HW 8

27	45	72	58	31	60	34	74
250	285	320	295	265	298	267	321.

$\bar{x} = 50.125$ $\quad \bar{y} = 287.625$

$S_x = 18.497$ $\quad S_y = 25.867$

$\sum x^2 = 22495$ $\quad \sum y^2 = 666509$ $\quad \sum xy = 118652$

$S_{xx} = 2394.875$ $\quad S_{xy} = 3314.375$ $\quad S_{yy} = 4683.875$

$\beta_1 = 1.384.$ $\quad b_0 = 218.25$

$S_{xx} = \sum x^2 - n\bar{x}^2$,

$= 22495 - 8 * 50.125^2 =$

$$S_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{S_{xx}}{n-1}$$

$$S_{xx} = (n-1) \cdot S_x^2 = 7 * 18.495^2$$

$$S_{yy} = (n-1) S_y^2 = 2394.\cancel{455}$$

$$S_{xy} = \sum XY - n \bar{X} \bar{Y}$$

$$= 118652 - 8 * 50.125 * 287.625$$

$$= 3314.375$$

ⓐ

① $$b_1 = \frac{S_{xy}}{S_{xx}} = \frac{3314.375}{2394.875} = 1.384$$

b

$$b_0 = \bar{y} - b_1 \bar{x} = 287.625 - 1.384 \cdot 50.125$$

$$= 218.25$$

② $$\hat{y} = 218.25 + b_1 * 65 = 308.21$$

③ $$s^2 = \frac{SSE}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2}$$

$$= \frac{4683.875 - 1.384 * 3314.375}{6}$$

$$= 16.13 \qquad s = 4.016$$

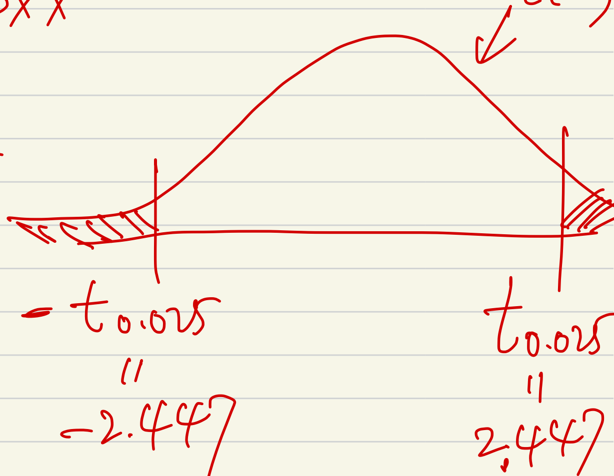
④ $H_0: \beta_1 = 0 \quad H_1: \beta_1 \neq 0$

$\alpha = 0.05$.

under H_0: $\dfrac{b_1}{S/\sqrt{S_{xx}}} \sim t(n-2)$

$t_{obs} = \dfrac{1.384}{4.016/\sqrt{2394.875}}$

$= 16.86 \in C$.



Reject H_0.

conclude $\beta_1 \neq 0$.

⑤ $R^2 = 1 - \dfrac{SSE}{SST} = 1 - \dfrac{96.78}{S_{yy}}$

4683.875

$= 1 - 0.02$

$= 0.98$.

⑦ 95% for the mean when $x = 65$

⑧

About rounding.

① what's the necessary sample size? $n \geq 71.2$

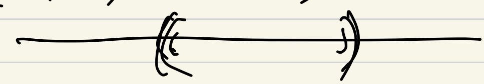
$n = 72.$

② what's the confidence interval?

$(71.678, \ 80.123)$

$\longrightarrow (71.7, \ 80.1) \ \times ??$

$(71.6, \ 80.2)$



Sample final:

1a Poisson $\lambda = 2/100$ feet.

$P(< 3$ flaws on a 200 foot screen$)$

$= P(X = 0, 1, 2$ when $\lambda = 4)$

$= e^{-4}\left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!}\right)$

$\boxed{\lambda = 4}$

1d. A : 20, B : 30.

mix 50. Grade 12.

