

内容大纲

- ❑ 第六章内容回顾和例题讲解
- ❑ 第七章内容回顾和例题讲解
- ❑ Excel 实验
 - 使用函数计算样本均值、样本方差等统计量
 - 绘制直方图和箱线图
 - 进行区间估计

第六章：样本及抽样分布 习题课

第六章作业提交截止时间

2018年12月4日

一、重点及难点

1. 重点

抽样分布中的常用统计量、经验分布函数

数理统计中的三个重要分布及性质

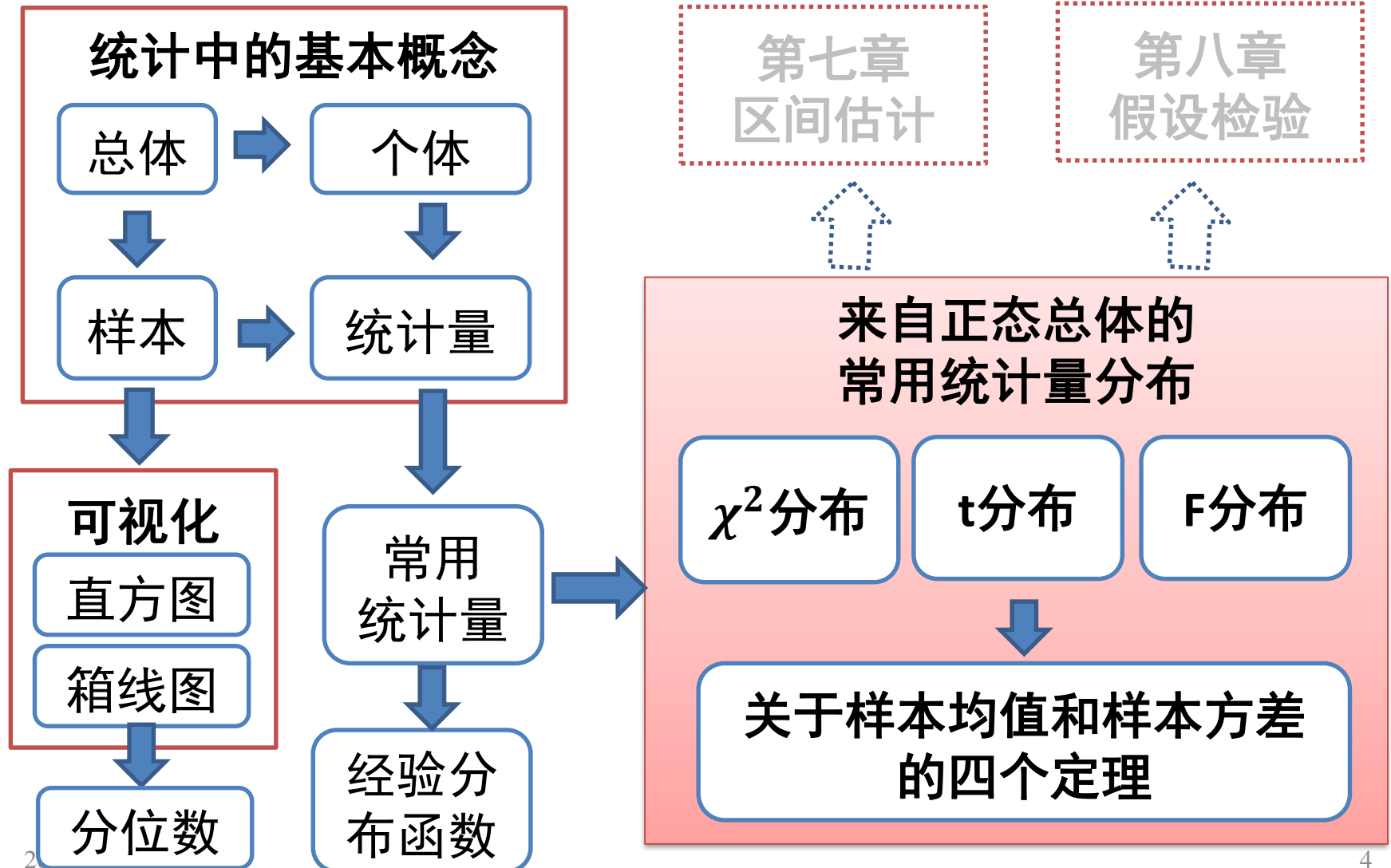
正确理解分位数的概念并会查表求其数值

2. 难点

χ^2 变量、t变量、F变量的构造及性质

运用与正态总体有关的样本数字特征解决实际问题

二、主要内容



总体、个体、样本

- **总体**：试验的全部可能的观察值。
- **个体**：总体中每一个可能的观察值。
- **样本**：设 X 是具有分布函数 F 的随机变量， X_1, X_2, \dots, X_n 是具有**同一分布**函数 F 、**相互独立**的随机变量，则称 X_1, X_2, \dots, X_n 为从分布函数 F （或总体 F 、或总体 X ）得到的**容量**为 n 的**简单随机样本**，简称**样本**。

经验分布函数

- 与总体分布函数 $F(x)$ 相对应的统计量
- 设 X_1, X_2, \dots, X_n 是总体 F 的一个样本。用 $S(x)$, $-\infty < x < +\infty$ 表示 X_1, X_2, \dots, X_n 中**不大于** x 的随机变量个数，定义**经验分布函数(样本分布函数)**为：

$$F_n(x) = \frac{1}{n} S(x), -\infty < x < +\infty$$

经验分布函数的观察值

- 一般地，设 x_1, x_2, \dots, x_n 是总体 F 的一个样本值，先将 x_1, x_2, \dots, x_n 按自小到大排序，并重新编号满足：

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

则**经验分布函数 $F_n(x)$ 的观察值**为：

$$F_n(x) = \begin{cases} 0, & \text{若 } x < x_{(1)} \\ \frac{k}{n}, & \text{若 } x_{(k)} \leq x < x_{(k+1)} \\ 1, & \text{若 } x \geq x_{(n)} \end{cases}$$

统计量

- 完全由样本所决定的量，叫做**统计量**
- 设 X_1, X_2, \dots, X_n 是来自总体 X 的样本，常用的样本数字特征统计量如下

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

反映了总体均值的信息

反映了总体方差的信息

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

反映了总体标准差的信息

统计量

- 完全由样本所决定的量，叫做**统计量**
- 设 X_1, X_2, \dots, X_n 是来自总体 X 的样本，常用的样本数字特征统计量如下

k阶样本原点矩

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k,$$

反映了总体
k阶原点矩的信息

$k = 1, 2, \dots$

k阶样本中心矩

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

反映了总体
k阶中心矩的信息

$k = 1, 2, \dots$

- 统计量与数字特征的区别：
样本数字特征**统计量**是**随机变量**
(随机变量的)**数字特征**是**常数**

概念对比

概率	数理统计	计算机科学(PR&ML)
样本空间	总体	
随机事件	样本	训练样本集
样本点	个体	单个样本
概率分布函数	经验分布函数	
期望	样本均值	
方差	样本方差	类内距离
	估计	学习
协方差	样本协方差	特征

统计学中的三大抽样分布

■ χ^2 分布

主要用于列联分析

$$\chi^2 = \sum_{i=1}^n \underbrace{X_i^2}_{\substack{\text{i.i.d. } \sim N(0,1)}} \sim \chi^2(n)$$

■ t分布（Student分布、学生氏分布）

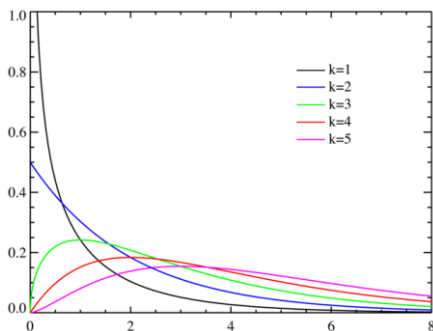
主要用于小样本分析

$$t = \frac{\underbrace{X}_{\sim N(0,1)}}{\sqrt{\underbrace{Y/n}_{\sim \chi^2(n)}}} \sim t(n)$$

■ F分布

主要用于方差分析

$$F = \frac{\underbrace{U/n_1}_{\sim \chi^2(n_1)}}{\underbrace{V/n_2}_{\sim \chi^2(n_2)}} \sim F(n_1, n_2)$$



χ^2 分布

独立同分布

- 设 X_1, X_2, \dots, X_n 是来自总体 $N(0, 1)$ 的**样本**，则称统计量 $\chi^2 = \sum_{i=1}^n X_i^2$ 服从**自由度为 n 的 χ^2 分布**，记为 $\chi^2 \sim \chi^2(n)$ 。
- χ^2 分布的**可加性**：设 $\chi_1^2 \sim \chi^2(n_1)$ ， $\chi_2^2 \sim \chi^2(n_2)$ ，并且 χ_1^2 与 χ_2^2 **相互独立**，则有

