Probability and Statistics
Tutorial 14

Siyi Wang

Southern University of Science and Technology

11951002@mail.sustech.edu.cn

December 21, 2020

Outline

Review

- One-sided hypotheses:

$$H_0: \theta \le \theta_0 \quad \text{versus} \quad H_1: \theta > \theta_0$$
$$H_0: \theta \ge \theta_0 \quad \text{versus} \quad H_1: \theta < \theta_0$$

- Two-sided hypotheses:

$$H_0: \theta = \theta_0 \quad \text{versus} \quad H_1: \theta \ne \theta_0$$

Review

- Consider a subset R of the sample space. If $X \in R$, then H_0 will be rejected. Then R is called the rejection region or critical region.

Review

Usual Steps of Hypothesis Testing:

1. Assume the hypothesis H_0 as true
2. Take a sample $\Rightarrow$ Compute the test statistic $W(X_1, \ldots, X_n)$ and its sampling distribution
3. Form a critical region by probability theory: if the realized test statistic is *not consistent with* or *too extreme under* the hypothesis, then reject it.

Review

- Type I error: H_0 true, reject
- Type II error: H_0 false, fail to reject

State|Decision	Fail to reject H_0	Reject H_0
H_0 True	Correct	Type I error
H_1 True	Type II error	Correct

- We want both errors to be small. However, decreasing type I error means a bigger R while decreasing type II error means a smaller R. So the two errors cannot be minimized simultaneously.
- In an extreme case, we always reject H_0, so R is the whole sample space. No type II error but huge type I error.

Significance Level

- If $H_0 : \theta \in \Theta_0$ is true, then $\alpha(\theta) = \mathbb{P}_\theta(X \in R)$ is the probability of a type I error.
  - If H_0 is simple, then it is a single probability.
  - If H_0 is composite, then there is a set of probabilities.

SIGNIFICANCE LEVEL

The maximum probability of a type I error, over the set of distributions specified by H_0,

$$\alpha = \sup_{\theta \in \Theta_0} \alpha(\theta),$$

is the significance level of the test.

- Usually, we target $\alpha \le 0.1$, 0.05 or even 0.01.

Review

处理假设检验问题的一般方法和步骤

- 根据实际问题，提出原假设 H_0 及备择假设 H_1
- 求出未知参数的较好的点估计
- 依据点估计构造一个检验统计量，然后分析当 H_0 成立时，该统计量有什么"趋势"，这个"趋势"就给出了 H_0 的拒绝域形式，即 H_0 的拒绝域形式由 H_1 确定
- 对于给定的显著性水平 α，按控制I类风险的检验原则，确定 H_0 的拒绝域
- 抽样，判断样本观察值是否落在拒绝域内，从而作出"拒绝"或"接受" H_0 的决策

Homework

1. 设某种清漆的9个样品，其干燥时间（单位：h）分别为6.0, 5.7, 5.8, 6.5, 7.0, 6.3, 5.6, 6.1, 5.0. 设干燥时间 X~N(μ,σ^2).
在下面两种情况下：（1）σ=0.6(h)；（2）σ未知，求μ的置信水平为0.95的置信区间.

Solution

1\. Solution $\bar{x}=6$, $n=9$, $s=0.574$, $\alpha=0.05$

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

μ的0.95置信区间为 $\left(\bar{X}-\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}, \bar{X}+\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}\right)$

i.e. $(5.608, 6.392)$

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

μ的0.95置信区间为 $\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(8), \bar{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(8)\right)$

i.e. $(5.559, 6.441)$.

Homework

2、有一大批糖果.现从中随机地取 16 袋，称得重量(以克计)如下：

506	508	499	503	504	510	497	512
514	505	493	496	506	502	509	496

设袋装糖果的重量近似地服从正态分布,试求：

（1）总体均值 μ 的置信水平为0.95的置信区间.

（2）总体标准差 σ 的置信水平为0.95的置信区间.

Homework

Solution

2. Solution. $\overline{X} = 503.75$, $S^2 = 38.47$, $S \approx 6.2$, $n = 16$

(1) $\dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$ $\quad\quad \alpha = 0.05$.

μ 的 0.95 置信区间: $\left(\overline{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1),\ \overline{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}(n-1)\right)$

i.e. $(500.45,\ 507.05)$.

