This final exam review (concluded)

next Mon 6 June 8-11 AM
here (Baskin Auditorium)

review problems
2(a) 2 independent samples, 0/1 outcome
4. see chap. 15 p. 102

all relevant
parameters
are
relevant
respectively?

\[ N = \frac{15}{0} \]
\[ \text{mean } p_1 = ? \]

\[ \text{mean } p_1 = 49.3\% \]

\[ \text{mean } \hat{p}_1 = 49.3\% \]

\[ \text{mean } \hat{p}_1 = 7 \text{ (50\%)} \]

\[ E_{III}(\hat{p}_1) = p_1 \]

\[ \text{S.E.}_{III}(\hat{p}_1) = \sqrt{\frac{p_1(1-p_1)}{n_1}} = \sqrt{\frac{(0.493)(0.507)}{592}} = 0.021 = 2.1\% \]

\[ \text{S.E.}_{2.1\%} \]

\[ \text{Confidence interval for } \hat{p}_1 \text{ of } p_1 \]
\[ N^2 = \begin{bmatrix} 15 \\ 15 \\ \text{or} \\ 15 \end{bmatrix} \]

\[ 1:1:2:5 \]

\[ \text{let} \ n_2 = 154 \]

\[ \text{mean } \hat{p}_2 = 48.4\% \]

\[ \text{log } n_2 \hat{p}_2 \text{ E}_{\text{IIID}}(\hat{p}_2) = \hat{p}_2 \]

\[ \text{est} \]

\[ \text{SE}_{\text{IIID}}(\hat{p}_2) \]

\[ \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \]

\[ = \sqrt{\frac{0.484(0.516)}{154}} = 4.0\% \]

---

**Inferential Summary**

<table>
<thead>
<tr>
<th>Quantity of Interest</th>
<th>( (\hat{p}_1 - \hat{p}_2) = \text{diff. in receipt rate between } ) &amp; ( ) in pop</th>
</tr>
</thead>
<tbody>
<tr>
<td>estimate</td>
<td>( (\hat{p}_1 - \hat{p}_2) = 49.8% - 48.4% = 0.4% )</td>
</tr>
<tr>
<td>practice</td>
<td>probably too small to matter practically</td>
</tr>
<tr>
<td>give or take for art.</td>
<td>( \text{SE}_{\text{IIID}}, \text{indiv} (\hat{p}_1 - \hat{p}_2) = 4.5% )</td>
</tr>
</tbody>
</table>
\[ \text{SE}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 4.5\% \]

\[ 95\% \text{ CI for } (\hat{p}_1 - \hat{p}_2) = (\hat{p}_1 - \hat{p}_2) \pm 2 \times \text{SE}(\hat{p}_1 - \hat{p}_2) = 0.9\% \pm 2(4.5\%) = 0.9\% \pm 9.0\% \]

\[ -8.1\% < (\hat{p}_1 - \hat{p}_2) < 9.9\% \]

null: \( (36.1\% \text{ not just due to unlucky sampling}) (\hat{p}_1 - \hat{p}_2 = 0) \)

since null value of 0 is (comfortably) inside our 95\% int, null looks ok:

not statistically significant

(b) (2 independent samples, continuous outcome)

like CS15 (skeletons)

(1) \( \bar{Y}_1 (\bar{T}) = 16.8 \text{ weeks} \)
\( n_1 = 592 \)

(2) \( \bar{Y}_2 = 24.3 \text{ weeks} \)
\( n_2 = 173 \text{ weeks} \)
\( s_1 = 15.8 \text{ weeks} \)
\( s_2 = 17.3 \text{ weeks} \)
mode has identical structure to (9) but with cont. outcome instead of 1/0

<table>
<thead>
<tr>
<th>Quantity of interest</th>
<th>((M_1 - M_2)) = (\text{pop. mean difference in weeks of paid work})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
<td>((\bar{y}_1 - \bar{y}_2) = (-7.5)) weeks</td>
</tr>
<tr>
<td>Practice (2, z)</td>
<td>(\frac{\bar{y}_1 - \bar{y}_2}{z} = \frac{16.8 - 24.3}{24.3} = -2.5 = -31%)</td>
</tr>
<tr>
<td>In other words, mean weeks of work for (\odot) people was 31% lower than for (\odot) practicing.</td>
<td></td>
</tr>
<tr>
<td>Give or take for (\pm).</td>
<td>(\hat{SE}_{II}, \text{indep} (\bar{y}_1 - \bar{y}_2) = 1.5 \text{ weeks})</td>
</tr>
</tbody>
</table>

\[
\hat{SE}_{II}, \text{indep} (\bar{y}_1 - \bar{y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{15.9^2}{592} + \frac{17.3^2}{154}} = 1.5 \text{ weeks}
\]

95\% CI for \((\mu_1 - \mu_2):\)
\((\bar{y}_1 - \bar{y}_2) \pm 2 \times \hat{SE} (\bar{y}_1 - \bar{y}_2) = \left[-7.5 \pm 2(1.5)\right] \text{ weeks}
= (-7.5 \pm 3.0) \text{ weeks}
\[ \text{statistic} = \frac{(\bar{Y} - \bar{X})}{\sqrt{\frac{ss_{xy}}{n}}} \]

0 is not in 95% CE. So null looks wrong: statistic yes

1 (calculator)

to review paired comparisons
for final: cs 14
p. 93 (discount pricing)

\[ \text{pop} \] all underground riders at hour
\[ \text{sample} \] the observed people

\[ N = \text{huge} \] like 100

\[ \text{weight} \]
\[ \text{SD} \sigma = 30 \text{ lb.} \]
\[ \text{mean} \mu = 155 \text{ lb.} \]

\[ \text{pop hist} \]
\[ \text{hist} \]

\[ \text{sum} s' = ? \] (ex. 31400 lb.)
\[ \text{sum} S = ? \] (ex. 31400 lb.)

\[ \text{like hist} \]
\[ \text{like hist} \]

\[ \text{long red} \]
\[ \text{SE}_{SS} (s') = \frac{\text{long red}}{\text{long red}} \]

\[ \text{SE}_{SS} (s') = \frac{31400}{200} = 157 \text{ lb.} \]
\[ P(\text{calculator break}) = P(S > 32,000) \]
\[ E_{\text{IID}}(S') = \frac{1}{4} = 200 \times (155 1/6) = 31,000 \text{ lb} \]
\[ SE_{\text{IID}}(S') = 5 \sqrt{n} = 30 \sqrt{200} = 424 1/6 \text{ lb} \]

\[ SE = 424 1/6 \text{ lb} \]

1% chance of failure

(6) This is very too often to fail: \( \frac{1}{100} \text{ trips/ day} \)

Every day (900 ft dry trips/day)

(8) 5% 424 1/6

99.98% tail of S

\[ \frac{32,000 - 31,000}{424 1/6} = 2.35 \]

Low

7100 1/6

X

\[ \frac{7100}{424 1/6} = 3.225 \]
\[ x = 31,000 + \frac{3.725}{424} \approx 32,579 \text{ lb}. \]

Small changes in engineering lead to large decreases in failure rate.