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

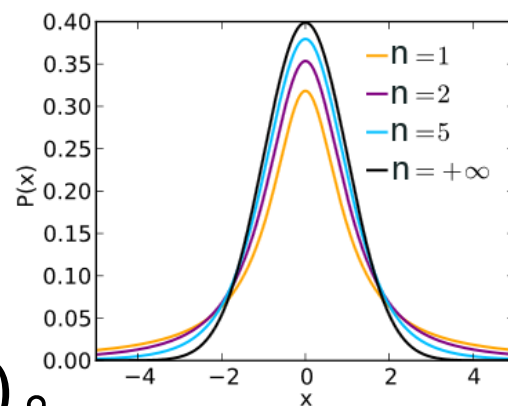
- χ^2 分布的**数学期望和方差**：若 $\chi^2 \sim \chi^2(n)$ ，则

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

t分布 (Student分布, 学生氏分布)

- 设 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 且 X, Y 相互独立, 则称随机变量

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



服从自由度为 n 的 t 分布, 记为 $t \sim t(n)$ 。

- t 分布关于 $x = 0$ 对称: $n = 1$ 时, $E(t)$ 不存在;
 $n \geq 2$ 时, $E(t) = 0$ 。
- $n \rightarrow \infty$ 时, t 分布趋向于标准正态分布。实际应用中, 当 $n \geq 45$ 时, 可用标准正态分布近似。

F分布

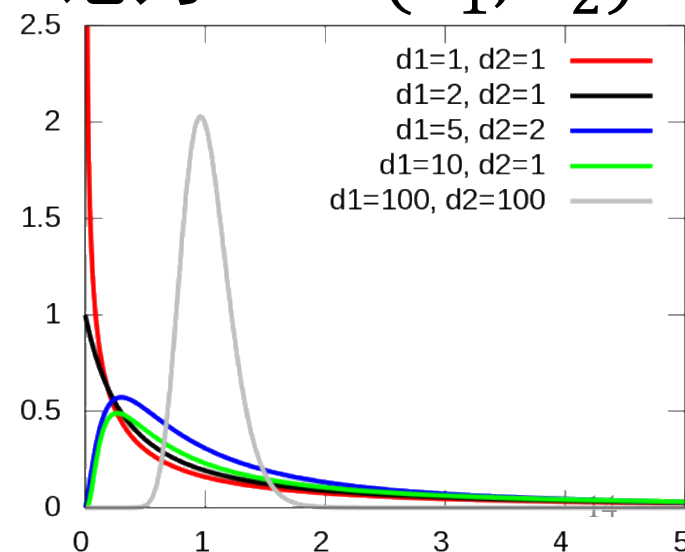
- 设 $U \sim \chi^2(n_1)$, $V \sim \chi^2(n_2)$, 且 U, V 相互独立, 则称随机变量

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

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

- 若 $F \sim F(n_1, n_2)$, 则

$$\frac{1}{F} \sim F(n_2, n_1)。$$



正态总体的样本均值与样本方差的分布

设 X_1, \dots, X_n 是来自总体 $X \sim N(\mu, \sigma^2)$ 的随机样本, 则

- **样本均值**的分布

证明了 \bar{X} 逼近 μ 的合理性

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

- **样本方差**的分布

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = n-1 \Rightarrow E[S^2] = \sigma^2$$

证明了 S^2 逼近 σ^2 的合理性

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$$

- 样本均值 \bar{X} 与样本方差 S^2 **相互独立**

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

正态总体的样本均值与样本方差的分布

设 X_1, \dots, X_{n_1} 与 Y_1, \dots, Y_{n_2} 是来自总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的随机样本, 且两样本**相互独立**, 则

- $\frac{s_1^2/s_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$

- 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时,

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

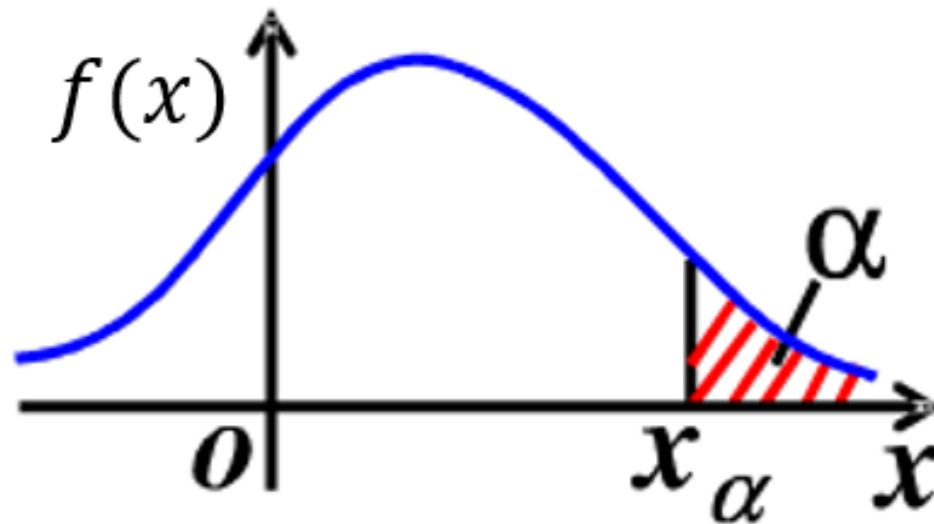
其中, $S_w = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$, $S_w = \sqrt{S_w^2}$ 。

分布的分位数（分位点）

设总体 X 的分布函数为 $F(x)$ ，密度函数为 $f(x)$ 。
定义满足条件

$$F(x_\alpha) = P\{X \leq x_\alpha\} = 1 - \alpha$$

的数 x_α 为分布函数 $F(x)$ 的上侧 α 分位数（上 α 分位点）。



χ^2 分布的分位数

- 对于给定正数 α , $0 < \alpha < 1$, 满足条件

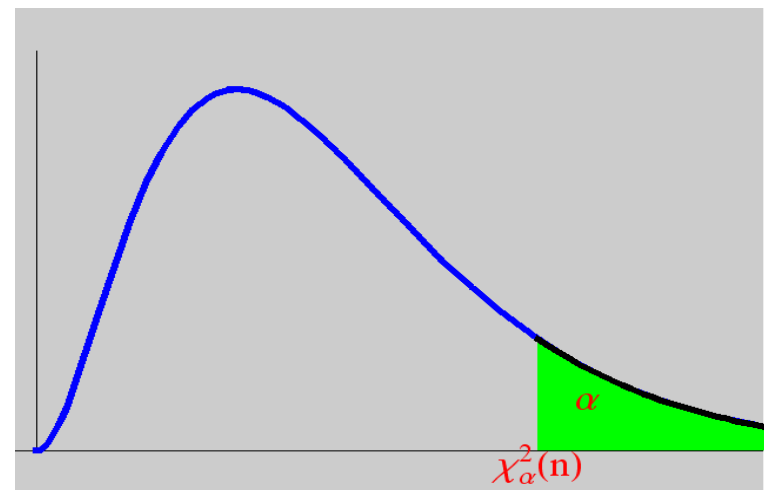
$$P\{\chi^2 > \chi_{\alpha}^2(n)\} = \alpha$$

的点 $\chi_{\alpha}^2(n)$ 就是 $\chi^2(n)$ 分布的上 α 分位点。

- R. A. Fisher证明当 n 充分大时 ($n > 40$),

$$\chi_{\alpha}^2(n) \approx \frac{1}{2} (z_{\alpha} + \sqrt{2n-1})^2$$

其中 z_{α} 是标准正态分布的上 α 分位点。



t分布的分位数

- 对于给定正数 α , $0 < \alpha < 1$, 满足条件

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

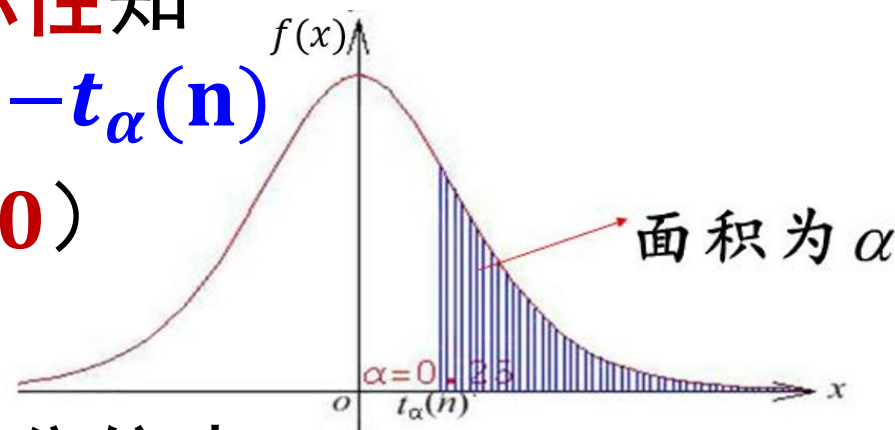
的点 $t_{\alpha}(n)$ 就是 $t(n)$ 分布的上 α 分位点。

