Solutions to Review Exercises

1. For each of the claims that follow, choose which of the following tests would be most appropriate.

   (a) One sample z/t test
   (b) Two sample z/t test
   (c) Paired difference test
   (d) One sample proportion test
   (e) Two sample proportion test
   (f) Analysis of variance
   (g) Chi-square goodness-of-fit t-test

   (i) (d) - A one sample proportion test is appropriate for testing if more than half of all babies born are boys.
   (ii) (e) - A two sample proportion test is appropriate for testing if the probability a student will pass Math 121 is the same as the probability a student will pass Math 134.
   (iii) (b) - A two sample z or t test is appropriate to test if the average cost of a pack of cigarettes is higher in Canada than in the U. S.
   (iv) (b) - A two sample z or t test is appropriate to test if smokers have a shorter life expectancy than non-smokers.
   (v) (b) - A two sample z or t test is appropriate to test if gas prices in Modesto are the same as in Sacramento.
   (vi) (f) The Analysis of Variance is the most appropriate test to determine if the average GPA is the same for students in each of the seven Claremont colleges.
   (vii) (g) - A chi-square goodness of fit test is most appropriate to determine if a coin is fair.
   (viii) (a) - A one sample z or t test is most suited to test if the average rent for a one-bedroom apartment in Modesto is $800.

2. Infant mortality.

   (a) Sketch the scatterplot.

   \[
   \begin{array}{c|c}
   \text{WEIGHT} & \text{AGE} \\
   \hline
   \bar{x} = 90 \text{ oz.} & \bar{y} = 50 \text{ days} \\
   s_x = 15 \text{ oz.} & s_y = 10 \text{ days} \\
   r = -0.50 \\
   \end{array}
   \]

   SD line \( \left\{ \begin{array}{l}
   \text{Slope} \quad (\text{sign of } r) \frac{s_y}{s_x} = -\frac{2}{3} \text{ days/oz.} \\
   \text{y-int} \quad \bar{y} - b \cdot \bar{x} = 110
   \end{array} \right. \)

   Regression line \( \left\{ \begin{array}{l}
   \text{Slope} \quad r \cdot \frac{s_y}{s_x} = -\frac{1}{3} \text{ days/oz.} \\
   \text{y-int} \quad \bar{y} - b \cdot \bar{x} = 80
   \end{array} \right. \)
(b) The age at which one of those very premature babies that weighs only 4 pounds at birth first sleeps through the night is \( \hat{y} = -1/3(64) + 80 \approx 58.7 \text{ days} \). The age at which one of those healthier babies that weighs 8 pounds at birth first sleeps through the night is \( \hat{y} = -1/3(128) + 80 \approx 37.3 \text{ days} \). The standard error for both predictions is \( SE(\hat{y}) = \sqrt{1 - r^2} \cdot s_y \approx 8.7 \). In practical terms, the difference of \( 58.7 - 37.3 = 21 \text{ days} \) until a baby first sleeps through the night is quite large.

3. (law) In 1969, the well-known pediatrician Dr. Benjamin Spock came to trial before a judge name Ford in Boston’s Federal court house. He was charged with conspiracy to violate the Military Service Act (in addition to his work on child development he was active in anti-war protests in the 60’s). A lawyer writing about the case that same ear in the Chicago Law Review said about the case, “Of all defendants at such trials, Dr. Spock, who had given wise and welcome advice on child-rearing to millions of mothers, would have liked women on his jury.”

The jury was drawn from a panel of 350 persons, called a venire, selected by Judge Ford’s clerk. This venire included only 102 women, even though 53% of the eligible jurors in the district were female. At the next stage in selecting the jury to hear the case, Judge Ford chose 100 potential jurors out of these 350 people. His choices included only 9 women.
#31a

Population

All eligible jurors in district

$P = 53\%$

$\sigma = \text{SE} = \sqrt{\frac{0.53(1-0.53)}{350}}$

Sample

observed venue

woman?

woman?

n = 350

$\hat{P} = \frac{102}{350} = 29\%$

$\text{SE} = 2.7\%$

$P(\hat{P} \leq 102 )$

Since $z = 8.8$ which is off the charts, $P(\hat{P} \leq 102 )$ is about 0%

[B.]

population

Venire

woman:

n = 350

$P = \frac{102}{350} = 29.1\%$

$\text{SE} = \sqrt{\frac{(0.29)(1-0.29)}{100}} = 0.045$

$P(\hat{P} \leq 9\%)$ is about 0%

[C.]

Judge Ford was not choosing at random; he was excluding women.
4. A newspaper article says that on average, college freshmen spend 7.5 hours a week going to parties. One administrator does not believe that these figures apply at her college, which has nearly 3,000 freshmen. She takes a simple random sample of 100 freshmen, and interviews them. On average, they report 6.6 hours a week going to parties, and the SD is 9 hours.

(a) Formulate the null and the alternative hypothesis.

\[ H_0 : \mu = 7.5 \text{ hr/week} \]
\[ H_a : \mu < 7.5 \text{ hr/week} \]

The administrator believes that college freshman do not spend 7.5 hours a week on average, going to parties; presumably, she believes that college freshman at her college spend fewer hours per week on average in comparison to what was reported by the newspaper article.

