17. 糖厂 99.3 98.7 100.5 ⋯ 102.1 100.5

额定100kg. 是否有系统偏差? $\alpha=0.05$

转为统计问题即 $H_0: \mu=\mu_0 \longleftrightarrow H_a: \mu\neq\mu_0$

检验统计量 $T=\dfrac{\sqrt{n}(\overline{X}-\mu_0)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}}$ $|T|>C$ 拒

$\int_C^{\infty} t(x;n-1)\,dx=\frac{\alpha}{2}$ $\overline{X}=99.9777$ $S^2=1.469444$

$T=\dfrac{\sqrt{9}\cdot(99.9777-100)}{\sqrt{1.469444}}=-0.054996$ $C=2.306$

不认为有偏差

18. 心跳 72 10个患者 54, 67, ⋯ 65, 69 均值异常? $\alpha=0.05$

统计问题: $H_0: \mu=72 \longleftrightarrow H_a: \mu\neq 72$

同上 $T=\dfrac{\sqrt{n}(\overline{X}-\mu_0)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}}$ $|T|>C$ 拒

$\overline{X}=67.4$ $S^2=35.15555$

$T=\dfrac{\sqrt{10}\cdot(67.4-72)}{\sqrt{35.1555}}=-2.4533575$ $C=2.262$

应该认为异常

20. 新法: 22 14 17 13 21 16 15 16 19 18

以往 $\mu_0 = 19$

统计问题: $H_0: \mu \geq \mu_0 \longleftrightarrow H_a: \mu < \mu_0$

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sqrt{\frac{1}{n-1}\sum(X_i - \overline{X})^2}}$$

否定域 $T < C$

其中 $\int_{-\infty}^{C} t(x; n-1)\,dx = \alpha$

$\overline{X} = 17.1 \quad S^2 = 8.5444$

$$T = \frac{\sqrt{10}\cdot(17.1 - 19)}{\sqrt{8.5444}} = -2.0554779$$

$C = -1.833$

应该认为新法效果好

21. 导线 $n=9 \quad S = 0.007 \quad \alpha = 0.05 \quad \sigma_0 = 0.005$ 能否认为 σ 显著大?

统计问题: $H_0: \sigma^2 \leq \sigma_0^2 \longleftrightarrow H_a: \sigma^2 > \sigma_0^2$

$$\chi^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 = (n-1)S^2 = 0.000392$$

否定域 $\chi^2 > C\sigma_0^2$ 其中 $\int_{C}^{\infty} \chi^2(x; n-1)\,dx = \alpha$

$C = 15.5073 \qquad C\sigma_0^2 = 0.00038768$

故能认为 σ 显著大

22.

机器1: 6.2, 5.7, ... 5.8, 6.0

机器2: 5.6, 5.9, ... 5.7, 5.5

$\alpha = 0.05$ 下 加工精度显著差异?

统计问题为 $H_0: \sigma_1^2 = \sigma_2^2 \longleftrightarrow H_a: \sigma_1^2 \neq \sigma_2^2$

检验量 $F = \frac{\sum(X_i - \bar{X})^2}{n_1 - 1} \Big/ \frac{\sum(Y_i - \bar{Y})^2}{n_2 - 1}$ 拒绝域 $F < C_1$ or $F > C_2$

$$\int_{-\infty}^{C_1} f(x; n_1-1, n_2-1)\,dx = \int_{C_2}^{\infty} f(x; n_1-1, n_2-1)\,dx = \frac{\alpha}{2}$$

$S_1^2 = 0.064$ $S_2^2 = 0.2444$ $F = 2.61818$

$C_1 = 0.2394$ $C_2 = 4.2951$

~~不~~认为 加工精度有差异

(*) 24. 10个患者, 甲,乙药 睡眠 $(1.9, 0.7), \cdots (3.4, 2.0)$

疗效有显著差异? $\alpha = 0.05$

统计问题 $X_1 \cdots X_{n_1} \sim N(\mu_1, \sigma_1^2)$ $Y_1 \cdots Y_{n_2} \sim N(\mu_2, \sigma_2^2)$