- 由概率密度函数的对称性知

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

- 当 n 充分大时 ($n > 40$)

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



z_{α} 是标准正态分布的上 α 分位点。

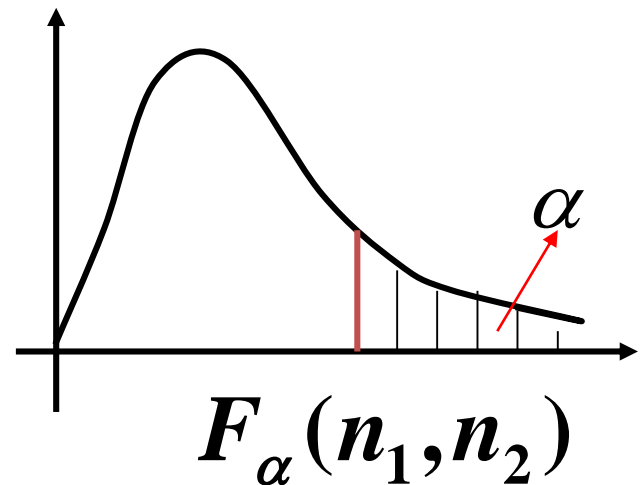
F分布的分位数

- 对于给定正数 α , $0 < \alpha < 1$, 满足条件

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

的点 $F_{\alpha}(n)$ 就是 $F(n_1, n_2)$ 分布的上 α 分位点。

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



三、典型例题

例1. 设 $X_1, X_2, X_3, Y_1, Y_2, Y_3, Y_4$ 是来自正态总体 $N(0, 1)$ 的独立样本, 则下列统计量服从什么分布?

$$X_1^2 + X_2^2 + X_3^2 \sim \chi^2(3) \quad \frac{\sqrt{3}X_1}{\sqrt{X_1^2 + X_2^2 + X_3^2}} \sim t(3)$$

$$\frac{4 \sum_{i=1}^3 X_i^2}{3 \sum_{i=1}^4 Y_i^2} \sim F(3, 4)$$

【思路】 根据 χ^2 变量、t 变量、F 变量的构成模式判断。

$\chi^2 \sim \chi^2(n)$: 如果 X_1, X_2, \dots, X_n i. i. d $\sim N(0, 1)$, $\chi^2 = X_1^2 + \dots + X_n^2$

$T \sim t(n)$: 如果 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, $T = X / \sqrt{\frac{Y}{n}}$

$F \sim F(n, m)$: 如果 $X \sim \chi^2(n)$, $Y \sim \chi^2(m)$, $F = \frac{X}{n} / \frac{Y}{m}$

例2. 设 X_1, X_2, \dots, X_{10} 独立同分布, $X_i \sim N(0, 0.3^2)$, 求 c 的值, 使 $c \sum_{i=1}^{10} X_i^2$ 服从 χ^2 分布, 并求自由度 n 。

【思路】 根据 χ^2 变量的构成模式计算。 χ^2 可以表示为 n 个标准正态分布随机变量的平方和。

解:

$$X_i \sim N(0, 0.3^2), \rightarrow \frac{X_i}{0.3} \sim N(0, 1)$$

$$\left(\frac{X_1}{0.3}\right)^2 + \left(\frac{X_2}{0.3}\right)^2 + \dots + \left(\frac{X_{10}}{0.3}\right)^2 \sim \chi^2(10)$$

$$\left(\frac{X_1}{0.3}\right)^2 + \left(\frac{X_2}{0.3}\right)^2 + \dots + \left(\frac{X_{10}}{0.3}\right)^2 = \frac{1}{0.09} \sum_{i=1}^{10} X_i^2$$

所以 $c = \frac{1}{0.09}$, 自由度 $n = 10$ 。

例3. 求总体 $N(20,3)$ 的容量分别为10,15的两独立样本均值差的绝对值大于0.3的概率。

【思路】 样本均值服从正态分布，正态分布之差服从正态分布。

解：两样本均值分别记为 \bar{X} 、 \bar{Y} ， 则 $\bar{X} \sim N(20, \frac{3}{10})$ ，

$$\bar{Y} \sim N(20, \frac{3}{15})$$

$$\bar{X} - \bar{Y} \sim N\left(20 - 20, \frac{3}{10} + \frac{3}{15}\right) = N\left(0, \frac{1}{2}\right)$$

所求概率为 $p = P\{|\bar{X} - \bar{Y}| > 0.3\} = 1 - P\{|\bar{X} - \bar{Y}| \leq 0.3\} = 1 - P\left\{\frac{-0.3}{\sqrt{1/2}} \leq \frac{\bar{X} - \bar{Y}}{\sqrt{1/2}} \leq \frac{0.3}{\sqrt{1/2}}\right\} = 2(1 - \Phi(0.42)) = 0.6744.$

例4. 从总体 $X \sim N(5, 2^2)$ 中随机抽取一容量为25的样本，求样本均值 \bar{X} 落在4.2到5.8之间的概率，以及样本方差 S^2 大于6.07的概率。

【思路】 样本均值服从正态分布，样本方差服从 χ^2 分布。

解：因为总体 $X \sim N(5, 2^2)$ ， $n = 25$ ，故

$$\bar{X} \sim \frac{1}{25} \sum_{i=1}^{25} X_i \sim N\left(5, \frac{4}{25}\right); \quad \frac{(25-1)S^2}{2^2} \sim \chi^2(24).$$

于是，样本均值 \bar{X} 落在4.2到5.8之间的概率为

$$\begin{aligned} P\{4.2 < \bar{X} < 5.8\} &= P\left\{\frac{4.2-5}{2/5} < \frac{\bar{X}-5}{2/5} < \frac{5.8-5}{2/5}\right\} \\ &= P\left\{-2 < \frac{\bar{X}-5}{2/5} < 2\right\} = 2\phi(2) - 1 = 2 \times 0.964 - 1 = 0.908 \end{aligned}$$

例4. 从总体 $X \sim N(5, 2^2)$ 中随机抽取一容量为25的样本，求样本均值 \bar{X} 落在4.2到5.8之间的概率，以及样本方差 S^2 大于6.07的概率。

【思路】 样本均值服从正态分布，样本方差服从 χ^2 分布。

解(续)：因为总体 $X \sim N(5, 2^2)$ ， $n = 25$ ，故

$$\bar{X} \sim \frac{1}{25} \sum_{i=1}^{25} X_i \sim N\left(5, \frac{4}{25}\right); \quad \frac{(25-1)S^2}{2^2} \sim \chi^2(24).$$

于是，样本方差 S^2 大于6.07的概率为

$$P\{S^2 > 6.07\} = P\left\{\frac{24S^2}{4} > \frac{6.07 \times 24}{4}\right\} = P\{\chi^2(24) > 36.42\} = 0.05$$

其中的0.05是反查 χ^2 分布表所得， $36.42 = \chi_{0.05}^2(24)$ ，即是自由度为24的 χ^2 分布的上侧0.05分位数。

例5. 设总体 $X \sim N(\mu, \sigma^2)$, 从总体中取一个容量为 $n = 16$ 的样本 $(X_1, X_2, \dots, X_{16})$, 求概率

$$(1) P \left\{ \frac{\sigma^2}{2} \leq \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \leq 2\sigma^2 \right\};$$

$$(2) P \left\{ \frac{\sigma^2}{2} \leq \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \leq 2\sigma^2 \right\}.$$

【思路】

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$$
$$\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

例5续. (1) $P\left\{\frac{\sigma^2}{2} \leq \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \leq 2\sigma^2\right\};$

解: (1) 因为 X_1, X_2, \dots, X_{16} 是来自正态总体的样本, 所以 $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n)$, 于是

$$\begin{aligned} P\left\{\frac{\sigma^2}{2} \leq \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \leq 2\sigma^2\right\} &= P\left\{8 \leq \frac{1}{\sigma^2} \sum_{i=1}^{16} (X_i - \mu)^2 \leq 32\right\} \\ &= P\{8 \leq \chi^2(16) \leq 32\} \\ &= P\{\chi^2(16) \leq 32\} - P\{\chi^2(16) \leq 8\} \\ &= (1 - P\{\chi^2(16) \geq 32\}) - (1 - P\{\chi^2(16) \geq 8\}) \\ &= (1 - 0.01) - (1 - 0.95) = 0.94 \end{aligned}$$

