Review Session

Topics "End of Course"/Part 3

① CONFIDENCE INTERVALS / HYPOTHESIS TESTS

- Mean → of Normal pop^n (?)
  → variance known/unknown
  → sample size big/small
- Population Proportion
- Difference between two Means
  → underlying distr. Normal
  → sample sizes small
  → variance unknown → equal
  → unequal

② LINEAR REGRESSION

→ "line of best fit" = least squares regression line $Y = \beta_0 + \beta_1 X + \varepsilon$

error

→ Estimated β_0, β_1 with $\hat{\beta}_0$ & $\hat{\beta}_1$

→ Hypothesis Test, C.I., Prediction Interval for β_1 with t-distribution

→ ANOVA (Analysis of Variance Approach) i.e. F-test for β_1

→ Correlation

③ SINGLE FACTOR EXPERIMENTS

→ ANOVA i.e. F-test on differences of many means at once

→ Fisher LSD (Least Significant Difference) Tests (one test for each pair of means)

Confidence Intervals & Tests (Test Statistics)

① C.I.s/Tests for mean μ

z − C.I./Test → N(0, 1) distribution

Ⓐ Underlying pop. = Normal
Variance σ^2 : Known
size of sample Size n : Irrelevant

100(1−α)% C.I.: $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ (20)

Test : $z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ (23)

value ass. in H_0

t − C.I./Test → t − distribution

Ⓑ Underlying pop. = Normal
Variance σ^2 : UNknown (estimate with s^2)
Size of sample n : small ($n < 40$)

100(1−α)% C.I.:

$\bar{x} \pm t_{n-1, \frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ (21)

Test : $t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ (24)

Ⓒ Underlying pop. : unknown
Variance σ^2 : unknown (est. with s^2)
sample size n : big ($n > 40$)

100(1−α)% C.I. : $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

Test : $z_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

} Not on formula sheet directly − need to know how to adapt what's there.

② C.I.s / Tests on Pop. Proportion

No choice !

Need $np, n(1-p) > 5$

$100(1-\alpha)\%$ C.I.: $\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ (22)

Test stat.: $Z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ (25)

③ C.I.s / Tests on Difference of 2 Means

→ assume variances UNKNOWN

→ Always a t-test / t-C.I.

Choice: (A) 2 variances same $\sigma^2 = \sigma_1^2 = \sigma_2^2$ (B) 2 vars NOT same $\sigma_1^2 \neq \sigma_2^2$

(A)

$100(1-\alpha)\%$ C.I.:

(26) $\bar{x}_1 - \bar{x}_2 \pm t_{\frac{\alpha}{2}, n_1+n_2-2} \, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

where $S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$ (26)

Test stat. $t_{n_1+n_2-2} = \dfrac{\bar{x}_1 - \bar{x}_2 \ (-0)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ (28)

S_i^2 = sample variance of popⁿ i

(B)

$100(1-\alpha)\%$ C.I.:

(27) $\bar{x}_1 - \bar{x}_2 \pm t_\nu \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

where ν = # degrees of freedom here (27)

Test stat.

$t_\nu = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ (29)

Sample Test 3 Q. 4 - 7

Measurements of sugar concentration in apple juice:

11.48, 11.45, 11.48, 11.47, 11.48, 11.50, 11.42, 11.49

(4.) Test $H_0 : \mu = 11.5$ v. $H_1 : \mu \neq 11.5$ (2-sided)

Using $\alpha = 0.05$.

different than

(NOT $\mu > 11.5$ bigger than
$\mu < 11.5$) smaller than

Questions to ask yourself?

- What are we testing? → Mean (single mean)
- Do we know variance? → No
- Is sample size big/small? → $n = 8$ small
- Is the underlying pop. normal? → We don't know!

We need an assumption to get unstuck.
What can we assume?

- We could calculate s^2 (sample variance) but we still don't know this is the underlying var.

But still stuck: cannot proceed if sample size small & variance unknown without assuming underlying pop. is normal.

So we are in setting (B) t-test territory <u>if</u> we assume underlying pop. normal (so we do).

For a 2-sided t-test on mean,

($H_1 : \mu \neq 11.5$) $\qquad T_0 = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$-distr.

with critical value $t_{n-1, \frac{\alpha}{2}}$ ← α if 1-sided

$= t_{7, 0.025}$

$= 2.365$

We need to compare t_0 with critical value

"$\boxed{\dfrac{\bar{x} - 11.5}{s/\sqrt{8}}}$"

$$\bar{x} = \frac{11.48 + 11.45 + \ldots + 11.49}{8} = 11.47125$$

To find s, find s^2 & square root!