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

σ^2 的 0.95 置信区间: $\left(\dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\ \dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)$ $\quad\quad (\alpha = 0.05)$

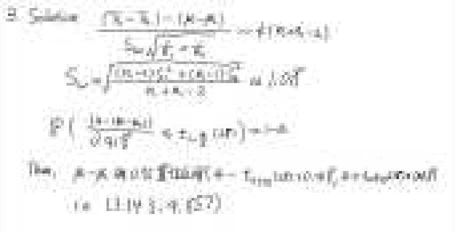
i.e. $(20.99,\ 92.18)$

Homework

3、为比较 I，II 两种型号步枪子弹的枪口速度，随机地取 I 型子弹 10 发，得到枪口速度的平均值为 $\overline{x_1}=500(m/s)$，方差 $s_1^2=1.10(m/s)$，随机地取 II 型子弹 20 发，得到枪口速度的平均值为 $\overline{x_2}=496(m/s)$，方差 $s_2^2=1.20(m/s)$. 假设两总体都可认为近似地服从正态分布.且生产过程可认为方差相等.求两总体均值差 $\mu_1-\mu_2$ 的置信水平为 0.95 的置信区间.

Homework

Solution



4、研究由机器 A 和机器 B 生产的钢管的内径，随机地抽取机器 A 生产的钢管18只，测得样本方差 $s_1^2=0.34(mm^2)$；随机地取机器 B 生产的钢管13只，测得样本方差 $s_2^2=0.29(mm^2)$. 设两样本相互独立，且设由机器 A 和机器 B 生产的钢管的内径分别服从正态分布 $N\left(\mu_1,\sigma_1^2\right)$, $N\left(\mu_2,\sigma_2^2\right)$,这里 μ_i,σ_i^2 $(i=1,2)$ 均未知. 试求方差比 σ_1^2/σ_2^2 的置信水平为0.90 的置信区间.

Homework

Solution

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

$$P\left(F_{\frac{\alpha}{2}}([illegible]) < \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} < F_{1-\frac{\alpha}{2}}([illegible])\right) = 1-\alpha$$

Then $\frac{\sigma_1^2}{\sigma_2^2}$ [illegible]

$$\left(\frac{S_1^2/S_2^2}{F_{1-\frac{\alpha}{2}}([illegible])}, \frac{S_1^2/S_2^2}{F_{\frac{\alpha}{2}}([illegible])}\right) \quad ([illegible])$$

[illegible] $([illegible], [illegible])$.

Homework

1.某电器元件平均电阻值一直保持 2.64Ω,今测得采用新工艺生产 36 个元件的平均阻值为 2.61Ω.假定在正常条件下,电阻值服从正态分布.而且新工艺不改变电阻的标准差.已知改变工艺前的标准偏差为 0.06Ω,问新工艺对产品的电阻值是否有显著性影响 $(\alpha=0.01)$?

Homework

Solution

Homework

2.某厂生产的某种钢索的断裂强度服从正态 $N(\mu,\sigma^2)$,其中 $\sigma=40$ (kg/cm^2),现在一批这种钢索的容量为 9 的一个样本测得断裂强度平均值为 $\bar{X}$,与以往正常生产的 μ 相比,$\bar{X}$ 较 μ 大 20(kg/cm^2).设总体方差不变,问在 $\alpha=0.01$ 能否认为这批钢索质量显著提高?

Homework

Solution

6\. Solution $n=9$, $\sigma=?$, $\alpha=0.01$, $\bar{x}-\mu_0=20$

$H_0: \mu=\mu_0$ vs $H_1: \mu\neq\mu_0$

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

H_0 的拒绝域 $\left\{ \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha} \right\}$

i.e. $\left\{ \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > 2.33 \right\}$

Since $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = 1.5 < 2.33$

Thus we fail to reject H_0

Homework

3. 某种零件的尺寸方差为$\sigma^2=1.21$，对一批这类零件检查6件得尺寸数据(mm)：32.56, 29.66, 31.64, 30.00, 21.87, 31.03. 设零件尺寸服从正态分布，问这批零件的平均尺寸能否认为是32.50mm ($\alpha=0.05$)?

Solution

7. Solution: $\bar{x} = 29.46$, $\mu_0 = $ [illegible], $\sigma = $ [illegible], $\alpha = 0.05$

$H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$

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

H_0的拒绝域为 $\left\{ \frac{|\bar{x} - \mu_0|}{\sigma / \sqrt{n}} > z_{1-\frac{\alpha}{2}} \right\}$

$$z_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$$

Since $\frac{|\bar{x} - \mu_0|}{\sigma / \sqrt{n}} = \frac{[illegible]}{0.4[illegible]} = 6.17 > 1.96$.

thus we reject H_0.

Ans. [illegible]

Supplement Exercises

1. 设 $x_1,\cdots,x_n$ 是来自 $N(\mu,1)$ 的样本，考虑如下假设检验问题

$$H_0:\mu=2 \quad vs \quad H_1:\mu=3,$$

若检验由拒绝域为 $W=\{\bar{x}\geq 2.6\}$ 确定.