例5续. (2) $P\left\{\frac{\sigma^2}{2} \leq \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \leq 2\sigma^2\right\}.$

解: (2) 因为 $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1)$, 于是

$$P\left\{\frac{\sigma^2}{2} \leq \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \leq 2\sigma^2\right\} = P\left\{8 \leq \frac{1}{\sigma^2} \sum_{i=1}^{16} (X_i - \bar{X})^2 \leq 32\right\}$$

$$= P\{8 \leq \chi^2(15) \leq 32\}$$

$$= P\{\chi^2(15) \geq 8\} - P\{\chi^2(15) \geq 32\}$$

$$\approx 0.90 - 0.005 = 0.895$$

例6. 为了估计灯泡寿命，抽取10个灯泡测试。求得样品均值 $\bar{X} = 1500$ 小时，样品标准差 $S = 20$ 小时。若灯泡寿命服从正态分布 $N(\mu, \sigma_0^2)$ ，以 \bar{X} 作为 μ 的估计，试求：

(1) 若已知 $\sigma_0^2 = 400$ ， \bar{X} 偏离总体均值 μ 不超过12小时的概率。

(2) 若未知 σ_0^2 ， \bar{X} 偏离总体均值 μ 不超过12小时的概率。

【思路】 (1) 已知 σ_0^2 ，根据 $\bar{X} \sim N(\mu, \sigma_0^2/n)$ 知

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

(2) 未知 σ_0^2 ，利用正态总体的样本均值和样本方差的分布

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

解：(1) 已知 $\sigma_0^2 = 400$ ，则 $\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0,1)$ ，

$$P\{|\bar{X} - \mu| \leq 12\} = P\left\{\left|\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}}\right| \leq \frac{12}{\sigma_0/\sqrt{n}}\right\} = P\left\{\left|\frac{\bar{X} - \mu}{20/\sqrt{10}}\right| \leq 1.897\right\}$$

$$= P\left\{-1.897 \leq \frac{\bar{X} - \mu}{20/\sqrt{10}} \leq 1.897\right\}$$

标准正态分布的对称性

$$= 2P\left\{\frac{\bar{X} - \mu}{20/\sqrt{10}} \leq 1.897\right\} - 1$$

$$= 2\Phi(1.897) - 1$$

$$\approx 2 \times 0.9713 - 1 = 0.94$$

所以，在知道总体方差 $\sigma_0^2 = 400$ 的情况下，以 \bar{X} 作为 μ 的估计，绝对误差不超过12小时的概率约为0.94。

解：(2) 未知 σ_0^2 ，则 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ ，

$$\begin{aligned}
 P\{|\bar{X} - \mu| \leq 12\} &= P\left\{\left|\frac{\bar{X} - \mu}{S/\sqrt{n}}\right| \leq \frac{12}{S/\sqrt{n}}\right\} = P\left\{\left|\frac{\bar{X} - \mu}{20/\sqrt{10}}\right| \leq 1.897\right\} \\
 &= P\left\{-1.897 \leq \frac{\bar{X} - \mu}{20/\sqrt{10}} \leq 1.897\right\} \\
 &= P\left\{\frac{\bar{X} - \mu}{20/\sqrt{10}} > -1.897\right\} - P\left\{\frac{\bar{X} - \mu}{20/\sqrt{10}} > 1.897\right\} \\
 &\quad \text{t分布的对称性} \\
 &= 1 - 2 \times P\left\{\frac{\bar{X} - \mu}{20/\sqrt{10}} > 1.897\right\} \\
 &\quad \text{查表 } t_{0.05}(9) = 1.8331 \\
 &\approx 1 - 2 \times 0.05 = 0.9
 \end{aligned}$$

所以，在不知道总体方差的情况下，以 \bar{X} 作为 μ 的估计，绝对误差不超过12小时的概率约为0.9。

第七章：参数估计 习题课

第七章作业提交截止时间

2018年12月11日

一、重点与难点

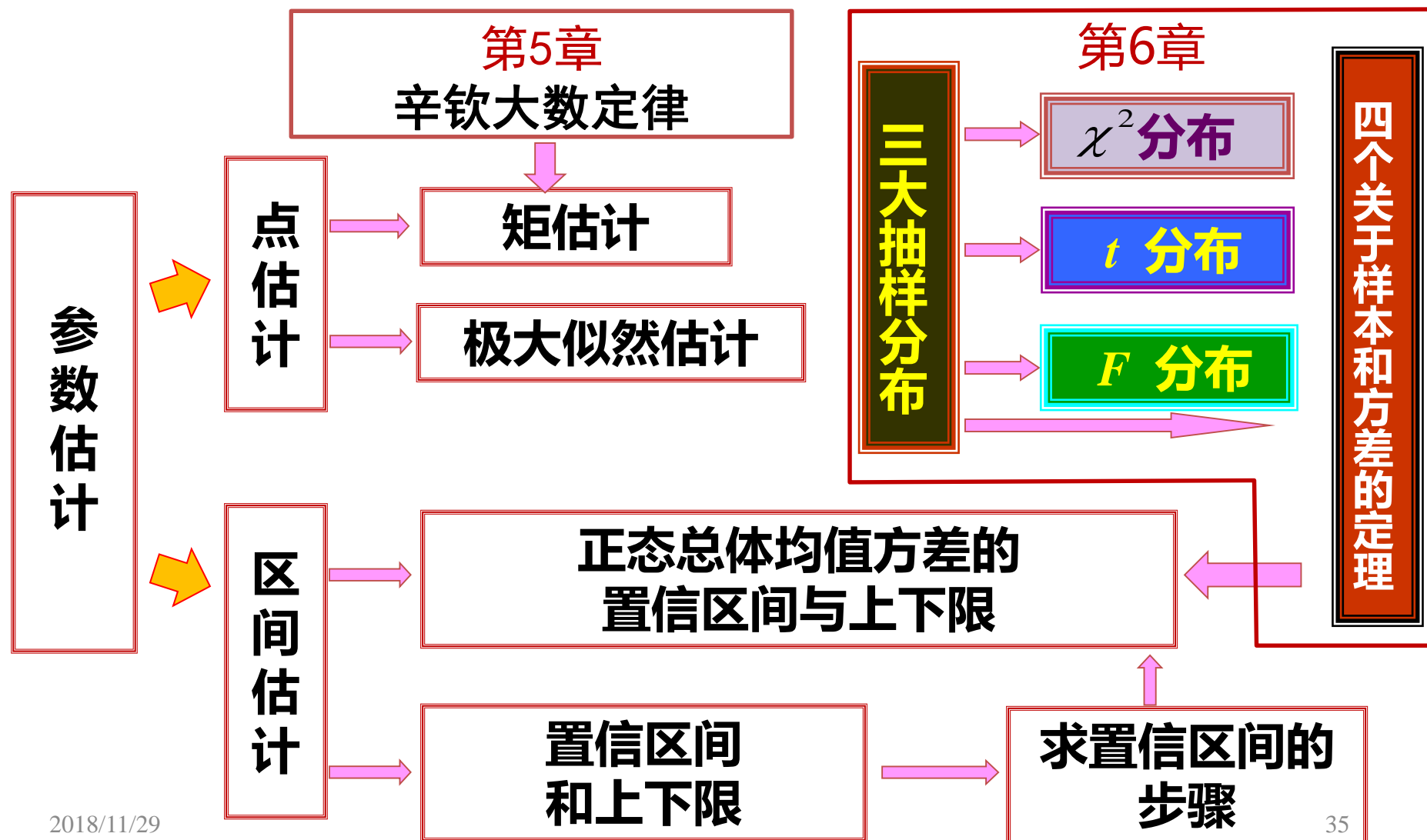
1.重点

两种常用的点估计法：矩估计，最大似然估计
估计量的评选标准
正态总体参数的区间估计

2.难点

应用点估计和区间估计解决实际问题

二、主要内容



点估计：矩估计法

- **基本思想**：以样本的 l 阶矩作为 $E(X^l)$ 的估计量，样本 l 阶矩的连续函数作为 $E(X^l)$ 的连续函数的估计量
- **理论依据**：辛钦大数定律
- **具体步骤**
 1. 求 $\mu_1 = E[X]$ ($\mu_2 = E[X^2], \dots$)
 2. 设样本的 l 阶矩 $A_1 = \mu_1$ ($A_2 = \mu_2, \dots$)
 3. 解上面的方程(组)，得 $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_k)$,
($\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_k), \dots$)

点估计：最大似然估计

- **基本思想**：根据样本选择参数 $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ ，使得该样本发生的**可能性最大**，即使得似然函数 L 达到最大值。
- **似然函数**：设总体 X 的概率密度函数为 $f(x; \theta)$ ，样本值 x_1, x_2, \dots, x_n 的似然函数为

