Math/Stat 3850 – Exam 2 Practice

There are 10 questions, worth a total of 100 points.
You may use any calculator, R and any internet resources for this exam, although you
are not allowed to ask someone else for help.

(10) 1. Each of these studies has won the prestigious Ig Nobel prize.
For each, state the null hypothesis.

(a) Marina de Tommaso, Michele Sardaro, and Paolo Livrea, for measuring the relative
pain people suffer while looking at an ugly painting, rather than a pretty painting,
while being shot [in the hand] by a powerful laser beam.

Solution: The quality of the painting has no effect on the pain people suffer.

(b) Atsuki Higashiyama and Kohei Adachi, for investigating whether things look dif-
ferent when you bend over and view them between your legs.

Solution: Things don’t look any different when you bend over and view them
between your legs.

(c) Patricia Yang, David Hu, Jonathan Pham, and Jerome Choo, for testing the bio-
logical principle that nearly all mammals empty their bladders in about 21 seconds.

Solution: All mammals empty their bladders in about 21 seconds.

(10) 2. Mammograms are usually read by two different clinicians, but it would be more efficient
to use one clinician plus computer aided detection. A study in NEJM reported:

The proportion of cancers detected was 199 of 227 (87.7%) for double reading and
198 of 227 (87.2%) for single reading with computer-aided detection (P=0.89).
The overall recall rates were 3.4% for double reading and 3.9% for single reading
with computer-aided detection; the difference between the recall rates was small
but significant.

(a) Is there a significant difference between the detection rates?

(b) When the clinicians see something on the mammogram, they recall patients for a
follow-up. What was the difference between the recall rates?

(c) How could such a small difference be significant?

Solution: a. No. b. 0.5%. c. There were enough people in the study to make
the probability of observing such a difference small, given the null hypothesis of
no difference.
3. Consider the confidence interval for the mean of a variable. For each part, say if the confidence interval gets wider, gets narrower, or stays the same:

(a) The sample size $n$ increases.
(b) The sample standard deviation $s$ increases.
(c) The sample mean $\bar{x}$ increases.
(d) The level of confidence $C$ increases.

**Solution:** (a) narrower (b) wider (c) same (d) wider

4. In 1798, Henry Cavendish measured the density of the Earth using a torsion balance. His measurements are the first column (density) in the dataset Cavendish in the HistData library. Each measurement is the density of the earth as a multiple of the density of water.

(a) State and carry out a hypothesis test that his results are significantly different from the modern accepted density of 5.517 g/cm$^3$.

(b) The third value, 4.88, was later discovered to be incorrect. Cavendish actually measured 5.88 and copied it wrong. Change the value to 5.88, and run your test again. How does this affect the results?

**Solution:** (a) $t = -1.6834$ with 28DF. His results are not significantly different from the modern value ($P = 0.1034$). (b) Now $P = 0.3753$ and his results differ even less significantly from the modern value.

5. In 1876, Darwin studied a variety of plants to compare cross-fertilized plants with self-fertilized plants. In one experiment, he planted thirty pairs of “candytuft” flowers, putting one self- and one cross-fertilized plant together in a pot. When the plants reached maturity, he measured their heights in inches. This data is on our web page as candytuft.csv.

(a) Which type of plant is taller, on average?

(b) State a hypothesis test that the mean heights of self- and cross-fertilized plants are different.

(c) Perform the test and report the results.

**Solution:** Using a matched pairs t-test gives $t = 3.0227$ with 29DF and $P = 0.005196$. Cross-fertilized plants are significantly taller.

(Independent samples $P = 0.001163$ is incorrect.)
6. (a) Look at the candytuft data from question 5 with appropriate graphical or summary methods. How comfortable are you using a t-test on this data? Why or why not? (b) Test for a difference in plant heights using the Wilcoxon test and report the P-value. Does this change your conclusions from question 5?

Solution: There are two outliers in the self-fertilized plant measurements. One is a bit high, probably ok. The other is a value of 0 (in fact, the plant died), and is a major outlier. Using a t-test with this outlier is questionable. It is less of a problem with Wilcoxon, which reports $P = 0.0004877$. The conclusions don’t change.

(Unpaired test gives $P = 0.0001002$, which is the wrong thing to do.)

The next two problems use the dataset oly12 from the VGAMdata package that you will probably need to install. It has individual competitor information from the Summer 2012 London Olympic Games.

7. According to this data, which country won the most medals? How many did that country win? (You need to sum Gold, Silver, and Bronze)

Solution:

```
oly12 %>% group_by(Country) %>%
  summarize(medals = sum(Gold, Silver, Bronze)) %>%
  arrange(desc(medals))
```

The US won the most medals, with 79 (according to this data - in fact, Wikipedia has it at 103).

8. Which country is the heaviest? Compute the mean weight of male athletes for all countries with at least 10 competitors, and report the top three.

Solution:

```
oly12 %>% group_by(Country) %>%
  filter(!is.na(Weight) & Sex == "M") %>%
  summarize(nc = n(), mw = mean(Weight)) %>%
  filter(nc >= 10) %>%
  arrange(desc(mw))
```

# A tibble: 97 x 3
The next two problems use the data set DrinksWages from library(HistData). This data, gathered by Karl Pearson in 1910, was a survey of people working in various trades (bakers, plumbers, goldbeaters, etc.). The trades are assigned class values of A, B, or C based on required skill.

For each trade, he counted the number of workers who drink (drinks), number of sober workers (sober), and recorded wage information (wage). There is also a column n = drinks + sober which is the total number of workers surveyed for each trade.

9. (10) Consider this dplyr pipeline:

DrinksWages %>% group_by(class) %>% summarize(mean(wage))

What question does it answer? Explain.

**Solution:** This computes the mean wages for each class (A, B, or C) of worker.

10. (10) Find the three trades with the highest proportion of drunks (drinks/n). Consider only trades with 10 or more workers in the survey.

**Solution:** DrinksWages %>% filter(n >= 10) %>% arrange(desc(drinks/n)) gives cabmen, tailors, and mason as the top three drunkest trades.