

Exam 1
Psych 2111, Fall 16

Vocabulary

1. Random assignment: A property of good experiment design, whereby members of the sample are randomly assigned to different levels of the independent variable
2. Random selection: A property of good experiment design, whereby every member of the population has an equal chance of being in the sample
3. Descriptive statistic: A statistic that is being used only for describing properties of the sample, without making inferences about a population
4. Discrete variable: A variable that can take on only a discrete (i.e., spaced) set of possible values
5. Percentile: A value for a variable that is determined by being greater than some specified percentage of the scores in a distribution

Conceptual questions

1. When I was a grad student at the University of Michigan, there were 11 people named Matt Jones. I always wondered whether that was the most common name on campus. What parameter was I interested in?

Mode

2. Give examples of variables that are defined on each of the following scales.

Nominal: Favorite vegetable

Ordinal: Level of education

Interval or ratio (circle one): Time to run a 100-meter dash

3. The personality measure for neuroticism you're developing has a mean of 50 and a standard deviation of 10. To make the scores more comparable to other measures in the literature, you convert them to z-scores. Now you have a standardized distribution. What will the mean and standard deviation of this distribution be?

Mean: 0

Standard deviation: 1

Every standardized distribution (i.e., distribution of z-scores) has a mean of zero and a standard deviation of one.

4. Write whether each of the following studies qualifies as an experiment (yes or no).

Subjects are first given a survey assessing their current level of depression. They are divided into two groups, Depressed and Not Depressed. Then everyone is given crossword puzzles, to test whether the mean completion time is faster for the Not Depressed group.

No (depression was not manipulated)

A set of rats is divided into two groups. The rats are all shown pictures of cats. Every time a cat is shown, rats in the first group are given a drop of chocolate milk. Rats in the second group see the pictures with no chocolate milk. Then all the rats are released into a room with a real cat, and the researchers count how many rats from each group are eaten within 15 minutes.

Yes (milk delivery is manipulated)

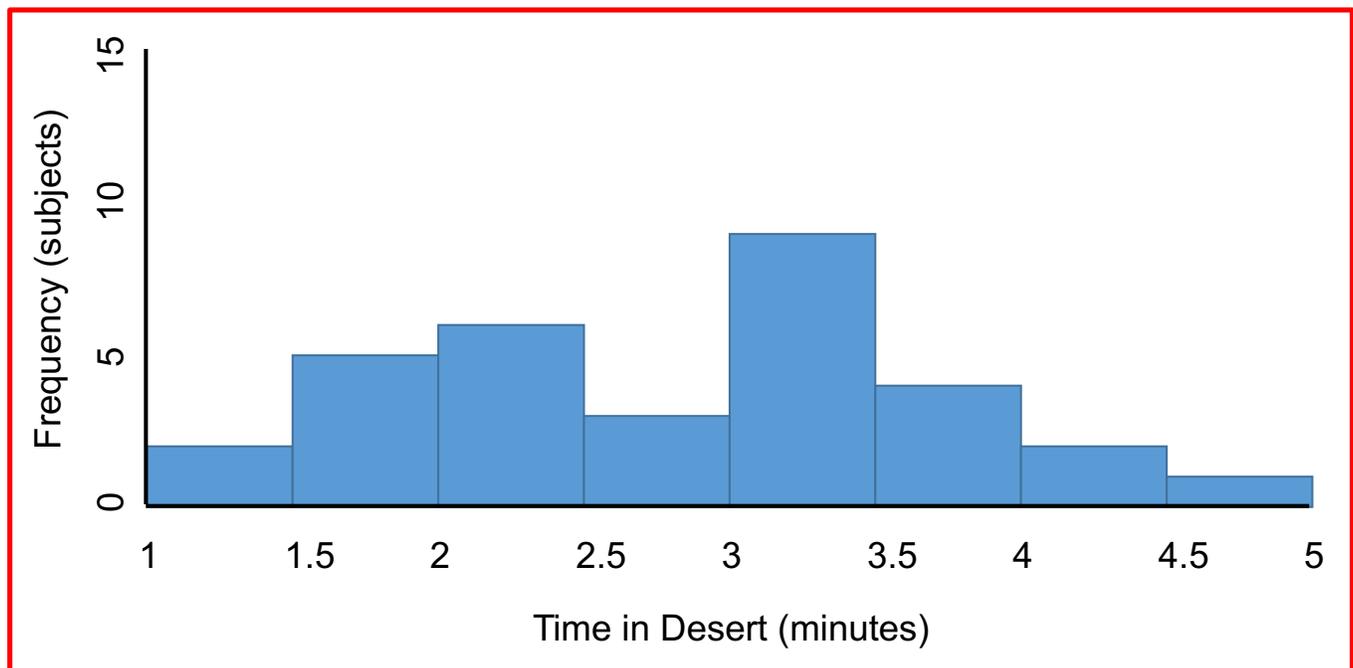
Human subjects are released individually in a remote desert, with nothing but a compass and a map. The researchers measure each subject's time to reach the nearest town, and then calculate the mean and median times of their sample.