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

- **求解依据**：极值处一阶导为0

点估计：最大似然估计

■ 具体步骤：

1. 写出似然函数 $L(\theta) = \prod_{i=1}^n f(x_i; \theta)$
2. 取对数 $\ln L(\theta) = \sum_{i=1}^n \ln f(x_i; \theta)$
3. 对 θ 求导 $\frac{\partial \ln L(\theta)}{\partial \theta}$ ，并令 $\frac{\partial \ln L(\theta)}{\partial \theta} = 0$ ，求解方程得未知参数 θ 的最大似然估计值 $\hat{\theta}$

对数似然方程

若分布中包含多个未知参数，需对每个参数 θ_i 分别求导，求解**对数似然方程组**

$$\frac{\partial \ln L(\theta_i)}{\partial \theta_i} = 0, i = 1, 2, \dots, k$$

似然方程(组)可能无法求解

点估计：最大似然估计

- 最大似然估计量的不变性

设 θ 的函数 $u = u(\theta)$, $\theta \in \Theta$ 具有**单值反函数**
 $\theta = \theta(u)$, $u \in U$, 又设 $\hat{\theta}$ 是 X 的概率分布中参数
 θ 的最大似然估计, 则 $\hat{u} = u(\hat{\theta})$ 是 $u(\theta)$ 的最大似
然估计。

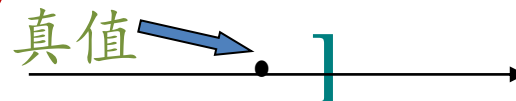
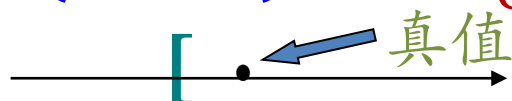
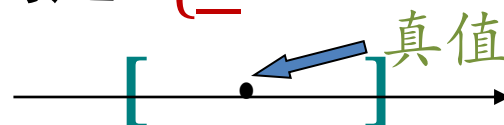
估计量的评选标准

设 X_1, X_2, \dots, X_n 是总体 X 的一个样本， $\theta \in \Theta$ 是包含在总体 X 的分布中的待估参数，估计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$

- **无偏性**：若 $E(\hat{\theta})$ 存在，且 $\forall \theta \in \Theta$ ， $E(\hat{\theta}) = \theta$ ，则 $\hat{\theta}$ 是 θ 的无偏估计量
- **有效性**： $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是 θ 的**无偏**估计量，若 $\forall \theta \in \Theta$ ， $D(\hat{\theta}_1) < D(\hat{\theta}_2)$ ，则 $\hat{\theta}_1$ 较 $\hat{\theta}_2$ 有效
- **相合性**： $\forall \theta \in \Theta$ ， $\hat{\theta} \xrightarrow{P} \theta$ ，则 $\hat{\theta}$ 是 θ 的相合估计量

区间估计

- **区间估计**：根据样本给出未知参数的一个范围，并希望知道这个范围包含该参数的可信程度
- θ 的**置信水平（置信度）为 $1 - \alpha$ 的双侧置信区间**：给定 $0 < \alpha < 1$ ，确定**统计量** $\underline{\theta}(X_1, \dots, X_n)$ 和 $\bar{\theta}(X_1, \dots, X_n)$ ，使得区间 $(\underline{\theta}, \bar{\theta})$ 满足 $P\{\underline{\theta} < \theta < \bar{\theta}\} \geq 1 - \alpha$ 。
- θ 的**置信度为 $1 - \alpha$ 的单侧置信区间**：给定 $0 < \alpha < 1$ ，确定**统计量** $\underline{\theta}(X_1, \dots, X_n)$ ，使得区间 $(\underline{\theta}, +\infty)$ 满足 $P\{\theta > \underline{\theta}\} \geq 1 - \alpha$ ；或**统计量** $\bar{\theta}(X_1, \dots, X_n)$ ，使得区间 $(-\infty, \bar{\theta})$ 满足 $P\{\theta < \bar{\theta}\} \geq 1 - \alpha$ 。



精度与样本容量的关系

- 置信区间长度(长)：估计的精确程度(低)
- 置信水平(高)：估计的可信程度(高)
- **样本容量固定**
 - 置信水平 $1 - \alpha$ 增大 \rightarrow 置信区间长度相应增大
 - 可信程度增大 \rightarrow 区间估计精度相应下降
- **置信水平 $1 - \alpha$ 固定**
 - 样本容量增大 \rightarrow 置信区间长度相应减小
 - 可信程度不变 \rightarrow 区间估计精度相应提高

求解置信区间的一般步骤

设 X_1, X_2, \dots, X_n 是一个样本， θ 是未知参数：

1. 寻求**枢轴量**函数 $W = W(X_1, X_2, \dots, X_n; \theta)$ ，其分布**不依赖于** θ 和其它**未知参数**
2. 对于给定置信水平 $1 - \alpha$ ，求**常数** a, b 使得
$$P\{a < W(X_1, \dots, X_n; \theta) < b\} = 1 - \alpha$$
3. 根据 $a < W(X_1, \dots, X_n; \theta) < b$ 得到与之等价的 θ 的不等式 $\underline{\theta}(X_1, \dots, X_n) < \theta < \bar{\theta}(X_1, \dots, X_n)$ ，则 $(\underline{\theta}, \bar{\theta})$ 为 θ 的置信度为 $1 - \alpha$ 的置信区间。

单个正态总体参数的区间估计

待估参数		枢轴量	枢轴量的分布	双侧置信区间的上、下限
μ	σ^2 已知	$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$N(0,1)$	$\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$
	σ^2 未知	$\frac{\bar{X} - \mu}{S/\sqrt{n}}$	$t(n-1)$	$\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)$
σ^2	μ 已知	$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}$	$\chi^2(n)$	$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\alpha/2}^2(n)}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\alpha/2}^2(n)}$
	μ 未知	$\frac{(n-1)S^2}{\sigma^2}$	$\chi^2(n-1)$	$\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)}$

单侧置信区间的上下限：将双侧置信区间上下限中的

2018/11/29 $z_{\alpha/2}$ 替换成 z_{α} , $t_{\alpha/2}$ 替换成 t_{α} , $\chi_{\alpha/2}^2$ 替换成 χ_{α}^2 即可。

两个正态总体未知参数的置信区间

待估参数		枢轴量	枢轴量的分布	双侧置信区间上下限
$\mu_1 - \mu_2$	σ_1^2, σ_2^2 均已知	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$N(0, 1)$	$\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
	$\sigma_1^2 = \sigma_2^2$, 但未知	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{m} + \frac{1}{n}}}$	$t_{\alpha/2}(m + n - 2)$	$\bar{X} - \bar{Y} \pm t_{\alpha/2}(n_1 + n_2 - 2) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

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

单侧置信区间的上下限：将双侧置信区间上下限中的 $z_{\alpha/2}$ 替换成 z_{α} ， $t_{\alpha/2}$ 替换成 t_{α} 即可。

两个正态总体未知参数的置信区间

待估参数		随机变量	随机变量的分布	双侧置信区间的上、下限
σ_1^2 σ_2^2	μ_1, μ_2 为未知	$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$	$F(m-1, n-1)$	$\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(n_1-1, n_2-1)},$ $\frac{S_1^2}{S_2^2} \frac{1}{F_{1-\alpha/2}(n_1-1, n_2-1)}$

单侧置信区间的上下限：将双侧置信区间上下限中的 $F_{\alpha/2}$ 替换成 F_{α} 即可。

(0-1) 分布参数的置信区间

待估参数	枢轴量	枢轴量的分布	双侧置信区间的上、下限
p	$\frac{n\bar{X} - np}{\sqrt{np(1-p)}}$	$N(0,1)$	$\frac{1}{2a}(-b - \sqrt{b^2 - 4ac}),$ $\frac{1}{2a}(-b + \sqrt{b^2 - 4ac})$

$$a = n + z_{\alpha/2}^2$$

$$b = -(2n\bar{X} + z_{\alpha/2}^2)$$

$$c = n\bar{X}^2$$

三、典型例题

例1. 设总体 X 的概率分布律为: $\begin{pmatrix} 1 & 2 & 3 \\ \theta & \theta/2 & 1 - 3\theta/2 \end{pmatrix}$,
其中 $\theta > 0$ 未知, 现得到样本观测值2,3,2,1,3。求 θ 的矩估计和极大似然估计。

