[redacted]

STATISTICS
SECTION II
Part A
Questions 1-5

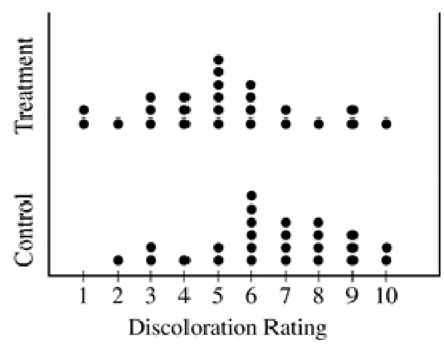
Spend about 65 minutes on this part of the exam.
Percent of Section II grade—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. The department of agriculture at a university was interested in determining whether a preservative was effective in reducing discoloration in frozen strawberries. A sample of 50 ripe strawberries was prepared for freezing. Then the sample was randomly divided into two groups of 25 strawberries each. Each strawberry was placed into a small plastic bag.

   The 25 bags in the control group were sealed. The preservative was added to the 25 bags containing strawberries in the treatment group, and then those bags were sealed. All bags were stored at 0°C for a period of 6 months. At the end of this time, after the strawberries were thawed, a technician rated each strawberry's discoloration from 1 to 10, with a low score indicating little discoloration.

   The dotplots below show the distributions of discoloration rating for the control and treatment groups.



   (a) The standard deviation of ratings for the control group is 2.141. Explain how this value summarizes variability in the control group.

   (b) Based on the dotplots, comment on the effectiveness of the preservative in lowering the amount of discoloration in strawberries. (No calculations are necessary.)

   (c) Researchers at the university decided to calculate a 95 percent confidence interval for the difference in mean discoloration rating between strawberries that were not treated with preservative and those that were treated with preservative. The confidence interval they obtained was (0.16, 2.72). Assume that the conditions necessary for the t-confidence interval are met.

   Based on the confidence interval, comment on whether there would be a difference in the population mean discoloration ratings for the treated and untreated strawberries.

1. (a) This is the average (typical) distance of each strawberries' discoloration rating from the mean.

E - interpreted correctly and in context

P - correct defn - no context
OR commented the control group appeared normal and applied the empirical rule

I - empirical rule without saying normal or wrote the formula for standard deviation

1. (b) The preservative seems to lower the amount of discoloration (is effective). The median is 6 for this group vs. 7 for the control group. Also, the ratings are more symmetric around 5 (mound-shaped), where the control group is more skewed left with most of it's ratings at 6 or higher.

The preservative does appear to have been effective in lowering the amount of discoloration in strawberries. The discoloration ratings for strawberries that received the preservative are clearly centered at a value that is lower than the center of the rating distribution for the control group.

By looking at the dotplots it can be seen that the preservative was somewhat effective because it yielded a greater number of low scores, indicating less discoloration overall (15 strawberries with the treatment had discoloration scores of 5 or less, whereas only 6 in the control group had discoloration scores of 5 or less.)

E - indicates that the preservative appears to be effective and explicitly links this decision to comparison of a characteristic of relative standing from the dotplots for the two groups.

P - says it appears to be effective because the ratings appear lower for the treatment group, but does not link... or correctly compares one or more characteristics but never states that it was effective at lowering discoloration.

I - stays it is not effective because the centers are roughly t he same, or says it's effective with incorrect or no justification.

1. (c) Yes, based on this confidence interval, there appears to be a difference in the population mean discoloration ratings. The C.I. does not contain 0, which would indicate no difference. you are 95% confident that the mean difference in discoloration ratings between the treatment and control groups is between .16 and 2.72, meaning the preservative was between .16 and 2.72 lower than the control group.

E - indicates 0 is not included in the C.I. so there is a difference in population mean and states the conclusion in the context of the experiment.

P - indicates that 0 is not included so there is a difference in pop. means but does not state in context, or correctly interprets 95% conf. in context and indicates there is diff in pop. means without indicating 0 is not included.

I - concludes it's not effective or says that no conclusion can be made or states a conclusion that refers to sample means instead of pop. means.

4 EEE

3 EEP

2 EEI
EPP
PPP

1 EPI
EII
PPI

(any order)

[redacted]

2. As dogs age, diminished joint and hip health may lead to joint pain and thus reduce a dog's activity level. Such a reduction in activity can lead to other health concerns such as weight gain and lethargy due to lack of exercise. A study is to be conducted to see which of two dietary supplements, glucosamine or chondroitin, is more effective in promoting joint and hip health and reducing the onset of canine osteoarthritis. Researchers will randomly select a total of 300 dogs from ten different large veterinary practices around the country. All of the dogs are more than 6 years old, and their owners have given consent to participate in the study. Changes in joint and hip health will be evaluated after 6 months of treatment.

(a) What would be an advantage to adding a control group in the design of this study?

(b) Assuming a control group is added to the other two groups in the study, explain how you would assign the 300 dogs to these three groups for a completely randomized design.

(c) Rather than using a completely randomized design, one group of researchers proposes blocking on clinics, and another group of researchers proposes blocking on breed of dog. How would you decide which <u>one</u> of these two variables to use as a blocking variable?

3. Big Town Fisheries recently stocked a new lake in a city park with 2,000 fish of various sizes. The distribution of the lengths of these fish is approximately normal.

(a) Big Town Fisheries claims that the mean length of the fish is 8 inches. If the claim is true, which of the following would be more likely?

- A random sample of 15 fish having a mean length that is greater than 10 inches

or

- A random sample of 50 fish having a mean length that is greater than 10 inches

Justify your answer.

(b) Suppose the standard deviation of the sampling distribution of the sample mean for random samples of size 50 is 0.3 inch. If the mean length of the fish is 8 inches, use the normal distribution to compute the probability that a random sample of 50 fish will have a mean length less than 7.5 inches.

