M358K: November 20th, 2020.

t-procedures [cont'd].

Review:

Normal population distribution with both the mean parameter μ and the standard deviation parameter σ are unknown.

Let $X_1, X_2, \ldots, X_n$ be a random sample from this population dist'n.

Define: $\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ .... sample mean

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad \ldots \text{ sample variance}$$

S... sample standard deviation

Consider the following statistic:

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

If we want to construct a confidence interval @ the confidence level C, we have:

$$\boxed{\bar{X} - t^* \cdot \frac{S}{\sqrt{n}} < \mu < \bar{X} + t^* \frac{S}{\sqrt{n}}}$$

Example. [Ramachandran · Tsokos]

The following is a data set from a normal distribution:

7.2, 5.7, 4.9, 6.2, 8.5, 2.8.

Construct a 95% confidence interval for the population mean.

Method I.

$\bar{x}$ = mean(x) = 5.883333

s = sd(x) = 1.958996

$t^* = ?$

- df = 6−1 = 5
- lower tail probability associated w/ confidence level 0.95 : 0.975

t^* = qt(0.975, df = 5) = 2.570582

(3.827493, 7.939174)

Method II.

t.test(x) ....

$\blacksquare$

Hypothesis Tests for the μ.

- If the sample is large ($n \geq 30$), then use the z procedure w/ s substituted for σ.
- If the sample is small ($n < 30$), then ...

We test:

$$H_0: \mu = \mu_0 \quad \text{vs.} \quad H_a: \begin{cases} \mu < \mu_0 \\ \mu \neq \mu_0 \\ \mu > \mu_0 \end{cases}$$

Test statistic (under the null):

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(df = n-1)$$

t... observed value of the TS

With a significance level α, our rejection region:

$$RR = \begin{cases} t < -t_{\alpha,n-1} \\ |t| > t_{\alpha/2,n-1} \\ t > t_{\alpha,n-1} \end{cases}$$

where $t_{\alpha,n-1}$ is such that for $T \sim t(df=n-1)$ we have

$$\mathbb{P}[T > t_{\alpha,n-1}] = \alpha$$

Decision: If the observed value of the test statistic falls in the rejection region, then reject the null.
If the observed value of the test statistic falls outside of the rejection region, then fail to reject the null.

Example. [Ramachandran · Tsokos]

A manufacturer of fuses who claims that w/ a 20% overload the fuses will blow in less than 10 minutes on average. To test this claim, a random sample of 20 of these fuses was subjected to a 20% overload. The times it took them to blow had the mean of 10.4 minutes and the sample standard deviation of 1.6 minutes. We can assume that the data come from the normal dist'n.

Do the data support or refute the manufacturer's claim?

→: X ... the population dist'n, i.e., the reaction time

$X \sim \text{Normal}(\text{mean} = \mu, \text{sd} = \sigma)$

$H_0: \mu = \mu_0 = 10$ vs. $H_a: \mu > \mu_0 = 10$

$\bar{x} = 10.4$; $s = 1.6$; $n = 20$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{10.4 - 10}{1.6/\sqrt{20}} = 1.118$$

$df = n - 1 = 20 - 1 = 19$

Using my t-table, I conclude that
the p-value is between 10% and 15%.

p-value: 1-pt(1.118, df = 19) = 0.1387521