【思路】矩估计: 概率分布中只有一个未知参数, 只需求出 $E(X)$, 令 $E(X) = \bar{X}$ 。

解: (1) 矩估计

$$E(X) = \sum x_k p_k = \theta + 2 \times \frac{\theta}{2} + 3 \times \left(1 - \frac{3\theta}{2}\right) = 3 - \frac{5}{2}\theta$$

$$\bar{X} = (2 + 3 + 2 + 1 + 3)/5 = 11/5$$

令 $E(X) = \bar{X}$, 即 $3 - \frac{5}{2}\theta = \frac{11}{5}$, 则 $\hat{\theta} = \frac{8}{25}$ 。

所以 θ 的矩估计为 $\frac{8}{25}$ 。

例1. 设总体 X 的概率分布律为： $\begin{pmatrix} 1 & 2 & 3 \\ \theta & \theta/2 & 1 - 3\theta/2 \end{pmatrix}$ ，其中 $\theta > 0$ 未知，现得到样本观测值2,3,2,1,3。求 θ 的矩估计和极大似然估计。

【思路】极大似然估计：对所给样本值写出似然函数，再求 θ 的估计值。

解(续)： (2) 先写出似然函数

$$L(\theta) = \frac{\theta}{2} \left(1 - \frac{3\theta}{2}\right) \frac{\theta}{2} \theta \left(1 - \frac{3\theta}{2}\right) = \frac{1}{16} \theta^3 (2 - 3\theta)^2$$

则对数似然函数为：

$$\ln L(\theta) = -\ln 16 + 3 \ln \theta + 2 \ln(2 - 3\theta)$$

θ 的极大似然估计值为 $\hat{\theta} = \max_{\theta} \ln L(\theta)$ 。

$$\text{由 } \frac{d \ln L(\theta)}{d\theta} = \frac{3}{\theta} - \frac{6}{2-3\theta} = 0, \text{ 可得 } \hat{\theta} = 0.4.$$

例2. 设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 未知。 X_1, X_2, \dots, X_n 是来自 X 的样本, 试确定常数 c , 使 $Y = c[(X_1 - X_2)^2 + (X_3 - X_4)^2 + (X_5 - X_6)^2]$ 是 σ^2 的无偏估计。

【思路】 无偏估计即要求 $E(Y) = \sigma^2$ 。

解: $E[(X_1 - X_2)^2] = D(X_1 - X_2) + [E(X_1 - X_2)]^2$
(因为 X_1, X_2 相互独立)
 $= D(X_1) + D(X_2) = 2\sigma^2$ 。

同理, $E[(X_3 - X_4)^2] = E[(X_5 - X_6)^2] = 2\sigma^2$ 。

故, $E(Y) = 6c\sigma^2$ 。

若 Y 是 σ^2 的无偏估计, 则 $E(Y) = \sigma^2$ 。

即 $6c\sigma^2 = \sigma^2$, $c = \frac{1}{6}$ 。

例3. 设总体 X 服从二项分布 $B(N, p)$, X_1, X_2, \dots, X_n 是来自 X 的样本, 试求参数 p 的极大似然估计量 \hat{p} , 并证明 \hat{p} 是 p 的无偏估计。

【思路】极大似然估计：先写出似然函数, \hat{p} 令对数似然函数取极大值。无偏估计即要求 $E(\hat{p}) = p$ 。

解：设样本值为 x_1, x_2, \dots, x_n , 似然函数为

$$L(p) = \prod_{i=1}^n C_N^{x_i} p^{x_i} (1-p)^{N-x_i}$$

对数似然函数为：

$$\ln L(p) = \sum_{i=1}^n \ln C_N^{x_i} + \sum_{i=1}^n x_i \ln p + \sum_{i=1}^n (N - x_i) \ln(1-p)$$

对数似然方程为：

$$\frac{d \ln L(p)}{dp} = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} \left(nN - \sum_{i=1}^n x_i \right) = 0$$

解(续)：由对数似然方程

$$\frac{d \ln L(p)}{dp} = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} \left(nN - \sum_{i=1}^n x_i \right) = 0$$

得 p 的最大似然估计值为 $\hat{p} = \frac{\sum_{i=1}^n x_i}{nN} = \frac{\bar{x}}{N}$,

最大似然估计量为 $\hat{p} = \frac{\bar{X}}{N}$ 。

因为 $E(\hat{p}) = E\left(\frac{\bar{X}}{N}\right) = \frac{E(\bar{X})}{N} = \frac{Np}{N} = p$, 所以 \hat{p} 是 p 的无偏估计。

例4. 设总体 X 的数学期望为 μ , 方差为 σ^2 , X_1, X_2 是来自 X 的样本, 则下面的无偏估计量哪个更有效?

$$(1) \hat{\mu}_1 = \frac{1}{4}X_1 + \frac{3}{4}X_2; \quad (2) \hat{\mu}_2 = \frac{1}{3}X_1 + \frac{2}{3}X_2;$$

$$(3) \hat{\mu}_3 = \frac{3}{8}X_1 + \frac{5}{8}X_2$$

【思路】 $\hat{\mu}_1$ 、 $\hat{\mu}_2$ 、 $\hat{\mu}_3$ 都是 μ 的无偏估计, 那么方差 $D(\hat{\mu}_i)$ 最小的 $\hat{\mu}_i$ 是最有效的。

X_1, X_2 相互独立

解: $D(\hat{\mu}_1) = D\left(\frac{1}{4}X_1 + \frac{3}{4}X_2\right) = \frac{1}{16}D(X_1) + \frac{9}{16}D(X_2) = \frac{5}{8}\sigma^2$

$$D(\hat{\mu}_2) = \frac{5}{9}\sigma^2, \quad D(\hat{\mu}_3) = \frac{17}{32}\sigma^2。$$

$D(\hat{\mu}_3)$ 最小, 所以无偏估计量 $\hat{\mu}_3$ 最有效。

例5. 设总体 $X \sim N(\mu, \sigma_0^2)$, σ_0^2 已知。问抽取多大容量的样本, 才能使总体均值 μ 的置信水平为0.95的置信区间长度不大于 l 。

【思路】 已知 σ^2 , 求正态总体均值 μ 的置信区间。

解: 设抽取样本容量为 n , 由于总体方差 σ_0^2 已知, μ 的置信水平为 $1 - \alpha = 0.95$ 的置信区间为:

$$\left(\bar{X} - \frac{\sigma_0}{\sqrt{n}} z_{\alpha/2}, \bar{X} + \frac{\sigma_0}{\sqrt{n}} z_{\alpha/2} \right)$$

置信区间的长度为: $2 \times \frac{\sigma_0}{\sqrt{n}} z_{\alpha/2} \leq l$ 。

查标准正态分布分布表得 $z_{\alpha/2} = z_{0.025} = \Phi(0.975) = 1.96$

解得 $n \geq \left(\frac{2\sigma_0^2 z_{\alpha/2}}{l} \right)^2 = \frac{15.37\sigma_0^2}{l^2}$, 样本容量至少为 $\left\lceil \frac{15.37\sigma_0^2}{l^2} \right\rceil$ 。

例6. 为了评价灯泡寿命状况，抽取10个灯泡测试。求得样品均值 $\bar{X} = 1500$ 小时，样品标准差 $S = 20$ 小时。若灯泡寿命服从正态分布 $N(\mu, \sigma_0^2)$ ，试求：

- (1) 该批灯泡寿命均值 μ 及标准差 σ_0 的置信区间（置信度0.95）。
- (2) 该批灯泡寿命均值 μ 的置信度0.95的单侧置信下限。

【思路】（1）未知 σ_0^2 ，求正态总体均值 μ 的置信区间；未知 μ ，求标准差 σ_0 的置信区间。

（2）未知 σ_0^2 ，求正态总体均值 μ 的单侧置信下限。

解: (1) 总体方差 σ_0^2 未知, μ 的置信度为 $1 - \alpha = 0.95$ 的置信区间为:

$$\text{查表 } t_{\alpha/2}(9) = t_{0.025}(9) = 2.2622$$

$$\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right) = \left(1500 \pm \frac{20}{\sqrt{10}} t_{0.025}(9) \right) \\ = (1485.7, 1514.3)$$

μ 未知, σ_0^2 的置信水平为 $1 - \alpha = 0.95$ 的置信区间为:

$$\left(\sqrt{\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}}, \sqrt{\frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)}} \right) \\ = \left(\sqrt{\frac{9 \times 20^2}{19.022}}, \sqrt{\frac{9 \times 20^2}{2.7}} \right) = (13.77, 36.51)$$