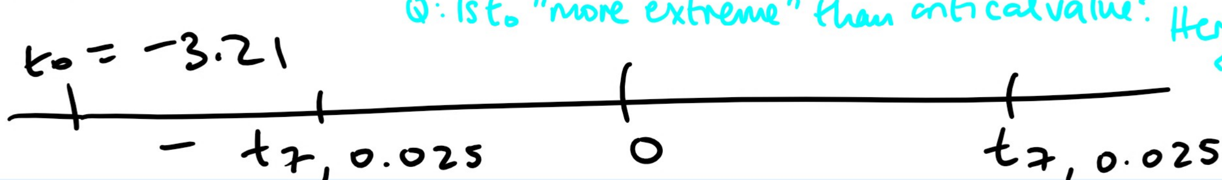
$$(16)\ s^2 = \frac{1}{n-1}\left(\sum x_i^2 - n\bar{x}^2\right) = \frac{1}{7}\left(\sum x_i^2 - 8(11.47125)^2\right)$$

$$\leadsto s = 0.0253\ldots$$

So $t_0 = -3.21$ ← Compare this to critical value.

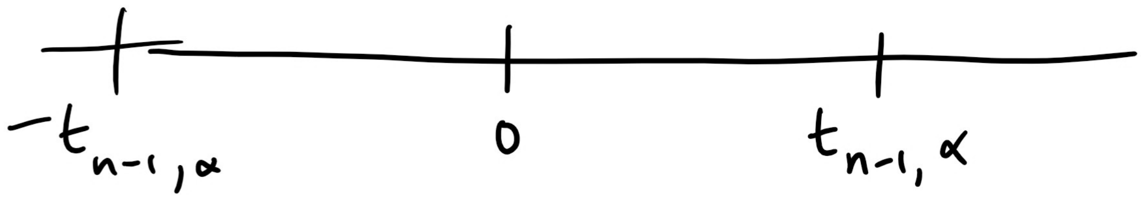
When we do a 2-sided test:

Q: Is t_0 "more extreme" than critical value?



Yes, so Reject H_0 for H_1

Aside: 1-sided test



Check: is to more extreme than whichever of $t_{n-1,\alpha}$ or $-t_{n-1,\alpha}$ makes sense

i.e. is $|t_0| > t_{n-1,\alpha}$?

If Yes reject H_0 for H_1.

(5.) Based on (4) what could be true?

(a) Type I Error? $\longrightarrow$ Reject H_0 when H_0 true?

(b) Type II Error? $\longrightarrow$ Accept H_0 when H_0 false? X

(c) True pop. mean is $\neq 11.5$? $\longrightarrow$ Our test outcome said "we reject the ass. $\mu = 11.5$"

(d) $P(\text{Type I Error}) = 0.025 \longrightarrow \alpha = 0.05 = P(\text{Type I Error})$

(e) $P(\text{Type II Error}) = 0.95$

↳ We didn't work out anything in (4.) with β = Type II Error. (We would, say, have needed to know the true value of the mean)

(6.) Find p-value: P-value

$= P(T_0$ [test stat.] $\underbrace{\text{as extreme as}}_{\text{2-sided}}$ observed value$)$ [−3.21]

$= P(T_0 < -3.21) + P(T_0 > 3.21)$

(If 1-sided test only take the relevant prob. of this pair.)

e.g. $H_1: \mu < 11.5$; P-value $= P(T_0 < -2.31)$

$= 2\boxed{P(T_0 > 3.21)}$ (as t-dist. is symmetric)

= We cannot calculate exactly using t-table (careful: if Z-test you can look it up exactly in Z-table!)

We can estimate (say which interval it lies in)

t_7-distr.

Look in row 7. Where does 3.21 lie?

3.21 falls between 2.998 and 3.499

So $P(T_0 > 3.21)$ falls " 0.01 and 0.005 (corresponding column headers)

So P-value falls between 2×0.01 and 2×0.005

i.e. in $(0.01, 0.02)$.

(7.) Assumption required?

Underlying distr. is normal. (See above under (4.))

(8.) How to check this assumption?

To check any assumption about which prob. distribution something has: do a prob plot & check to see if points fall close to a straight line.