19.03.13

RECALL THAT CONFIDENCE INTERVALS FOR REGRESSION COEFFICIENT β_1 AND β_2 :

LEVEL OF SIGNIFICANCE

$$Pr\left(-t_{\frac{\alpha}{2}} \leq t \leq t_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$Pr\left(-t_{\frac{\alpha}{2}} \leq \frac{\hat{\beta}_2-\beta_2}{se(\hat{\beta}_2)} \leq t_{\frac{\alpha}{2}}\right) = 1-\alpha$$

OR $$Pr\left[\hat{\beta}_2 - t_{\frac{\alpha}{2}}\cdot se(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t_{\frac{\alpha}{2}}\cdot se(\hat{\beta}_2)\right] = 1-\alpha$$

OR, IN COMPACT FORM, $100(1-\alpha)\%$ CONFIDENCE INTERVAL FOR β_2:

$$\hat{\beta}_2 \pm t_{\frac{\alpha}{2}}\cdot se(\hat{\beta}_2).$$

EXAMPLE: MEAN HOURLY WAGE (Y) ON EDUCATION (X)

$\hat{\beta}_2 = 0.7240$

$se(\hat{\beta}_2) = 0.07$

$n = 13$, $df = n-2 = 11$, $\alpha = 0.05$

$1-\alpha = 0.95$

CRITICAL $t_{\frac{\alpha}{2}} = 2.201$

$$Pr\left[\hat{\beta}_2 - t_{\frac{\alpha}{2}}\cdot se(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t_{\frac{\alpha}{2}}\, se(\hat{\beta}_2)\right] = 1-\alpha$$

SO, CONFIDENCE INTERVAL FOR REGRESSION COEFFICIENT β_2 IS

$$0.5700 \leq \beta_2 \leq 0.8780.$$

OR $$\underset{\hat{\beta}_2}{0.7240} \pm \underset{t_{\frac{\alpha}{2}}}{2.2010}\,(\underset{se(\hat{\beta}_2)}{0.07}).$$

INTERPRETATION ?

GIVEN THE CONFIDENCE COEFFICIENT OF 95% ($100(1-\alpha)$), IN 95 OUT OF 100 CASES INTERVALS WILL CONTAIN THE TRUE β_2.

WRONG INTERPRETATION → THE PROBABILITY IS EQUAL TO 95 PERCENT THAT THE "SPECIFIC INTERVAL" ABOVE CONTAIN THE TRUE β_2.

PLEASE AVOID THIS.

WHY THIS IS WRONG ?

B/C THE INTERVAL LIKE ABOVE IS NOW FIXED AND NO LONGER RANDOM. SO THE PROBABILITY THAT

"THE SPECIFIC FIXED INTERVAL INCLUDES THE TRUE β_2 IS THEREFORE 1 OR 0.

<u>EXERCISE</u> GIVEN THE SAME INFORMATION ON

$\hat{\beta}_2$, $se(\hat{\beta}_2)$, n, CONSTRUCT

A CONFIDENCE INTERVAL FOR β_2 WHEN $\alpha = 0.01$

$\hat{\beta}_2 = 0.7240$, $se(\hat{\beta}_2) = 0.07$, $n = 13$, $df = n-2 = 11$

$\alpha = 0.01$, $1-\alpha = 0.99$

CRITICAL $t_{\frac{\alpha}{2}}$ = 3.106.

$$Pr\left[\hat{\beta}_2 - t_{\frac{\alpha}{2}} \cdot se(\hat{\beta}_2) \le \beta_2 \le \hat{\beta}_2 + t_{\frac{\alpha}{2}} \cdot se(\hat{\beta}_2)\right] = 1-\alpha.$$

CONFIDENCE INTERVAL FOR β_2 IS

$$0.7240 \pm 3.106(0.07)$$

IN REPEATED SAMPLING (i.e., IN LONG RUN SENSE), IF WE CONSTRUCT 100 CASES OF CONFIDENCE INTERVAL LIKE ABOVE, 99 OUT OF 100 INTERVALS WILL CONTAIN THE TRUE BUT UNKNOWN β_2.

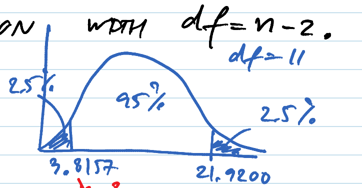
<u>CONFIDENCE INTERVALS FOR σ^2</u>

WE CAN ALSO CONSTRUCT CONFIDENCE INTERVAL FOR σ^2:

$$Pr\left(\chi^2_{1-\frac{\alpha}{2}} \le \chi^2 \le \chi^2_{\frac{\alpha}{2}}\right) = 1-\alpha$$

WHY CHI-SQUARE?

B/C UNDER THE NORMALITY ASSUMPTION, THE VARIABLE σ^2 FOLLOWS THE CHI-SQUARE DISTRIBUTION WITH $df = n-2$.



<u>EXAMPLE</u> WAGE - EDUCATION

$\hat{\sigma}^2 = 0.8936$ (WHY?)