查表

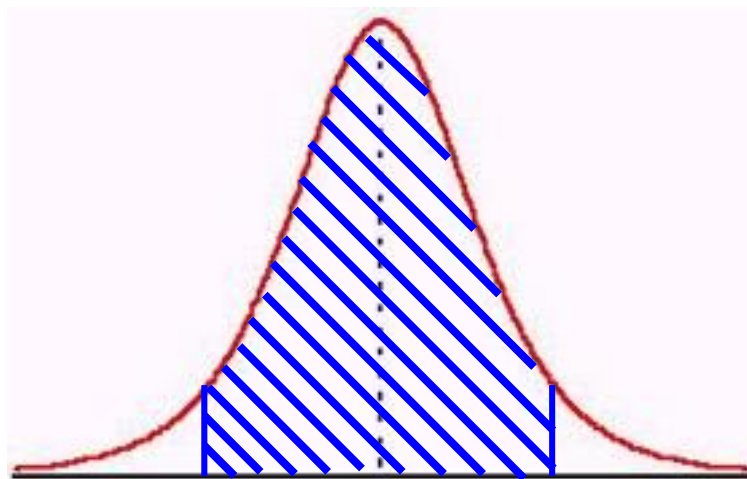
$$\chi_{\alpha/2}^2(9) = \chi_{0.025}^2(9) \\ = 19.022$$

$$\chi_{1-\alpha/2}^2(9) = \chi_{0.975}^2(9) \\ = 2.7$$

解：(2) 总体方差 σ_0^2 未知， μ 的置信度为 $1 - \alpha = 0.95$ 的单侧置信下限为：

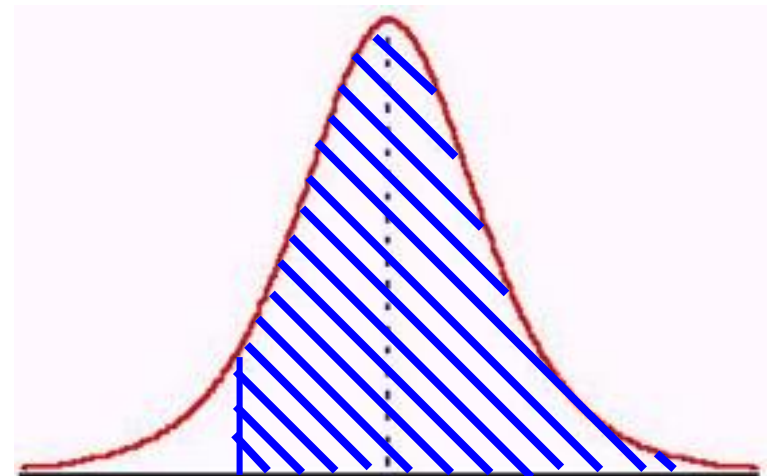
$$\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha}(n-1) = 1500 - \frac{20}{\sqrt{10}} t_{0.05}(9) = 1488.41$$

其中，查表得 $t_{\alpha}(9) = t_{0.05}(9) = 1.8331$ 。



1485.7

1514.3



1488.41

例7. 某工厂利用两条自动化流水线灌装番茄酱，分别从两条流水线上抽取样本： X_1, X_2, \dots, X_{12} 和 Y_1, Y_2, \dots, Y_{17} 。算出 $\bar{X} = 10.6\text{g}$ ， $\bar{Y} = 9.5\text{g}$ ， $S_1^2 = 2.4$ ， $S_2^2 = 4.7$ 。假设这两条流水线灌装的番茄酱重量均服从正态分布，且互相独立，其均值分别为 μ_1, μ_2 。

(1) 若两总体方差 $\sigma_1^2 = \sigma_2^2$ ，求 $\mu_1 - \mu_2$ 置信度为0.95的置信区间。

(2) 求 σ_1^2/σ_2^2 置信度为0.95的置信区间。

【思路】 (1) 两个正态总体方差未知，但 $\sigma_1^2 = \sigma_2^2$ ，求正态总体均值差 $\mu_1 - \mu_2$ 的置信区间。

(2) 正态总体均值 μ_1, μ_2 未知，求方差比 σ_1^2/σ_2^2 的置信区间。

解：(1) 两个正态总体方差未知，但 $\sigma_1^2 = \sigma_2^2$ ，总体均值差 $\mu_1 - \mu_2$ 的置信度为 $1 - \alpha = 0.95$ 的置信区间是

$$\left(\bar{X} - \bar{Y} \pm t_{\alpha/2}(n_1 + n_2 - 2) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

$$= \left(10.6 - 9.5 \pm t_{0.025}(27) S_w \sqrt{\frac{1}{12} + \frac{1}{17}} \right) = (-0.401, 2.601)$$

$$\text{其中 } S_w = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{11 \times 2.4 + 16 \times 4.7}{12 + 17 - 2} = 3.76,$$

查表得 $t_{0.025}(27) = 2.0518$ 。

所以均值差置信度0.95的置信区间为 $(-0.401, 2.601)$ 。

解: (2) 两个正态总体均值 μ_1, μ_2 未知, 方差比 σ_1^2/σ_2^2 的置信度为 $1 - \alpha = 0.95$ 的置信区间是

$$\begin{aligned} & \left(\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(n_1 - 1, n_2 - 1)}, \frac{S_1^2}{S_2^2} \frac{1}{F_{1-\alpha/2}(n_1 - 1, n_2 - 1)} \right) \\ &= \left(\frac{2.4}{4.7} \frac{1}{F_{0.025}(11, 16)}, \frac{2.4}{4.7} \frac{1}{F_{0.975}(11, 16)} \right) \\ &= \left(\frac{2.4}{4.7} \frac{1}{F_{0.025}(11, 16)}, \frac{2.4}{4.7} \frac{1}{F_{0.975}(11, 16)} \right) = (0.1737, 1.7004) \end{aligned}$$

其中 $\frac{1}{F_{0.975}(11, 16)} = F_{0.025}(16, 11) = 3.33$, $F_{0.025}(11, 16) = 2.94$

所以方差比置信度0.95的置信区间为(0.1737, 1.7004)。

Excel 实验

- 使用函数计算样本均值、样本方差等统计量
- 绘制直方图和箱线图
- 区间估计

使用函数计算样本的统计量

■ 常用函数及其功能

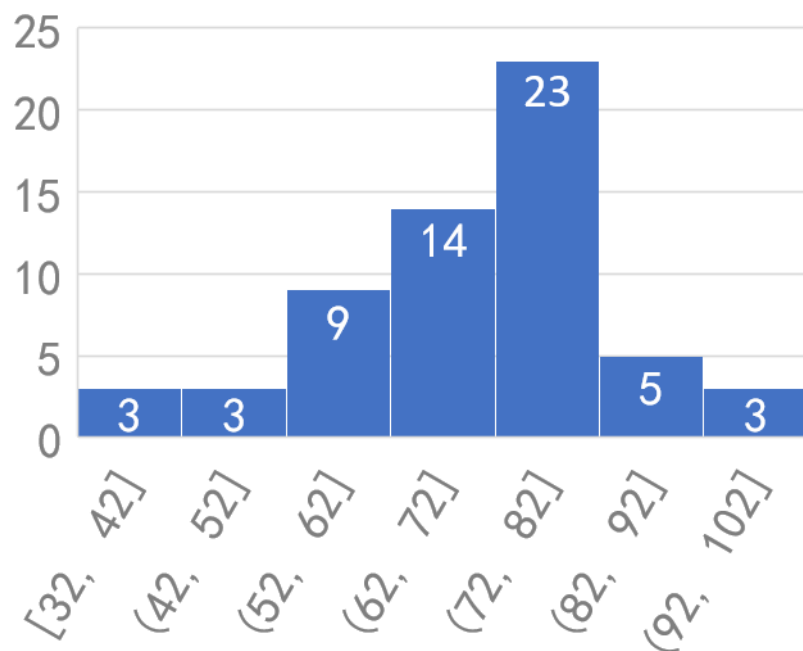
函数	功能	函数	功能	函数	功能
COUNT	计数	SQRT	平方根	NORMSINV	z值计算
MAX	最大值	AVERAGE	平均值	CHIINV	卡方值
MIN	最小值	VAR	方差	TINV	T值
SUM	求和	STDEV	标准差	FINV	F值
SUMSQ	平方和	MEDIAN	中位数		
ABS	绝对值	QUARTILE	四分位数		

