This Final Exam Review

Time: Session

Next Final: Thu 15 Jun Noon-3pm
Time: In Media Theater mid

Last ex. of INT ReCt. & REG: pp. 140-

Y = 4 P. 140

Reg Line For Pred

X from Y Line

Reg Line For Pred Y from X

Wife has 18 yr of schooling =
This is 18 - 12 = 2
So we predict her
husband will only
be 1.2 yrs above
ave = (0.5)(2) yrs

Above ave = 1.5 yrs above ave = 12 + 3 = 15
yrs of schooling. (92)

H. has 15 yrs
of schooling = this is 15 - 12 = 15 yrs
above ave. For H, we predict will
only be 1.2 = 0.5(3) above ave +
(0.5)(3 yrs) = 1.5 yrs above ave + 12 +
1.5 = 13.5 yrs.
% FINAL PROBLEMS, OPEN EVERYTHING
  . Prob models for sums and means
  . 2 independent samples
  . 1 sample inference on proportions
  . Correlation & regression
  . 2 sample paired comparisons

At least once: you will be asked if inference is appropriate.

#5 on final exam review (9)

\[
\begin{align*}
\hat{y} &= a + bx \\
\text{PRE} \left( \text{ht.} \right) &= 66.75 + \left( 0.25 \frac{\text{in}}{\text{yr}} \right) ( \text{yr} )
\end{align*}
\]

(i) If \( x = 12 \text{ yrs} \to \)
\[
\begin{align*}
\hat{y} &= 66.75 \text{ in} + (0.25 \frac{\text{in}}{\text{yr}})(12 \text{ yr}) \\
&= 69.75 \text{ in}
\end{align*}
\]

(ii) If \( x = 16 \text{ yrs} \to \)
\[
\begin{align*}
\hat{y} &= 66.75 \text{ in} + (0.25 \frac{\text{in}}{\text{yr}})(16 \text{ yr}) \\
&= 70.75 \text{ in} \quad (1 \text{ inch taller})
\end{align*}
\]
INTERPRETATION: THE QUESTION IS LONGITUDINAL BUT DATA ARE CROSS-SECTIONAL; IT IS DIFFICULT TO DRAW VALID LONG-TERM CONCLUSIONS FROM CROSS-SECTIONAL DATA. CORRECT INTERPRETATION:

ASSOC. WITH BAHT HT.

INCREASE OF 1 YR OR

ED. IS AN INCREASE OF 1/4 IN IN HT: IF YOU COMPARE 2 GROUPS OF MEN 25-34 WHOSE ED. LEVELS DIFFER BY 4 yrs, YOU EXPECT TO FIND THEIR HTS WILL DIFFER ON AVE. BY ABOUT 1 IN. (REASON: PLFS: BETTER NUTRITION)

ASSOC. (CORR.) ≠ CAUSATION: THIS SURVEY IS CROSS-SECTIONAL & OBSERVATIONAL, SO WE CAN'T
DEAL WITH HISTORICAL, CASUAL CONCLUSIONS LIKE THE ONE IN G1. INSTEAD: PCE (NUTRITION) CAUSING BOTH INCOME, HT TO APPEAR TO BE ASSOCIATED.

\[ \begin{align*}
\text{POP} & \quad \text{(a)} \quad \text{SAMPLE} \quad \text{POPULATION} \quad \text{IMPROB.} \\
\text{ALL ELIGIBLE} & \quad \text{THE OBS.} \quad \text{PEOPLE} \\
0 & \quad 0 = M \\
\text{OTHERS} & \quad \text{POSSIBLE} \\
\end{align*} \]

\[ N = \begin{cases} 
15 & \text{SRS} \\
5 & \text{SRS} \\
\end{cases} \]

\[ \begin{align*}
\text{MEAN } p & = ? \\
\text{HYP.} & \quad \text{15} \\
\text{40} \quad \text{h = 350} \\
\text{L.R. MEAN} & = \frac{57.3}{56.0} \\
\text{L.R.} & \quad \text{1} \\
\text{L.R.} & \quad \text{2} \\
\text{L.R.} & \quad \text{L.R.} \\
\end{align*} \]

\[ \text{MEAN } p = 53\% \]

JUST LIKE MR. FRANK AUFERST E5
1. \( E\text{IID}(p) = p = 53\% \)

2. \( SE\text{IID}(p) = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.53)(0.47)}{350}} = 2.7\% \)

3. \( SE 2.7\% \)

L.R. Hist. of \( p \)

\[ \frac{102}{350} = 29\% \]

Chance = 0%

(b)

\[ [\text{Likelihood}] \]

\[ \text{n} = 100 \]

\[ \text{JITTO} \]

Mean \( p = 29\% \)

Mean \( \hat{p} = ? \)

SE 4.5

\[ L.R. \]

Hist. of \( \hat{p} \)