(b) Is the difference between her students and the newspaper report significant?

We set up the population as if the Null hypothesis were true, so that \( E(\bar{x}) = \mu = 7.5 \) and \( SE(\bar{x}) = \sigma/\sqrt{n} \approx \frac{9}{\sqrt{100}} = 0.9 \) hrs/wk. The z-stat = \( \frac{6.6 - 7.5}{0.9} = -1 \) and according to the Empirical Rule the P-value is approximately \((100 - 68)/2 = 16\%\).

(c) What is your conclusion?

The evidence against the Null hypothesis is not strong, and we accept \( H_0 \). Thus, the difference between the sampled difference in average time spent going to parties for freshman at her school is not significantly different from the reported average of 7.5 hours a week in the newspaper article.

5. \( P(10 \text{ heads out of } 10 \text{ tosses}) = (0.5)^{10}. \)

6. \( P(1 \text{ head in } 1 \text{ more toss}) = 1/2 = 50\% \)

7. \( P(\text{above average for } 2 \text{ consecutive years}) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}. \)
   \[ P(\text{above average for } 4 \text{ consecutive years}) = \left(\frac{1}{2}\right)^4 = \frac{1}{16}. \text{ since the chance of sales growth falling above or below the median does not depend on the prior year’s sales.} \]

8. \( P(\text{income } \geq \text{ $5 million}) = \frac{35,207}{134,372,678} = 0.026\% \)
   \[ P(\text{income } \geq \text{ $5 million} | \text{income } \geq \text{ $1 million}) = \frac{35,207}{303,817} \approx 11.6\% \]

9. Read the Background, Methods, Results and Conclusions sections of the paper “Study of the Therapeutic Effects of Intercessory Prayer (STEP) in cardiac bypass patients; A multi-center randomized trial of uncertainty and certainty of receiving intercessory prayer” attached to the back of this exam.

(a) What are the hypotheses the researchers studied?

Whether (1) receiving intercessory prayer was associated with uncomplicated recovery after CABG surgery and
(2) being certain of receiving intercessory prayer was associated with uncomplicated recovery after CABG surgery.

(b) Formulate these as null hypotheses. What are the corresponding alternative hypotheses?

Let Group T represent those who received intercessory prayer, but did not know and Group C represent those that did not receive intercessory prayer, but did not know and Group TT represent those who received and knew that they received intercessory prayer.

(1) \( H_0 : p_T = p_C \) and \( H_a : p_T \neq p_C \)

(2) \( H_0 : p_{TT} = p_T \) and \( H_a : p_{TT} \neq p_T \)

(c) Was this an observational study or a controlled experiment? Patients at 6 US hospitals were RANDOMLY assigned to 1 of the three groups by the researchers, therefore, this was a randomized controlled experiment.

(d) What do the researchers conclude?

The researchers concluded that intercessory prayer itself had no effect on complication-free recovery from CABG, but the certainty of receiving intercessory prayer was associated with a higher incidence of complications.

(e) The investigators arranged for prayers to be said for those in the study who were assigned to the treatment groups, but they could not control whether others (e.g. family members of the patients) were saying prayers. Could this have affected the conclusions? No, because family members could have been praying for those who were in the Treatment groups, as well as those who were in the Control groups.

(f) From the data regarding the groups

(1) who were prayed for but were uncertain of whether they were being prayed for, and
(2) who were prayed for and knew they were being prayed for,

Can you conclude that there was a difference in complications between the two groups? The null hypothesis for this test is: \( H_0 : p_T = p_{TT} \) and \( H_a : p_T \neq p_{TT} \)

\[
SE(\hat{p}_T)^2 = \frac{352/601(1-352/601)}{601} \\
SE(\hat{p}_{TT})^2 = \frac{315/604(1-315/604)}{604}
\]

The Standard Error for the significance test is given by:

\[
SE(\bar{d}) = \sqrt{\hat{p}(1-\hat{p}) \left( \frac{1}{601} + \frac{1}{604} \right)} \text{ where} \\
\hat{p} = \frac{352 + 315}{601 + 604} = \frac{667}{1205} \approx 0.554
\]

The test statistic is:

\[
z = \frac{\bar{d} - 0}{SE(\bar{d})} = \frac{\left( \frac{352}{601} - \frac{315}{604} - 0 \right)}{0.029} \approx 2.21
\]
The corresponding $P$-value is $2(1 - 0.9864) = 0.0272$, which is strong evidence against the Null hypothesis, so we reject the Null in favor of the alternative hypothesis. Thus, the number of complications after CABG surgery for those patients certain of receiving intercessory prayer was greater, on average, than those patients who were uncertain and received intercessory prayer. These results led researchers to state that certainty of receiving intercessory prayer was associated with a higher incidence of complications.

(g) Explain why the test you performed in part(f) is valid in this case.
The $z$-test can be used to compare the treatment and control groups in a randomized controlled experiment, even though the groups are not independent. (The treatment and control groups in an randomized controlled experiment are dependent, because if a participant is in the treatment group we KNOW that theyre not in the control group.)