■ 自定义函数

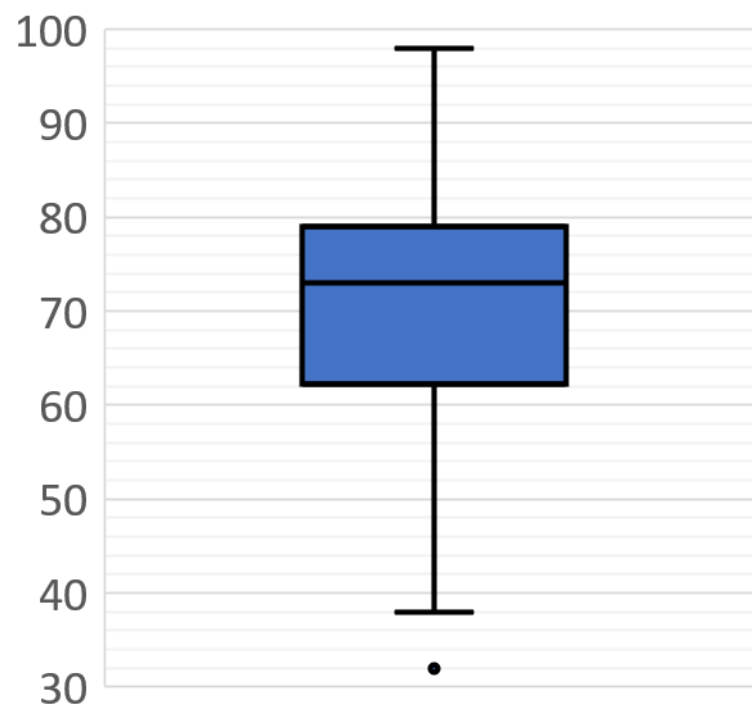
■ 示例操作

绘制直方图和箱线图

期中考试成绩直方图



期中考试成绩箱线图



使用Excel做区间估计

例：已知某种灯泡的寿命服从正态分布，现从一批灯泡中抽取16只，测得其寿命（单位：小时）如下所示：

1510 1450 1480 1460 1520 1480 1490 1460

1480 1510 1530 1470 1500 1520 1510 1470

求该灯泡平均使用寿命置信度为90%，95%，99%的置信区间。

使用Excel做区间估计

分析：总体方差未知，关于总体均值 μ 的置信度为 $1 - \alpha$ 的置信区间为

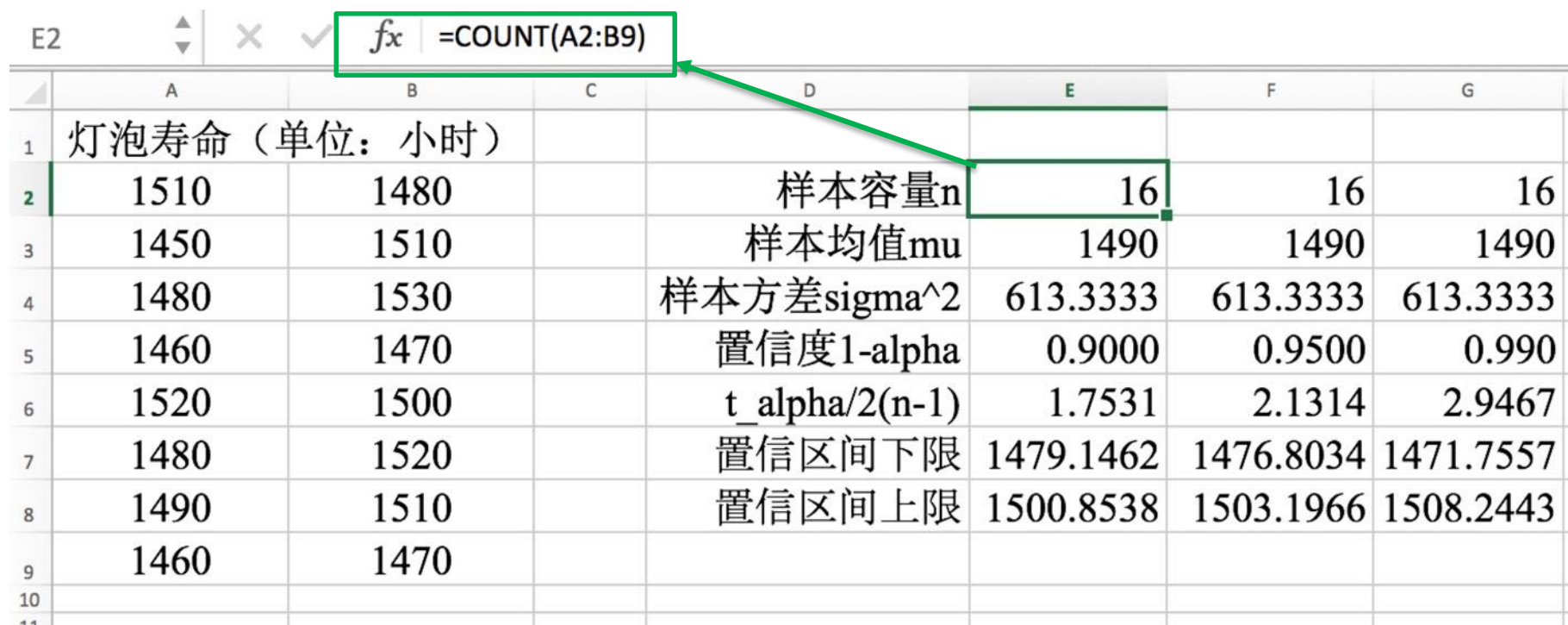
$$\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right)$$

其中， \bar{X} 为样本均值， S 为样本标准差， n 为样本容量。

- 在Excel里输入原始数据，然后编写置信区间求解公式求解。

使用Excel做区间估计

Excel操作截图（样本容量）



	A	B	C	D	E	F	G
1	灯泡寿命（单位：小时）						
2	1510	1480		样本容量n	16	16	16
3	1450	1510		样本均值mu	1490	1490	1490
4	1480	1530		样本方差sigma^2	613.3333	613.3333	613.3333
5	1460	1470		置信度1-alpha	0.9000	0.9500	0.990
6	1520	1500		t_alpha/2(n-1)	1.7531	2.1314	2.9467
7	1480	1520		置信区间下限	1479.1462	1476.8034	1471.7557
8	1490	1510		置信区间上限	1500.8538	1503.1966	1508.2443
9	1460	1470					
10							
11							

使用Excel做区间估计

Excel操作截图（样本均值）

E3

fx

=AVERAGE(A2:B9)

	A	B	C	D	E	F	G
1	灯泡寿命（单位：小时）						
2	1510	1480		样本容量n	16	16	16
3	1450	1510		样本均值mu	1490	1490	1490
4	1480	1530		样本方差sigma^2	613.3333	613.3333	613.3333
5	1460	1470		置信度1-alpha	0.9000	0.9500	0.990
6	1520	1500		t_alpha/2(n-1)	1.7531	2.1314	2.9467
7	1480	1520		置信区间下限	1479.1462	1476.8034	1471.7557
8	1490	1510		置信区间上限	1500.8538	1503.1966	1508.2443
9	1460	1470					
10							
11							

使用Excel做区间估计

Excel操作截图（样本方差）

E4			f_x =VAR(A2:B9)				
	A	B	C	D	E	F	G
1	灯泡寿命（单位：小时）						
2	1510	1480		样本容量n	16	16	16
3	1450	1510		样本均值mu	1490	1490	1490
4	1480	1530		样本方差sigma^2	613.3333	613.3333	613.3333
5	1460	1470		置信度1-alpha	0.9000	0.9500	0.990
6	1520	1500		t_alpha/2(n-1)	1.7531	2.1314	2.9467
7	1480	1520		置信区间下限	1479.1462	1476.8034	1471.7557
8	1490	1510		置信区间上限	1500.8538	1503.1966	1508.2443
9	1460	1470					
10							
11							

使用Excel做区间估计

Excel操作截图（T值）

$TINV(\alpha, n)$ 返回的是 $t_{\alpha/2}(n)$

E6			f_x	=TINV(1-E5,E2-1)			
	A	B	C	D	E	F	G
1	灯泡寿命（单位：小时）						
2	1510	1480		样本容量n	16	16	16
3	1450	1510		样本均值mu	1490	1490	1490
4	1480	1530		样本方差sigma^2	613.3333	613.3333	613.3333
5	1460	1470		置信度1-alpha	0.9000	0.9500	0.990
6	1520	1500		t_alpha/2(n-1)	1.7531	2.1314	2.9467
7	1480	1520		置信区间下限	1479.1462	1476.8034	1471.7557
8	1490	1510		置信区间上限	1500.8538	1503.1966	1508.2443
9	1460	1470					
10							

使用Excel做区间估计

Excel操作截图（置信区间） $\left(\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) \right)$

E7							
	A	B	C	D	E	F	G
1	灯泡寿命（单位：小时）						
2	1510	1480		样本容量n	16	16	16
3	1450	1510		样本均值mu	1490	1490	1490
4	1480	1530		样本方差sigma^2	613.3333	613.3333	613.3333
5	1460	1470		置信度1-alpha	0.9000	0.9500	0.990
6	1520	1500		t_alpha/2(n-1)	1.7531	2.1314	2.9467
7	1480	1520		置信区间下限	1479.1462	1476.8034	1471.7557
8	1490	1510		置信区间上限	1500.8538	1503.1966	1508.2443
9	1460	1470					
10							
11							