

Stat 535 HW 1

Jason Rich

January 25, 2015

2.5

one sample Z Test of $\mu = 30$ vs $\text{not} = 30$

$\mu = 30$ $s = 1.2$ $n = 16$ $\bar{x} = 31.200$ $SE_{\bar{x}} = 0.300$

95% CI: (30.6120, 31.7880)

a.)

$Z_{.001} = 4$

$P_{.001} = 5.8995355$

b.)

This is a two-sided test

c.)

99% CI is $\bar{x} + / - z_{\alpha/2} s \sqrt{n}$

where the $+/-$ greater than the mean, and less than the mean respectively

99% CI = (31.1985 , 31.2015)

d.)

P for $\mu > 30$

$P_{.002} = 0.0013499$

2.10

a.)

The standard error is calculated using algebra as and the t statistic formula

SE: $-1.88 = \frac{11.5}{s(\sqrt{25})}$, where s = the standard deviation;

$$s(-1.88)(\sqrt{25}) = 11.5$$

$$s = -1.2234043$$

The Standard error is: $s\sqrt{25} = -6.11702$

b.)

two-sided test

c.)

the t-value = 1.6607141

The rejection criteria is $|t\text{-stat}| > t\text{-value}(\alpha/2)$

Thus, at the $\alpha = .05$ $|t\text{-stat}| = -1.88 < t_{\alpha/2} = 1.6607141$, and I would reject the H_0 .

d.)

$$CV = 2.0638986$$

$$95\% \text{ CI} = (-1.1249088, 24.1249088)$$

2.14

a.)

Degress of freedom for the *t-test* is N-1. DF = 11

b.)

The *t-stat* test statistic = $(\bar{x} - \mu) / \text{se}(\bar{x})$

$$T = 2.0291667$$

$$\text{standard deviation} = 0.0969948$$

$$95\% \text{ lower bound} = 0.0783827$$

2.19

Since $x \sim N(\mu, 9)$, and the 95% two-sided CI on μ is

$$CI = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Know that our interval has a total length on 1.0, half is 0.5. Thus our CV is $z(\alpha/2) = z_{0.025} = 1.96$

$$\bar{x} - (1.96)(3)/\sqrt{n} = 0.5$$

$$\sqrt{n} = 11.76$$

$$n = 138.2976$$

2.21

At the limit, the t-distribution (CLT) become more and more normal. Since the sample size is small, major departs from normality will not effect the predictive ability of the data.

2.25

$$H_0: \mu_1 - \mu_2 = 10; H_1: \mu_1 - \mu_2 > 10$$

$$\bar{y}_1 = 162.5, \bar{y}_2 = 155.0$$

$$\sigma_1 = 1, \sigma_2 = 1$$

$$n_1 = 10,$$

$$n_2 = 12$$

$$z_0 = \frac{\bar{y}_1 - \bar{y}_2 - 10}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$z_0 = 162.5 - 155.0 - 10 / \sqrt{1^2/10 + 1^2/12} = -5.85$$

$$z_{0.01} = 2.225; \text{ do not reject } H_0$$

$$99\% \text{ CI} = 6.3974509 \mu_1 - \mu_2 8.6025491$$

2.28

a.)

Using $\alpha = 0.05$

$$H_0: \sigma_1 = \sigma_2 ; H_1: \sigma_1 \neq \sigma_2$$

$$F_{0.025,7,8} = 4.53$$

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{101.17}{94.73} = 1.0679827$$

Do not Reject the H_0

b.)

$$H_0: \mu_1 = \mu_2 ; H_1: \mu_1 \neq \mu_2$$

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

$$S_p^2 = \frac{(8-1)101.17 + (9-1)94.73}{8+9-2} = 97.74$$

$$S_p = 9.$$

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

$$t_0 = \frac{12.5 - 10.2}{9.89 \sqrt{\frac{1}{8} + \frac{1}{9}}} = 0.479$$

$$t_{0.05,15} = 1.753$$

2.31

a.)

$$95\% \text{ CL: } \frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{(1-\alpha/2), n-1}^2} \quad \frac{(20-1)0.8891}{32.852} \leq \sigma^2 \leq \frac{(20-1)0.8891}{8.907}$$

b.)

$$\sigma^2 = 1.0, \text{ using } \alpha = 0.05$$

$$H_0: \sigma^2 = 1$$

$$H_1: \sigma^2 \neq 1$$

$$\text{Test Statistic: } \chi^2 = \frac{SS}{\sigma_0^2} = 15.019$$

$$\chi_{0.025,19}^2 \text{ Value} = 32.852;$$

$$\chi_{0.975,19}^2 \text{ Value} = 8.907$$

The evidence does not support the claim that $\sigma^2 \neq 1$

c.)

The normality assumption, is argued by many, as the most important assumption when analyzing the σ^2 then when analyzing the mean. Small variations from normality may cause issues the both test of H_0 and CI's.

d.)

A check of the normal probability plot does not show any significant issue with normality

2.32

a.)

$$H_0: \mu_1 = \mu_2;$$

$$H_1: \mu_1 \neq \mu_2$$

An Analysis of the data reveals Two Sample t-test

data: cal_1 and cal_2

$$t = 0.43, df = 11, p\text{-value} = 0.674$$

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: -0.001024, 0.001524

sample estimates mean of cal_1 mean of cal_2

$$0.266250 \quad 0.266000$$

b.)

$$p\text{-value} = 0.674$$

c.)

$$\bar{d} - t_{\alpha/2, n-1} \frac{S_d}{\sqrt{n}} \leq \mu_1 - \mu_2 \leq \bar{d} + t_{\alpha/2, n-1} \frac{S_d}{\sqrt{n}}$$

$$0.00025 - 2.201 \frac{0.002}{\sqrt{12}} \leq \mu_1 - \mu_2 \leq 0.00025 + 2.201 \frac{0.002}{\sqrt{12}}$$

$$-0.00102 \leq \mu_1 - \mu_2 \leq 0.000152$$

2.37

a.)

Two Sample t-test

data: sol_1 and sol_2

$$t = -2.38, df = 14, p\text{-value} = 0.032$$

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: -0.83, 0.043

sample estimates mean of sol_1 mean of sol_2

9.925

10.362

b.)

95% CI: -0.83, 0.043

c.)

A review of the normal QQ-Plot reveals no issue with normality

question A

A linear combination of independent random variables

$$y_{ij} = \beta_0 + \beta_1 x_i + e_{ij}$$

If the random variables are iid $rvs N(\mu, \sigma^2)$ the rvs are distributed multivariate normal distribution (aka: multinormal distribution)

question B

CLT is the cornerstone of modern statistical analysis. It states, at the limit, as the sample size goes to infinity, the data becomes more and more normal (technically the data becomes more standard normal).

For the data to be considered “normal”, the sample size should be at least (no smaller than) 25 to 30 (depending on the data, and the analyst experience with data and the model)

question C

Since $x \sim N(\mu, \sigma^2)$, and the 95% two-sided CI on μ is

$$CI = \bar{x} \pm z(\alpha/2) * \sigma / \sqrt{n}$$

Know that our interval has a total length of 1.0, half is 0.5. Thus our CV is $z(\alpha/2) = z_{0.025} = 1.96$

$$\bar{x} \pm (1.96) \frac{\sigma^2}{\sqrt{n}} = 0.5$$

$\sqrt{n} = n$ which is the sample size required for the experiment

question D

Is the appropriate allocation of experimental materials and the order in which the individual runs of the experiment are to be performed are randomly determined. Furthermore, the statistical methods underlying the analysis of the experiment require iid “random” variable, and randomization usually validates this requirement!