{Cov}\left(\boldsymbol{X_1} - \overline{\boldsymbol{X}}, \boldsymbol{X_n} - \overline{\boldsymbol{X}}\right) \\
&= \operatorname{Cov}(\boldsymbol{X_1}, \boldsymbol{X_n}) - \operatorname{Cov}(\boldsymbol{X_1}, \overline{\boldsymbol{X}}) - \operatorname{Cov}(\overline{\boldsymbol{X}}, \boldsymbol{X_n}) + \operatorname{Cov}(\overline{\boldsymbol{X}}, \overline{\boldsymbol{X}}) \\
&= \boldsymbol{0} - \frac{\boldsymbol{1}}{\boldsymbol{n}} - \frac{\boldsymbol{1}}{\boldsymbol{n}} + \frac{\boldsymbol{1}}{\boldsymbol{n}} \\
&= -\frac{\boldsymbol{1}}{\boldsymbol{n}}
\end{aligned}
$$

例 8

设 $X_1, X_2, \cdots, X_n$ 是来自标准正态总体 $N(0,1)$ 的一个样本，且 $\overline{X}$ 为样本均值，S^2 为样本方差，求

(1)$E\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$；(2)$D\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$.

例 8

设 $X_1, X_2, \cdots, X_n$ 是来自标准正态总体 $N(0,1)$ 的一个样本，且 $\overline{X}$ 为样本均值，S^2 为样本方差，求

(1)$E\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$；(2)$D\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$.

解：(1) $$E\left(\overline{X}^2 - \frac{1}{n}S^2\right) = E\left(\overline{X}^2\right) - \frac{1}{n}E(S^2)$$

例 8

设 $X_1, X_2, \cdots, X_n$ 是来自标准正态总体 $N(0,1)$ 的一个样本，且 $\overline{X}$ 为样本均值，S^2 为样本方差，求

(1)$E\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$；(2)$D\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$.

解：(1)

$$\begin{aligned} E\left(\overline{X}^2 - \frac{1}{n}S^2\right) &= E\left(\overline{X}^2\right) - \frac{1}{n}E(S^2) \\ &= D\left(\overline{X}\right) + \left[E\left(\overline{X}\right)\right]^2 - \frac{1}{n}E(S^2) \end{aligned}$$

例 8

设 $X_1, X_2, \cdots, X_n$ 是来自标准正态总体 $N(0,1)$ 的一个样本，且 $\overline{X}$ 为样本均值，S^2 为样本方差，求

(1)$E\left(\overline{X}^2-\frac{1}{n}S^2\right)$；(2)$D\left(\overline{X}^2-\frac{1}{n}S^2\right)$.

解：(1)

$$E\left(\overline{X}^2-\frac{1}{n}S^2\right)=E\left(\overline{X}^2\right)-\frac{1}{n}E(S^2)$$

$$=D\left(\overline{X}\right)+\left[E\left(\overline{X}\right)\right]^2-\frac{1}{n}E(S^2)$$

$$=\frac{D(X)}{n}+\left[E\left(X\right)\right]^2-\frac{1}{n}D(X)$$

例 8

设 $X_1, X_2, \cdots, X_n$ 是来自标准正态总体 $N(0,1)$ 的一个样本，且 $\overline{X}$ 为样本均值，S^2 为样本方差，求

(1)$E\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$；(2)$D\left(\overline{X}^2 - \dfrac{1}{n}S^2\right)$.

解：(1)

$$
\begin{aligned}
E\left(\overline{X}^2 - \frac{1}{n}S^2\right) &= E\left(\overline{X}^2\right) - \frac{1}{n}E(S^2) \\
&= D\left(\overline{X}\right) + \left[E\left(\overline{X}\right)\right]^2 - \frac{1}{n}E(S^2) \\
&= \frac{D(X)}{n} + \left[E\left(X\right)\right]^2 - \frac{1}{n}D(X) \\
&= \frac{1}{n} + 0 - \frac{1}{n} = 0
\end{aligned}
$$

(2) 因为 $\overline{X}$ 与 S^2 相互独立，所以

$$D\left(\overline{X}^2-\frac{1}{n}S^2\right)=D(\overline{X}^2)+\frac{1}{n^2}D(S^2)$$

由于 $\overline{X}\sim N\left(0,\frac{1}{n}\right)$，故 $\dfrac{\overline{X}}{\sqrt{\frac{1}{n}}}\sim N(0,1)$，从而 $\left(\dfrac{\overline{X}}{\sqrt{\frac{1}{n}}}\right)^2=n\overline{X}^2\sim\chi^2(1)$

