January 29, 2014
Quiz 2

Your name: Solutions

Quiz rules:

(a) This quiz is closed book, but you are allowed a two-sided sheet of paper of notes and a calculator.
(b) Each question is worth 6 points.
(c) A normal table is provided on the last page.
(d) You have 50 minutes to complete this quiz.
(e) If you fail to show your work and/or explain how you arrived at your answer then no points will be awarded.

1. In a large class, the average score on the midterm was a 60% with an SD of 10%. The average score on the final was also a 60% with an SD of 20%. The correlation coefficient $r = 0.7$
   (a) Predict the final exam score of a student who scored a 70% on the midterm.

(1) Convert X into Z_x

$$\frac{70\% - 60\%}{10\%} = 1$$

(2) $r * Z_x = Z_y$

$$0.7 * 1 = 0.7$$

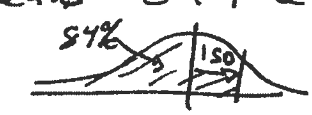
(3) Convert Z_y into Y

$$.7 * 20\% + 60\% = 74\%$$

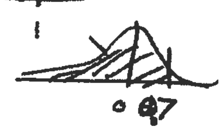
Note: Predict that people who got 70% on midterm will get 74% on final.

   (b) Does your answer to (a) contradict the regression effect? Explain why or why not.

Does not contradict regression effect. The regression effect is most easily seen on the Z-score scale. Note $Z_x = 1$ but $Z_y = 0.7$. This means on the midterm, people who scored at the



84th percentile are predicted to score at the



75.8th percentile. Note: It may appear to defy regression effect only because X and Y are of the same scale and have the same average. The differ[...] SDs muc[...] things u[...]

2. To study the effect of a new dog food on the growth of puppies, researchers run a randomized study. The researchers recruit 18 dogs – six large (Bernese Mountain Dogs), six medium (English Bulldog) and six small (West Highland Terrier). If birth is considered "week 0," the researchers are interested in the amount of weight the dogs gain from week 6 to week 52.

   The researchers randomly assigned half of the dog to treatment (i.e., the new dog food) and half to the control (i.e., the usual dog food). The table on the next page summarizes the total weight of the dogs at baseline (at week 6) and final (at week 52).

January 29, 2014
Quiz 2

		Baseline (total lbs)	Final (total lbs)
Treatment Group	• 4 Bernese • 4 Bulldogs • 1 Westies	33	614
Control Group	• 2 Bernese • 2 Bulldogs • 5 Westies	21	406

(a) In this randomized experiment, the treatment group increased its total poundage by 581lbs (=614-33). The control group increased by 385lbs. The treatment group increased by 196lbs more than the control (581lbs-385lbs=196lbs). Given the design of the study, can you attribute that difference in increases to the difference in food? If so, identify the features of the study which allow you to do this. If not, identify alternative explanations for the observed difference. In either case, be as clear as possible and use statistical concepts to illuminate your argument.

Best: Can not attribute this solely to food, because of a potential confounder. Breed type has an impact on the outcome of interest, weight. It is also imbalanced between the treatment and control. More big and medium dogs are in the treatment and small dogs in the control. Breed type is a highly plausible alternative explanation for the difference in the change in weights.

OK: Randomization occurred, which means in general no confounders caused bias in a systematic way. But in this particular case there was a chance imbalance leading to confounding.

(b) Regardless of how you answered in part a, there is ~~an~~ improvement to how the researchers assigned the dogs to the treatment/control. Identify this improvement and discuss how it will help the study.

Improvement: Pair-match dogs on characteristics at baseline (e.g, breed type, weight at week 6). Then randomize one of the dogs within the pair to treatment and the other to control.

How this helps: If we do this improved design, we will get 3 large 3 medium and 3 small dogs in both treatment and control every time. This forces the randomly assigned groups to be more similar than the current procedure allows.

3. In baseball, one measure of a pitcher's performance is ERA, the number of runs surrendered per game pitched. For example, a pitcher who gives up 43 runs in 10 games would have ERA = 43/10 = 4.3. A lower ERA is better.

Consider two pitchers, Darvish and Tanaka. Suppose Darvish has ERA = 1.0 and Tanaka has ERA = 2.0 in the first half of the season. Suppose also that in the second half, Darvish has ERA = 3.0 and Tanaka has ERA = 4.0. Which of the following is true? Choose one.

- Darvish must have the higher ERA for the entire season.
- Tanaka must have the higher ERA for the entire season.
- Either Darvish or Tanaka could have higher ERA for the entire season.

Use numerical examples to illustrate your choice. (Hint: A typical pitcher pitches about 20-30 games in a season if healthy, but fewer if injured.)

This is an example of Simpson's Paradox.

Tanaka higher ERA

	1st half		2nd half		overall
	ERA	Games	ERA	Games	
Darvish	1	10	3	10	2
Tanaka	2	10	4	10	3

Darvish higher ERA

	1st half		2nd half		overall
	ERA	Games	ERA	Games	
Darvish	1	1	3	15	≈2.88
Tanaka	2	20	4	2	≈2.18

Intuition: Seems counterintuitive until you realize they don't have to have played equal number of games. Usually people default to the assumption of equal base rates.

