L9

Wednesday, October 17, 2018
9:05 PM

<<Lecture9_Students-1.pdf>>

Stat 324 – Introduction to Statistics for Engineers

LECTURE 10: T TESTS, CI/2-SIDED EQUIVALENCE, REVISITING POWER

5.4 - 5.7 OF OTT AND LONGNECKER.

Revisiting Paint Data

Recall our paint data from earlier notes:

A car manufacturer ... warehouse ... thousands of painted blocks ... 16 blocks are selected at random, and the paint thickness is measured ...:

1.29, 1.12, 0.88, 1.65, 1.48, 1.59, 1.04, 0.83, 1.76, 1.31, 0.88, 1.71, 1.83, 1.09, 1.62, 1.49

Suppose the specification says the mean thickness should be 1.50 mil.

We want to know whether the device is off this mark on average, so that the machine should be re-calibrated to correct its population mean thickness, μ.

How could we do a hypothesis test about the true mean point thickness? We don't know σ !

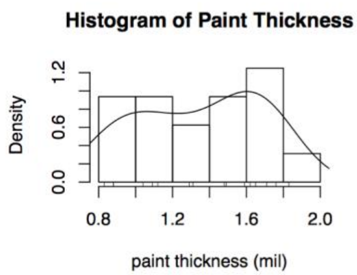
Revisiting Paint Data

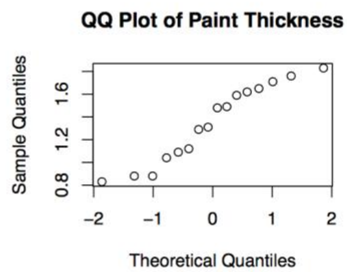
A. Define hypotheses: $H_o: \mu = 1.5$ vs $H_o: \mu \neq 1.5$

B. Check Assumptions: (1) Independent Observations?

(2) Observations from Normal Population or
sample size large enough that CLT kicks in?

The plots below tell us it is plausible the sample come from a Normal population because ….





One Sample T Test for Mean μ

C. Choosing a test statistic

Since we don't know σ

(and our sample size is so small that our sample sd probably isn't a good approximation)

$$T = \frac{\bar{X} - \mu_o}{S/\sqrt{n}} \sim \underline{t_{df=n-1}}$$

Where μ_o is the value of μ under null, and in our example: $H_o: \mu = 1.50$

Values of $\bar{X}$ far from μ_0 or equivalently values of t far away from 0 fall __into__ the rejection region and indicate __strong__ evidence against the null H_o.

D. Choosing an α level (to set Type 1 error rate).

p-value	Strength of evidence against the null hypothesis
Greater than 0.1 (or so)	no evidence
Between 0.05 (or so) and 0.1	weak evidence
Between 0.01 (or so) and 0.05 (or so)	moderate evidence
Between 0.001 (or so) and 0.01 (or so)	strong evidence
Less than 0.001 (or so)	very strong evidence

While it is common to use $\alpha = 0.05$, this should __not__ be used too strictly. P-values of 0.051 and 0.049 are very nearly the same amount of evidence as a p value of 0.05 and should be interpreted accordingly (moderate evidence against the null)

$\bar{x} = 1.348$

Revisiting Paint Data

$H_0: \mu = 1.5$ $\quad S_x = 0.3385$

$H_A: \mu \neq 1.5$

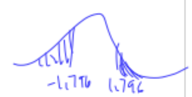
E. Calculate test statistic from data: $t = \frac{\bar{x} - \mu_o}{s/\sqrt{n}} = \frac{1.348 - 1.5}{0.3385/\sqrt{16}} = -1.796$

F. Determine level of evidence by calculating p value or comparing test statistic to critical value

i. Calculate p value or ii. compare observed test statistic to critical values from relevant t curve.

$H_A: \mu \neq 1.5$ dosent matter $\pm$

i. (2-sided) pvalue $= P(T \text{ value as or more extreme than } t_{obs} \mid H_o \text{ is true}) =$



$P(T_{15} < -1.796) + P(T_{15} > 1.796)$

Pvalue $= 2(0.025 <$ value $< 0.05)$ $\qquad$ pvalue $= 0.05 <$ pvalue < 0.1

Ti - 84

$\boxed{\text{stat}}$ $\boxed{\text{edit}}$

list values

$\boxed{\text{stat}}$ $\boxed{\text{calc}}$ $\boxed{\text{1-var stats}}$

$\bar{x}$ sample mean

S_x sample SD

σ_x population SD (Don't use)

in R

t.test(data, mu = 1.5,

This p value gives us __weak__ evidence against the null; there is some evidence the true mean may not be 1.5 but __insufficient__ evidence at the 5% level to reject the null.