$H_0: \mu_1 = \mu_2 \longleftrightarrow H_a: \mu_1 \neq \mu_2$

检验量 (成对数据)

$T = \frac{\bar{z}}{\sqrt{S^2/n}}$ $z_i = x_i - y_i$ $|T| > C$ 拒绝

$$\int_C^{\infty} t(x, n-1)\,dx = \frac{\alpha}{2}.$$

$\bar{z} = 1.58$ $S^2 = 1.51288$ $T = 4.06214$

$C = 2.262$

故应当认为有显著差别

25. 上题以为2组输入，没有成对关系

检验 $H_0: \mu_1 = \mu \longleftrightarrow H_a: \mu_1 \neq \mu_2$

$|T| > c$ 时拒绝

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

其中 $\int_c^{\infty} t(x; m)\, dx = \frac{\alpha}{2}$

$\overline{X} = 2.33$ $\overline{Y} = 0.75$ $S_1^2 = 4.009$ $S_2^2 = 3.20055$ $n_1 = n_2 = 10$

$c = 2.1009$

$T = 1.8608$

$$\left(m^* = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2} = 17.76,\ m = 18 \right)$$

故认为没有显著差异

(*) 26. $X \sim N(\mu_1, \sigma_1^2)$ $Y \sim N(\mu_2, \sigma_2^2)$ $X_1 \cdots X_n \sim X$ $Y_1 \cdots Y_m \sim Y$

用广义似然法求 $H_0: \mu_1 = \mu_2 \longleftrightarrow H_a: \mu_1 \neq \mu_2$ （已知 $\sigma_1^2 = \sigma_2^2$）否定域

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

$$L(\hat{\Theta}_0) = \sup_{\theta \in \Theta_0} L(x, \theta)$$

$$L(\hat{\Theta}) = \sup_{\theta \in \Theta} L(x, \theta)$$

$$L(x, \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(X_i - \mu_1)^2}{2\sigma^2}} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(Y_i - \mu_2)^2}{2\sigma^2}}$$

$$= \frac{1}{(\sqrt{2\pi})^{n+m} \sigma^{n+m}} e^{-\frac{\sum (X_i - \mu_1)^2 + \sum (Y_i - \mu_2)^2}{2\sigma^2}}$$

$L(\hat{\Theta}_0)$ 最优参数：$\mu_1 = \mu_2 = \mu^* = \frac{\sum X_i + \sum Y_i}{n+m} = \frac{n\overline{X} + m\overline{Y}}{n+m}$，$\sigma^2 = \frac{\sum (X_i - \mu^*)^2 + \sum (Y_i - \mu^*)^2}{n+m} = \frac{nS_1^2 + mS_2^2 + \frac{mn}{m+n}(\overline{X} - \overline{Y})^2}{n+m}$

$L(\hat{\Theta})$ 最优参数：$\mu_1 = \overline{X}$ $\mu_2 = \overline{Y}$ $\sigma^2 = \frac{\sum (X_i - \overline{X})^2 + \sum (Y_i - \overline{Y})^2}{n+m} = \frac{nS_1^2 + mS_2^2}{n+m}$

$$\Lambda = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} = \left(\frac{nS_1^2 + mS_2^2}{nS_1^2 + mS_2^2 + \frac{mn}{m+n}(\bar{X}-\bar{Y})^2} \right)^{\frac{n+m}{2}}$$

故可看出 检验统计量

记 $S^2 = \frac{1}{n_1+n_2-2}(nS_1^2 + mS_2^2) = \frac{1}{n_1+n_2-2}\left(\sum(X_i-\bar{X})^2 + \sum(Y_i-\bar{Y})^2\right)$

$$T = \frac{\bar{X}-\bar{Y}}{S\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

$$\Lambda = \left(\frac{1}{1+\frac{1}{n_1+n_2-2}T^2} \right)^{\frac{n+m}{2}}$$

故否定域 $|T| > C$

其中 $\int_C^{\infty} t(x; n_1+n_2-2)\,dx = \frac{\alpha}{2}$

#