(1) 当 $n=20$ 时求检验犯两类错误的概率；

(2) 如果要使得检验犯第二类错误的概率 $\beta\leq 0.01$，n 最小应取多少？

(3) 证明：当 $n\to\infty$ 时，$\alpha\to 0$，$\beta\to 0$.

Solution

解 (1) 由定义知，犯第一类错误的概率为

$$\alpha = P(\bar{x} \geq 2.6 \mid H_0) = P\left(\frac{\bar{x}-2}{\sqrt{1/20}} \geq \frac{2.6-2}{\sqrt{1/20}}\right) = 1-\Phi(2.68) = 0.0037,$$

这是因为在 H_0 成立下，$\bar{x} \sim N(2, 1/20)$. 而犯第二类错误的概率为

$$\beta = P(\bar{x} < 2.6 \mid H_1) = P\left(\frac{\bar{x}-3}{\sqrt{1/20}} < \frac{2.6-3}{\sqrt{1/20}}\right) = \Phi(-1.79)$$

$$= 1-\Phi(1.79) = 0.0367,$$

这是因为在 H_1 成立下，$\bar{x} \sim N(3, 1/20)$.

(2) 若使犯第二类错误的概率满足

$$\beta = P(\bar{x} < 2.6 \mid H_1) = P\left(\frac{\bar{x}-3}{\sqrt{1/n}} < \frac{2.6-3}{\sqrt{1/n}}\right) \leq 0.01,$$

即 $1-\Phi\left(\frac{0.4}{\sqrt{1/n}}\right) \leq 0.01$，或 $\Phi(0.4\sqrt{n}) \geq 0.99$，查表得：$0.4\sqrt{n} \geq 2.33$，由此给出 $n \geq 33.93$，因而 n 最小应取 34，才能使检验犯第二类错误的概率 $\beta \leq 0.01$.

Solution

(3) 在样本量为 n 时，检验犯第一类错误的概率为

$$\alpha = P(\bar{x} \geq 2.6 \mid H_0) = P\left(\frac{\bar{x}-2}{\sqrt{1/n}} \geq \frac{2.6-2}{\sqrt{1/n}}\right) = 1 - \Phi(0.6\sqrt{n}),$$

当 $n \to \infty$ 时，$\Phi(0.6\sqrt{n}) \to 1$，即 $\alpha \to 0$.

检验犯第二类错误的概率为

$$\beta = P(\bar{x} < 2.6 \mid H_1) = P\left(\frac{\bar{x}-2}{\sqrt{1/n}} < \frac{2.6-3}{\sqrt{1/n}}\right) = \Phi(-0.4\sqrt{n}) = 1 - \Phi(0.4\sqrt{n})$$

当 $n \to \infty$ 时，$\Phi(0.4\sqrt{n}) \to 1$，即 $\beta \to 0$.

注：从这个例子可以看出，要使得 α 与 β 都趋于0，必须 $n \to +\infty$ 才可实现，这一结论在一般场合仍成立，即要使得 α 与 β 同时很小，必须样本量 n 很大. 由于样本量 n 很大在实际中常常是不可行的，故一般情况下人们不应要求 α 与 β 同时很小.

4. 设总体为均匀分布 $U(0,\theta)$, $x_1,\cdots,x_n$ 是样本，考虑检验问题

$$H_0:\theta \geqslant 3 \quad vs \quad H_1:\theta < 3,$$

拒绝域取为 $W=\{x_{(n)} \leqslant 2.5\}$，求检验犯第一类错误的最大值 α. 若要使得该最大值 α 不超过 0.05, n 至少应取多大?

Supplement Exercises

Solution

解 均匀分布 $U(0,\theta)$ 的最大次序统计量 $x_{(n)}$ 的密度函数为

$$f_n(x)=\begin{cases}\dfrac{nx^{n-1}}{\theta^n}, & 0<x<\theta,\\ 0, & \text{其他。}\end{cases}$$

因而检验犯第一类错误的概率为

$$\alpha(\theta)=P(x_{(n)}\leqslant 2.5\mid H_0)=\int_0^{2.5} n\frac{x^{n-1}}{\theta^n}\mathrm{d}x=\left(\frac{2.5}{\theta}\right)^n,$$

它是 θ 的严格单调递减函数，故其最大值在 $\theta=3$ 处达到，即

$$\alpha=\alpha(3)=\left(\frac{2.5}{3}\right)^n.$$

若要使得 $\alpha(3)\leqslant 0.05$，则要求 $n\ln(2.5/3)\leqslant \ln 0.05$，这给出 $n\geqslant 16.43$，即 n 至少为 17.

Supplement Exercises

5. 在假设检验问题中，若检验结果是接受原假设，则检验可能犯哪一类错误？若检验结果是拒绝原假设，则又有可能犯哪一类错误？

Thank you!