(2) 因为 $\overline{X}$ 与 S^2 相互独立，所以

$$D\left(\overline{X}^2-\frac{1}{n}S^2\right)=D(\overline{X}^2)+\frac{1}{n^2}D(S^2)$$

由于 $\overline{X}\sim N\left(0,\frac{1}{n}\right)$，故 $\dfrac{\overline{X}}{\sqrt{\frac{1}{n}}}\sim N(0,1)$，从而 $\left(\dfrac{\overline{X}}{\sqrt{\frac{1}{n}}}\right)^2=n\overline{X}^2\sim\chi^2(1)$

利用 χ^2 分布的性质有

$$D(\overline{X}^2)=\frac{2}{n^2},D(S^2)=\frac{2\sigma^4}{n-1}=\frac{2}{n-1}$$

(2) 因为 $\overline{X}$ 与 S^2 相互独立，所以

$$D\left(\overline{X}^2-\frac{1}{n}S^2\right)=D(\overline{X}^2)+\frac{1}{n^2}D(S^2)$$

由于 $\overline{X}\sim N\left(0,\frac{1}{n}\right)$，故 $\frac{\overline{X}}{\sqrt{\frac{1}{n}}}\sim N(0,1)$，从而 $\left(\frac{\overline{X}}{\sqrt{\frac{1}{n}}}\right)^2=n\overline{X}^2\sim\chi^2(1)$

利用 χ^2 分布的性质有

$$D(\overline{X}^2)=\frac{2}{n^2},D(S^2)=\frac{2\sigma^4}{n-1}=\frac{2}{n-1}$$

故 $D\left(\overline{X}^2-\frac{1}{n}S^2\right)=D(\overline{X}^2)+\frac{1}{n^2}D(S^2)=\frac{2}{n^2}+\frac{1}{n^2}\frac{2}{n-1}=\frac{2}{n(n-1)}$

例 9

设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，且 $\overline{X}_n$ 为样本均值，S_n^2 为样本方差，又设 X_{n+1} 也服从 $N(\mu, \sigma^2)$ 分布，并且与 $X_1, X_2, \cdots, X_n$ 相互独立，求统计量 $Y = \dfrac{\overline{X}_n - X_{n+1}}{S_n}\sqrt{\dfrac{n}{n+1}}$ 服从何种分布?

例 9

设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本，且 $\overline{X}_n$ 为样本均值，S_n^2 为样本方差，又设 X_{n+1} 也服从 $N(\mu, \sigma^2)$ 分布，并且与 $X_1, X_2, \cdots, X_n$ 相互独立，求统计量 $Y = \dfrac{\overline{X}_n - X_{n+1}}{S_n}\sqrt{\dfrac{n}{n+1}}$ 服从何种分布?

解：因为 $\overline{X}_n \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$，且 X_{n+1} 与 $\overline{X}_n$ 相互独立，从而

$$\overline{X}_n - X_{n+1} \sim N\left(0, \left(1 + \frac{1}{n}\right)\sigma^2\right)$$

标准化得

$$\frac{\overline{X}_n - X_{n+1}}{\sigma\sqrt{\dfrac{n+1}{n}}} \sim N(0,1)$$

$$\frac{\overline{X}_n - X_{n+1}}{\sigma\sqrt{\dfrac{n+1}{n}}} \sim N(0,1)$$

由于 $\dfrac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$，且 S_n^2 与 $\overline{X}_n, X_{n+1}$ 均相互独立，则由 t 分布的定义有

$$\frac{\overline{X}_n - X_{n+1}}{\sigma\sqrt{\dfrac{n+1}{n}}} \sim N(0,1)$$

由于 $\dfrac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$，且 S_n^2 与 $\overline{X}_n, X_{n+1}$ 均相互独立，则由 t 分布的定义有

$$\frac{\dfrac{\overline{X}_n - X_{n+1}}{\sigma\sqrt{\dfrac{n+1}{n}}}}{\sqrt{\dfrac{\dfrac{(n-1)S_n^2}{\sigma^2}}{n-1}}} = \frac{\overline{X}_n - X_{n+1}}{S_n}\sqrt{\frac{n}{n+1}} \sim t(n-1)$$

概率论与数理统计

概率论与数理统计教研团队[1]

[1] 数学科学学院
哈尔滨工程大学

2024 年　春

大工至善
大学至真