$\alpha = 0.05$, $df = n-2 = 13-2 = 11$

$\chi^2_{0.025} = 21.9200$, $\chi^2_{0.975} = 3.8157$.

$\chi^2_{0.975}$: $1-\frac{\alpha}{2}$ $\chi^2_{0.025}$: $\frac{\alpha}{2}$

SO $0.04484 \le \sigma^2 \le 2.25760$.

NOTE: $Pr\left[(n-2)\dfrac{\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}}} \le \sigma^2 \le (n-2)\dfrac{\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2}}}\right] = 1-\alpha$

GIVEN THE CONFIDENCE COEFFICIENT OF 95 PERCENT, INTERVALS LIKE WE CONSTRUCTED ABOVE WILL CONTAIN THE TRUE σ^2

<u>EXERCISE</u>: $\hat{\sigma}^2 = 0.8936$

$\alpha = 0.01$, $df = n-2 = 13-2 = 11$

EXERCISE 9:

$$= 0.8936$$

$$\alpha = 0.01, \quad df = n-2 = 13-2 = 11$$

$$\chi^2_{\frac{\alpha}{2}} = \underline{\bigcirc} \qquad \chi^2_{1-\frac{\alpha}{2}} = \underline{\bigcirc}$$

$$\underline{\bigcirc} \leq \sigma^2 \leq \underline{\bigcirc}$$

INTERPRETATION → IF WE ESTABLISHED 99 PERCENT CONFIDENCE LIMIT ON σ^2 AND IF WE MAINTAIN A PRIORI THAT THESE LIMITS WILL INCLUDE THE TRUE σ^2, WE WILL BE RIGHT IN THE LONG RUN 95 PERCENTS OF THE TIME.

21.03.13

HYPOTHESIS TESTING: CONCEPTUAL IDEA

LET'S HAVE A LOOK FROM NON STATISTICAL HYPOTHESIS TESTING FIRST.

A CRIMINAL TRIAL IS AN EXAMPLE OF HYPOTHESIS TESTING W/O THE STATISTICS.

IN A TRIAL, A JURY MUST DECIDE BETWEEN TWO HYPOTHESES.

THE NULL HYPOTHESIS IS H_0: THE DEFENDANT IS INNOCENT

THE ALTERNATIVE HYPOTHESIS OR RESEARCH HYPOTHESIS IS

H_1: THE DEFENDANT IS GUILTY.

THE JURY DOES NOT KNOW WHICH HYPOTHESIS IS TRUE. THEY MUST MAKE A DECISION ON THE BASIS OF EVIDENCE PRESENTED.

IN THE LANGUAGE OF STATISTICS,

CONVICTING THE DEFENDANT ↔ REJECTING THE NULL HYPOTHESIS IN FAVOR OF THE ALTERNATIVE HYPOTHESIS.

(ENOUGH EVIDENCE TO SUPPORT H_1)

IF THE JURY ACQUITS IT. ↔ THERE IS NOT ENOUGH EVIDENCE TO SUPPORT H_1.

NOTE THAT THE JURY IS NOT SAYING THAT THEY DEF... IS INNOCENT, ONLY THAT THERE IS NOT ENOUGH EVIDENCE TO SUPPORT THE ALTERNATIVE HYPOTHESIS.

THAT IS WHY WE SHOULD NOT SAY THAT WE ACCEPT THE NULL HYPOTHESIS, ALTHOUGH MOST PEOPLE IN ACADEMICS WILL SAY "WE ACCEPT THE NULL HYPOTHESIS".

THERE ARE TWO POSSIBLE ERRORS.

TYPE I ERROR: WHEN WE REJECT A TRUE NULL HYPOTHESIS. (PUT AN INNOCENT PERSON INTO JAIL)

TYPE II ERROR: WHEN WE DON'T REJECT A FALSE NULL HYPOTHESIS. (RELEASE AN GUILTY PERSON BACK

TYPE II ERROR : WHEN WE DON'T REJECT A FALSE NULL HYPOTHESIS. (RELEASE AN GUILTY PERSON BACK TO THE SOCIETY)

H_0	T	F
REJECT	TYPE I ERROR	
DON'T REJECT		TYPE II ERROR

i.e., REJECT H_0 WHEN IT IS TRUE.

i.e., DON'T REJECT H_0 WHEN IT IS FALSE.

NON STATISTICAL HYPOTHESIS TESTING

THE CRITICAL CONCEPTS ARE AS FOLLOWS:

1. WE HAVE H_0 AND H_1
2. WE BEGIN W/ THE ASSUMPTION THAT THE NULL HYPOTHESIS IS TRUE.
3. THE GOAL ♡ IS TO DETERMINE WHETHER THERE IS ENOUGH EVIDENCE TO INFER THAT THE ALTERNATIVE HYPOTHESIS IS TRUE, OR THE NULL IS NOT LIKELY TO BE TRUE.
4. THERE ARE TWO POSSIBLE DECISIONS:

DECISION - 1 CONCLUDE THAT THERE IS ENOUGH EVIDENCE TO SUPPORT H_1 $\Leftrightarrow$ REJECT THE NULL (H_0)

OR

DECISION 2 CONCLUDE THAT THERE IS NOT ENOUGH EVIDENCE TO SUPPORT H_1 $\Leftrightarrow$ ~~FAIL TO~~ REJECT THE NULL HYPOTHESIS.