\( SE(\hat{p}) = \sqrt{\frac{p(1-p)}{n}} = \sqrt{(0.29)(0.71)} \]

\[ -4.4 \]

\[ -4.5\% \]
1. Mean $\hat{p}_1 = 9.8\%$

2. Independent samples (like $\odot$ in prefer. treatment)

$$\frac{\overline{1} - \overline{2}}{\overline{2}} = \frac{(12.5\%) - (9.8\%)}{(9.8\%)} = \frac{2.7\%}{9.8\%} = +28\%$$

I.e. Prop. of people 65+ in U.S. went up 28% in only 20 yrs - (highly) pract. sig.; but is this diff. statistically (or does $\odot$ not make sense?)
\[ N = \begin{bmatrix} 203 \\ 0 \end{bmatrix} \]  

\[ N = \begin{bmatrix} 203 \\ 0 \end{bmatrix} \]  

\[ n_1 = 203, \quad n_2 = 0 \]

\[ \text{Mean } \hat{p} = 9.8\% \]

**Sample = Pop, No Uncertainty; No Need for Stat. Inference.**

**Stat. Sig. Doesn't Make Sense**

\[ \hat{p}_1 = 49.3\%, \quad \hat{p}_2 = 48.4\% \]

This difference is not large in practical terms.

\[ \hat{p}_1 - \hat{p}_2 = 0.9\% \]

\[ SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \]

\[ = \sqrt{\left(\frac{0.493}{203}\right) + \left(\frac{0.484}{0}\right)} = 4.5\% \]

95% CI for \((\hat{p}_1 - \hat{p}_2)\): \((\hat{p}_1 - \hat{p}_2) \pm 1.96 \times SE(\hat{p}_1 - \hat{p}_2)\)
\[
= 0.9\% \pm 2(4.5\%) = 0.9\% \pm 9.0\%
\]
\[
\frac{\bar{y} - \bar{z}}{\bar{z}} = \frac{16.8 - 24.3}{24.3} = -31\% \text{ drop}
\]

Income support caused a 31\% drop in wks. of PA 1 work. (Highly practical sig.) Stat. sig.? Model is just like model in CS about skeletons. (Refer to that for details; don't skip model details on final if problem asks for model.)

\(\bar{y}_1 = 16.8\ \text{wks}\)  \(s_1 = 15.9\ \text{wks}\)  \(n_1 = 592\)

\(\bar{y}_2 = 24.3\)  \(s_2 = 17.3\)  \(n_2 = 154\)
\((\bar{y}_1 - \bar{y}_2) = -7.5 \text{ WKS.}\)

\[
SE(\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{15.9^2}{500} + \frac{17.3^2}{154}}
\]

\[= 2.4 \text{ WKS}\]

\[95\% \text{ CI: } (\bar{y}_1 - \bar{y}_2) \pm 2 \times SE(\bar{y}_1 - \bar{y}_2) = -7.5 \pm 4.8 \text{ WKS}\]

\[
\begin{array}{cccc}
-12.3 & -7.5 & -3.7 & 0
\end{array}
\]

0 is not in this CI, so this diff. is **stat. sig.**

**#1**

**Pop.**

**All n. M. E. S.**

**At push + r.**

**Sample**

**The obs.**

**A. M. E. H.**

** Imag. Data**

**Possible**

**F. K.**

\[
N = \text{BIG}
\]

\[
\text{Mean } \mu = 155 \text{ LB}
\]

\[
\sigma = 30 \text{ LB}
\]
\[ P(\text{break} \mid \text{break}) = P(S > 32,000 \text{ lb.}) \]
\[ = P \left( \frac{S}{n} > \frac{32,000}{n} \right) = P \left( \frac{S}{200} > \frac{32,000}{200} \right) \]
\[ = P(\bar{Y} > 160 \text{ lb}) =? \]

4. \[ E_{\text{IID}}(S) = \mu = (200)(155 \text{ lb}) = 31,000 \text{ lb.} \]

5. \[ \text{SE}_{\text{IID}}(S) = \sigma \sqrt{n} = (300 \text{ lb}) \sqrt{200} = 424 \text{ lb.} \]

6. \[ \text{SE} 424 \]

1% 98% 1% of S

3,000 32,000

-2.35 2.35

CHANCE OF BREAKDOWN = 1% 2.35

1 in every 100 fully loaded trips; so such trips every day, so would break every \( \frac{100}{90} = 1.1 \) days; Not good ENOUGH.
\[
\begin{align*}
X &= (3.75) (424 \text{ LB}) + 31,000 \text{ LB} \\
   &= 32,590 \text{ LB} \quad \text{ENG. THE ESC TO HOL OPENLY AN EXTRA \fbox{590 LB}} \\
\text{ABOUT 4 EXTRA PEOPLE} \text{ DROPS FAILURE RATE FROM } \frac{1}{100} \text{ TO } \frac{1}{10000}. \\
\text{EVEN THIS WOULD BE FAILING ABOUT EVERY } \frac{10000}{90} = 111 \text{ DAYS} \quad \text{(STILL NOT GOOD ENOUGH?)} \\
\text{WE WILL HOL OPENLY REGULAR OFFICE (J, ETHAN, VANEER, KEN)} \quad \text{THIS WEEK.}
\end{align*}
\]