No (sex is not manipulated, although it can be considered a quasi-independent variable)

5. Here's a stem-and-leaf plot of the data from the desert study, in hours. Make a histogram for this dataset, using a bin width of half an hour.

```
1 | 3467889
2 | 112344679
3 | 1122344446669
4 | 138
```

The stem-and-leaf plot shows 32 values: 1.3, 1.4, 1.6, ..., 4.8.



Math questions

You don't need to show your work, but you can receive give partial credit for partial answers.

1. You ask a sample of 50 people how many siblings they have. Then you use your data to calculate frequencies (f) and cumulative frequencies (F). If $F(2) = 41$, $F(4) = 49$, and $f(4) = 3$, then how many people have exactly three siblings?

The number of people with at most two siblings is 41, and the number with at most four is 49. Therefore, the number of people with either three or four siblings is $49 - 41 = 8$. Of those 8 people, 3 have exactly four siblings, so the other $8 - 3 = \underline{5}$ have exactly three.

2. Calculate the standard deviation of [35, 41, 39, 38, 42], as a sample.

$$M = \frac{\Sigma X}{n} = \frac{35 + 41 + 39 + 38 + 42}{5} = 39$$

$$SS = \Sigma(X - M)^2 = (35 - 39)^2 + (41 - 39)^2 + (39 - 39)^2 + (38 - 39)^2 + (42 - 39)^2 = 30$$

$$s^2 = \frac{SS}{n - 1} = \frac{30}{4} = 7.5$$

$$s = \sqrt{7.5} \approx \underline{2.7}$$

3. Convert the data in Question 2 to a standardized distribution.

$$z = \frac{X - M}{s} \approx \frac{[35, 41, 39, 38, 42] - 39}{2.7} \approx \underline{[-1.5, .7, 0, -.4, 1.1]}$$

4. Imagine I chose to grade this question randomly, without looking at your answer. With 70% probability, I'll give you 5 points (out of 5). With 10% probability, I'll give you 4 points. With 20% probability, I'll give you 3 points. What's the expected value for your score?

$$E(R) = \Sigma_x(x \cdot p(x)) = 5 \cdot 70\% + 4 \cdot 10\% + 3 \cdot 20\% = \underline{4.5}$$

5. My 10-year-old son is 4'2", which gives him a z-score of -2.5 for boys his age. If the mean height for 10-year-old boys is 4'7", what is the standard deviation?

$$z = \frac{X - \mu}{\sigma}$$
$$-2.5 = \frac{4'2'' - 4'7''}{\sigma} = \frac{-5''}{\sigma}$$

$$\sigma = \frac{-5''}{-2.5} = \underline{2''}$$

R questions

1. List three R functions, showing an example command you would enter (don't write just the name of the function) and a verbal description of what the command does.

| | Command | Description |
|---------|---------------------------------------|---|
| Example | <code>> mean(X)</code> | computes the mean of a vector X |
| 1 | <code>> log(X)</code> | computes the natural logarithm of the value(s) of X |
| 2 | <code>> data.frame(X, Y, Z)</code> | creates a data frame containing the variables X, Y, Z |
| 3 | <code>> pnorm(X)</code> | computes the probability of getting a score $\leq X$, in a standard Normal distribution |

2. What statistic is being computed by the following R commands? (It may help to write what `step1` and `step2` are next to the first two lines.)

```
> step1 = sum(X)/length(X)    mean(X)
> step2 = (X - step1)^2       Squared deviations
> answer = sum(step2)         Sum of squared deviations
```

Statistic: Sum of Squares, SS

3. What is the result of the following commands?

```
> X = 7:11                    X = [7 8 9 10 11]
> X[1] + X[3]                 X[1] = 7, X[3] = 9
```

16

4. What is the result of the following computations?

```
> X = c(34, 68, 42, 94, 32)
> Y = c("male", "male", "female", "female", "male")
> Z = X[Y=="female"]         X[FALSE FALSE TRUE TRUE FALSE]
→ Z = [42 94]
```

5. What is the result of the following command? Remember that the Standard Normal distribution is centered on 0. The answer is a number, but if you can't figure out the exact result, try describing in words what this command does, or draw a picture.

```
> pnorm(0)
```

Probability of a score ≤ 0 , which is $\frac{1}{2}$