①. $$\frac{\binom{30}{5}\binom{20}{7}}{\binom{50}{12}}$$

②. $$1-\left(\frac{\binom{30}{12}}{\binom{50}{12}}+\frac{\binom{20}{12}}{\binom{50}{12}}\right)$$

Q5. (d). $H_0: \beta_1 = 0$. $H_1: \beta_1 \neq 0$.

under H_0. $$\frac{b_1}{S/\sqrt{S_{xx}}} \sim t(n-2)$$

Q3.

	Smokers	non-smoker	
TB	42	218	260
w/o TB	15	245.	260
	57	463	520

a

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$.

$\hat{p}_1 = \frac{42}{260}$

$\hat{p}_2 = \frac{15}{260}$ ...

$\hat{p}_0 = \frac{57}{520} = $ — —

b $P(TB \mid \text{smoker})$ can't be estimated using this data set. Because we don't have a random sample of smokers!

Q1.

a $H_0: \sigma_1^2 = \sigma_2^2 \quad H_1: \sigma_1^2 \neq \sigma_2^2$.

under H_0: $\frac{S_1^2}{S_2^2} \sim f(24, 24)$.

b cannot rej H_0

we can assume $\sigma_1^2 = \sigma_2^2$.

$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 > \mu_2$

under H_0: $\frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{25} + \frac{1}{25}}} \sim t(48)$

Sample final Q4.

a paired data

b

1%	grand
5%	2nd
10%	3rd
84%	nothing

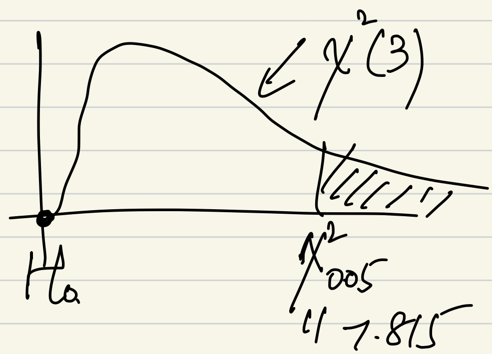
	Grand	2nd	3rd	nothing
O_i	30	120	210	1640
e_i	20	100	200	1680

H_0: Good fit H_1: not a good fit

under H_0. $\chi^2 = \sum \frac{(O_i - e_i)^2}{e_i} \sim \chi^2(3)$

$$\chi^2_{obs} = \frac{100}{20} + \frac{400}{100} + \frac{100}{200} + \frac{1600}{1680}$$

≈ 10.5



$C = \{\chi^2 > 7.815\}$

$\chi^2_{obs} \in C \Rightarrow$ Rej H_0

1025

Ex: claim: driven $> 20{,}000$ km/year

$n = 100$

would you agree with the claim

$\bar{X} = 23{,}500$, $S = 3{,}900$?

$\alpha = 0.05$. p-value.

H_0. $\mu = 20{,}000$ H_1 $\mu > 20{,}000$

under H_0, $\dfrac{\bar{X} - 20{,}000}{S/\sqrt{100}} \sim Z$.

$\dfrac{\bar{X} - 20{,}000}{390} \sim Z$.

p-value

Conclusion.

$$z_{obs} = \frac{23500 - 20000}{390}$$

$$= \frac{3500}{390} = 8.97$$

$$\text{p-value} = P(Z > 8.97) \approx 0$$

reject H_0. agree to H_1.

$$C = \left\{ \frac{\bar{X} - 20000}{390} > z_{0.05} = 1.645 \right\}$$

$$= \left\{ \bar{X} > 20641.55 \right\}$$

$$\bar{X}_{obs} = 23500 \in C.$$

P(type I error) = P(rej H_0 when H_0 true)

P(type II error) = P(not rej H_0 when H_0 false)

power of the test

= 1 − P(type II error)

= P(rej H_0 when H_0 false)

Thursday 12/16.

9:30 — 10:30 pm