4. Stanford biologists H. Craig Heller and Dennis Grahn developed a "cooling glove" that can rapidly cool body temperatures. There are several applications for this glove, one of which is to improve athletic performance. During exertion athletes generate excessive amounts of heat. Excessive body heat reduces performance – primarily through limiting endurance. (This is real! Try Googling "cooling glove stanford.")

Imagine that Stanford's athletic director is considering rolling out this technology for use by all of Stanford's varsity athletic teams. He wants to know if this glove will help Stanford's varsity athletes. The athletic director commissions a randomized study of the effect of the glove. Suppose the researchers recruit 30 male athletes from three of the high-endurance men's varsity sports teams (i.e., cross country, rowing, soccer). On each team, half of the men will receive the treatment (use of the glove during competition) and the other half will be assigned to control (normal playing conditions). The researchers then perform a reasonable randomized study, being careful to have similar athletes in the treatment and control groups. On all measurements of performance, the treatment group outperforms the control group.

(a) What features of this study qualify it as an experiment rather than an observational study? List at least two features.

Really important: (1) Randomization because it will, in general, create equivalent groups. This is important because it addresses known variables as well as unknown. (2) Researcher controlled design. People intentionally sought out participants, created a protocol and considered potential confounders in design of study.

OK to have: Prospective (instead of retrospective) study, measurement of baseline characteristics, sample from defined population.

(b) You may assume this study was properly run (e.g., no mismeasurements occurred, the researchers measured the right kinds of outcomes both for performance and increased risk). Nonetheless, this study is inadequate. It is not capable of directly answering the athletic director's question. There are at least two features of the way this study was designed that limits its usefulness. Identify at least one of these issues and discuss its impact.

Big issues (1) Varsity athletes are what we're interested in. But they only experimented with men. This completely misses women, who may react differently to the glove.
(2) Not all athletes are high-endurance (e.g. shot putter). But researchers only sampled from high-endurance sports. Effect may be quite different for other types of athletes.

Ok, but not great: Issues due to placebo or belief-in-effect from lack of blinding. This may occur, but the Athletic Director doesn't care if people are performing better because of physiology or psychology (or both!).

5. Out of all the states in the US, Montana has the highest temperature variability over the course of the year. The capital of Montana – Helena – has an average annual temperature of 43.9°F and an SD of 16.5°F. Most other states have SDs that are less than 10°F.

A weatherman would like to predict temperatures in Helena. He has heard that he can use one variable to predict another. He decides to try this out. First he collects the average temperature for each month in Helena. He then decides to use month (Jan = 1, Feb = 2, ..., Dec = 12) to predict temperature.

Before running a regression he runs a linear correlation between month and temperature, because he's heard correlation will show him "if there's any connection between X and Y." He finds that the correlation between month and temperature is -0.08.

(a) What does this correlation say about the association between month and temperature?

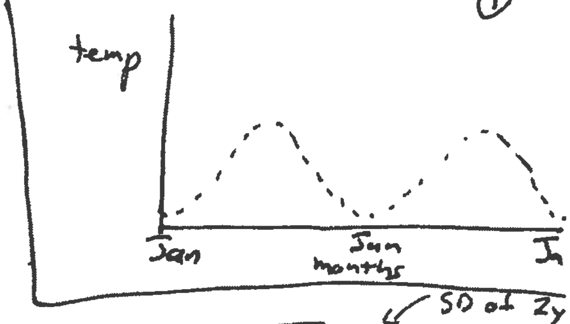
It means there is little linear association, because −0.08 is quite close to zero.

But: given what we know about temperatures across months, the plot probably looks something like:

. This is not linear, but there is a very high association.

(b) Do you think month can be used reliably to predict temperature? Explain your thinking by using statistical terminology.

Linear correlation is the wrong measurement of association here (see part a). Month is probably a good predictor of temperature — January this year tends to be a lot like January in prior years, but quite unlike July temperatures. Maybe with Global Warming things are changing from year-to-year (i.e., a secular trend). But mostly we see a strong cyclic association.



6. Stats 60 students study an average of 10 hours a week, with an SD of 2 hours. The correlation between hours spent and exam performance is 0.30. What percentage of Stats 60 students who study 7 hours a week would you expect to score below average on an exam? Or is there not enough information? (If not enough information, please make as precise a statement about the exam performance of these students as possible.)

If we assume the scatterplot is football shaped, then we have enough information to get percentiles. We couldn't predict score on the exam.

(1) Convert X in Z_x

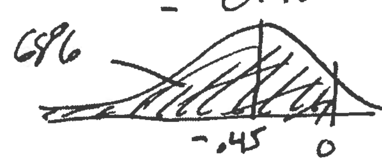
$$Z_x = \frac{7-10}{2} = -1.5$$

(2) $r \cdot Z_x = 0.30(-1.5) = -0.45 = Z_y$

Now find RMSE to find spread

$$RMSE = \sqrt{1-0.3^2} \cdot 1 \approx 0.95$$

SD of Z_y



$$Z = \frac{0-(-0.45)}{0.95}$$

$$Z = 0.47$$

From Z-table: $\approx 68\%$