THE STEPS TO DO HYPOTHESIS TESTING

1. DEFINE THE POPULATION IN STUDY
2. STATE THE HYPOTHESIS TO BE INVESTIGATED.

i.e.,

2-TAIL

$H_0 : \beta_2 = \beta_2^*$

$H_1 : \beta_2 \neq \beta_2^*$

OR

RIGHT-TAIL

$H_0 : \beta_2 \leq \beta_2^*$

$H_1 : \beta_2 > \beta_2^*$

OR

$H_0 : \beta_2 \geq \beta_2^*$

$H_1 : \beta_2 < \beta_2$

LEFT-TAIL

3. GIVE THE DESIRED SIGNIFICANCE LEVEL (α)
4. SELECT A SAMPLE FROM POPULATION
5. COLLECT THE DATA.
6. PERFORM YOUR COMPUTATION
7. REACH A CONCLUSION.

EXAMPLE: WAGE - EDUCATION (FROM HANDOUT)

$\hat{\beta}_2 = 0.724097$

$se(\hat{\beta}_2) = 0.069581$

$n = 13 \rightarrow df = n-2 = 13-2 = 11$

$\alpha = 0.05$

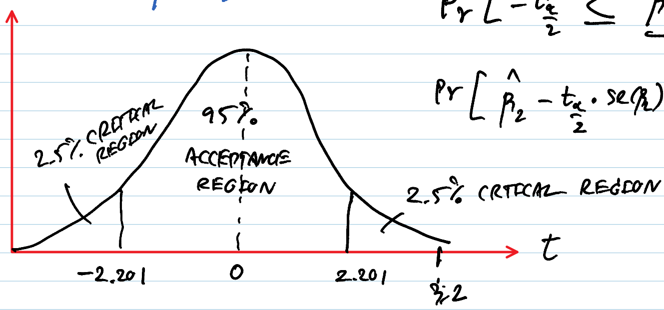
$H_0: \beta_2 = 0.5$

$H_1: \beta_2 \neq 0.5$

$t_{\frac{\alpha}{2}} = 2.201$

$Pr\left[-t_{\frac{\alpha}{2}} \leq \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \leq t_{\frac{\alpha}{2}}\right] = 1-\alpha$

$Pr\left[\hat{\beta}_2 - t_{\frac{\alpha}{2}} \cdot se(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t_{\frac{\alpha}{2}} \cdot se(\hat{\beta}_2)\right] = 1-\alpha$



$\hat{t} = \frac{\overset{\hat{\beta}_2}{0.724097} - \overset{\beta_2}{0.5}}{0.069581}$

$= \underline{\underline{3.2}}$

THE PROBABILITY OF OBSERVING $\hat{t} > 2.201$ IS ONLY 2.5% (VERY SMALL, ACTUALLY), BUT WE STILL OBSERVE IT,

THEREFORE, THE NULL HYPOTHESIS $H_0: \beta_2 = 0.5$ SHOULD NOT BE TRUE AND WE REJECT IT.

EXERCISE

$H_0: \beta_2 = 0$ (EDUCATION HAS NO INFLUENCE ON WAGE)

$H_1: \beta_2 \neq 0$ (EDUCATION HAS AN INFLUENCE ON WAGE)

EXAMPLE SALES & PRICE

SUPPOSE $\hat{Y}_i = 49.667 - 2.1576 X_i$

SALES — PRICE

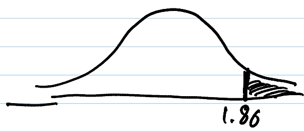
$\hat{\beta}_2 = -2.1576$

$se(\hat{\beta}_2) = 0.1204$

$df = 10-2 = 8$

$\alpha = 0.05$

$t_{0.05} = 1.86$



$H_0: \beta_2 \leq 0$

$H_1: \beta_2 > 0$

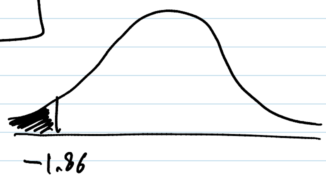
$\hat{t} = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \qquad \frac{-2.1576 - 0}{0.1204} = -17.92$

WE DO NOT HAVE ENOUGH EVIDENCE IN SUPPORTING H_1 AND SO WE FAIL TO REJECT H_0.

HOW ABOUT? $\Rightarrow$ $\boxed{H_0: \beta_2 \geq 0}$ $\rightarrow$ LEFT TAIL TEST

HOW ABOUT ? $\Rightarrow$ $\boxed{\begin{array}{l} H_0 : \beta_2 \geq 0 \\ H_1 : \beta_2 < 0 \end{array}}$ $\rightarrow$ LEFT TAIL TEST

$$\hat{t} = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \frac{-2.1576 - 0}{0.1204} = -17.92$$



As $\hat{t} < -1.86$, WE FOUND THAT THERE IS ENOUGH EVIDENCE TO SAY THAT $\beta_2 < 0$.