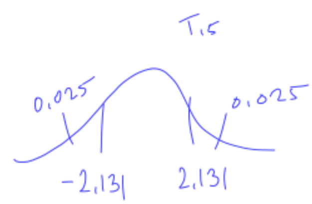
```
> #two-sided p value:
> 2*pt(q=-1.794521, df=15)
[1] 0.09290539
```

smaller pvalue = stronger evidence

Revisiting Paint Data

F. i. Calculate p value or ii. compare observed test statistic to critical values from relevant t curve.

$$t_{obs} = -1.796$$



ii. Rejection region critical values from significance level $\alpha = 0.05$,

Since 2 sided alternative:

Reject $\boldsymbol{H_0}$ if $t_{obs} < -t_{\alpha/2, n-1} = -t_{0.025, 15} = -2.131$

or $t_{obs} > t_{\alpha/2, n-1} = t_{.025, 15} = 2.131$

Since our t_{obs} = __-1.796__ is __not in rejection region__, we would

__fail to__ reject the null at the 5% level, similar to before.

Revisiting Paint Data

G. Compute Confidence Interval with Complementary Confidence Level

Recall the 95% t CI for μ we computed before:

$$\bar{X} \pm t_{.025,15} \frac{s}{\sqrt{n}} = 1.348 \pm 2.131 * \frac{0.339}{\sqrt{16}} = (1.167, 1.529)$$

This interval gives us a range of __plausible_____for μ. Notice it __covers_____ $\mu_0 = 1.5$.

When the two-sided t significance test at level α rejects $H_o: \mu = \mu_o$, the $100(1-\alpha)\%$ confidence interval for μ, will __not_____ contain the hypothesized value μ_o.

When the two-sided t significance test at level α fails to rejects $H_o: \mu = \mu_o$, the $100(1-\alpha)\%$ confidence interval for μ, __will_____ contain the hypothesized value μ_o.

Revisiting Power – was our study *underpowered*?

Suppose we wanted to detect a difference in mean paint thickness of 0.2 mm. Is 16 a large enough sample size? Calculate power with n=16 and $\alpha = 0.10$

- It turns out to be difficult to compute the power of a t-test exactly. Instead, the sample standard deviations from previous studies are sometimes used as an estimate for σ. Once we estimate σ , we calculate power with Z as before.

- Post-hoc (after data gathering and analysis) power tests are not ideal – this should be a before-experiment activity!

- If n is large, then these power calculations will be approximately correct using s instead of σ.

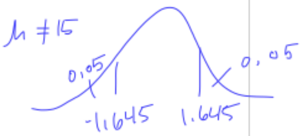
- If n is small and σ is unknown, consult a statistician to help with the details.

Revisiting Power – was our study *underpowered*?

Suppose we wanted to detect a difference in mean paint thickness of 0.2 mm. Is 16 a large enough sample size? Calculate power with n=16 , $\alpha = 0.10$, and $\sigma \approx 0.40$ use Z

$$Power = 1 - \beta_a = P(Reject\ H_o | H_o\ false, \mu_a = 17)$$

1. Define rejection region based on $\alpha = 0.10$, assumed $\sigma \approx 0.40$, and n $H_A: \mu \neq 15$
2. Calculate probability of landing in rejection region for specific alternative.



1. Rejection Region:

$z < -1.645$ or $z > 1.645$

$\bar{X}$ and null: $z = 1.645 = \dfrac{\bar{X}_{obs} - 1.5}{0.40/\sqrt{16}}$

$\bar{X} < -1.645\left(\frac{0.40}{\sqrt{16}}\right) + 1.5 = 1.34$ $\quad \bar{X} > 1.645\left(\frac{0.40}{\sqrt{16}}\right) + 1.5 = 1.66$

$\bar{X} < 1.34$

$\bar{X} > 1.66$

Revisiting Power – was our study *underpowered*?

Suppose we wanted to detect a difference in mean paint thickness of 0.2 mm. Is 16 a large enough sample size? Calculate power with n=16 , $\alpha = 0.10$, and $\sigma \approx 0.40$

$$Power = 1 - \beta_a = P(Reject\ H_o | H_o\ false, u_a = 17)$$

2. Calculate probability of landing in rejection region for specific alternative.

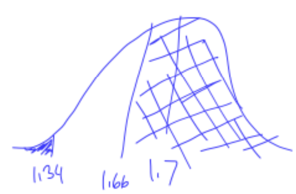
$P(\bar{X} < 1.34 \mid \mu_A = 1.7) + P(\bar{X} > 1.66 \mid \mu_A = 1.7)$

$P\left(z < \dfrac{1.34 - 1.7}{0.40/\sqrt{16}}\right) + P\left(z > \dfrac{1.66 - 1.7}{0.40/\sqrt{16}}\right)$

$P(z < -3.6) + P(z > -0.40)$

$< 0.0003 + (1 - 0.3446)$

smallest value basically 0 $\quad + \boxed{0.6654}$



power of 66% ok