(c) Suppose the distribution of fish lengths in this lake was nonnormal but had the same mean and standard deviation. Would it still be appropriate to use the normal distribution to compute the probability in part (b)? Justify your answer.

GO ON TO THE NEXT PAGE.

2. (a) A control gropu gives the researchers a comparison group to be used to evaluate the effectiveness of the treatments. It allows them to compare the changes in hip and joint health in the dogs receiving the two drugs vs. dogs receiving no treatment.

E - the advantage of using a comparison group is described in the context of this study

P - the advantage of using a control group is described, but not in context

I - says that control groups should always be used but gives no further explanation or an incorrect explanation.

2 (b) Roll a die for each dog:

1, 2 assign to gluc

3, 4 assign to chond

5, 6 control group

continue until all dogs have been assigned a group.

****DO NOT 'dump' the rest when one group is full!

Each dog will be assigned a number 001 - 300. Using a RNT, read 3 digit numbers ignoring repeats and numbers not assigned. The 1st 100 number selected will receive gluc, the next 100 chon., and the last 100 will be the control group.

E - randomization used correctly and the method can be implemented after reading the student response (so that two knowledgeable statistics users could use the same method.)

P - randomization is used, but the method could not be implemented after reading the student response.

I - randomization not used in a planned way OR the solution does not yield a completely randomized design.

2. (c) The key question is which variable has the strongest association with joint and hip health. The goal of blocking is to create groups of homogenous dogs. It is reasonable to assume that most clinics will see all kinds and breeds of dogs so there is no reason to suspect that joint and hip health will be strongly associated with clinic. On the other hand, different breeds of dogs tend to come in different sizes. This size of a dog is associated with joint and hip health so it would be better to block by breed.

Breed of dog is more likely going to affect joint and hip health than the clinic, so I would block by breed.

You could look at past studies/clients to clinics and see if there is a bigger difference between clinics or breed of dog in hip and joint health. Block the one that has the biggest difference.

E - argues var. with the stronger relationship to joint and hip health (response var.) should be used as the blocking var. OR states that the var. with the larger anticipated var. in the response measure should be used. A rationale is required, but a var. does not have to be selected.

P - indicated the purpose of blocking is to create homogeneous exp units but makes an error in application OR does not acknowledge that there may be more variability in one than the other OR does not recognize that both var are associated with variation in the response.

I - does not exhibit an understanding of the purpose of blocking.

4	EEE
3	EEP
2	EEI
	EPP
	PPP
1	EPI
	EII
	PPI

3. (a) The random sample of 15 fish is more likely to have a sample mean length greater than 10 inches. The sampling distribution of the sample mean is normal with mean 8 and standard deviation $\sigma/\sqrt{n}$. Thus, both sampling distributions will be centered at 8 inches, but the one when n=15 will have more variability than the sampling distribution of the sample means when n=50.

A random sample of 15 fish having a mean length that is greater than 10 inches is more likely. The standard deviation of the sampling distribution increases as the sample size decreases so greater deviation from the true population mean is more likely with a sample of 15 fish as opposed to 50.

E - says 15 is more likely AND the justification is based on the variability in the sampling distributions.

P - says 15 - makes correct statements about the sampling dist. of the sample mean or the probabilities but does not specifically refer to the variability in these two sampl. dist. OR remearks that the sample mean approaches the pop. mean as the sample size increases. (Law of Large Numbers)

I - no or incorrect justification or chooses 50.

3. (b)

samp. dist
of sample mean
fish length

$$z = \frac{7.5 - 8}{.3} = -1.67$$

$$Pr(z < -1.67) = .0475$$

Stand. dev. of samp. dist
of sample mean is .3,

means $\frac{s}{\sqrt{n}} = .3$

already done!

compute the prob.

NOT a test,
is $\bar{x} = 7.5$ sign. less
than 8....

E - Prob. correct and a reasonable sketch or calculation is shown.

P - incorrect but plausible calc. is shown. ex using $.3/\sqrt{50}$
OR switches 7.5 and 8 to get a z of 1.67.

I - answer provided with no justification or incorrect.

3. (c) Yes, because the sample size is 50, which is large (<30). According to the CLT, the sampling distribution will be normal even for nonnormal populations with a large sample size.

E - yes because of the CLT (sampl. dist.) and large n.

P - yes, but weak justification - ex. mentioning CLT without reference to sample size, and mentioning sample size without reference to CLT.

Part a - E = 2 pts, P = 1

b and c - E = 1pt P = 1/2

4 = 4 pts
3 = 3 or 3 1/2
2 = 2 or 2 1/2
1 = 1 or 1 1/2

[redacted]

4. Investigators at the U.S. Department of Agriculture wished to compare methods of determining the level of *E. coli* bacteria contamination in beef. Two different methods (A and B) of determining the level of contamination were used on each of ten randomly selected specimens of a certain type of beef. The data obtained, in millimicrobes/liter of ground beef, for each of the methods are shown in the table below.

		Specimen									
		1	2	3	4	5	6	7	8	9	10
Method	A	22.7	23.6	24.0	27.1	27.4	27.8	34.4	35.2	40.4	46.8
	B	23.0	23.1	23.7	26.5	26.6	27.1	33.2	35.0	40.5	47.8

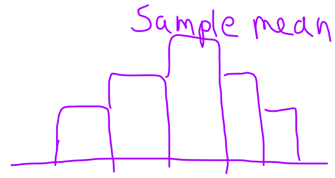
Is there a significant difference in the mean amount of *E. coli* bacteria detected by the two methods for this type of beef? Provide a statistical justification to support your answer.

1pt. matched pairs t-test for means!

• SRS from pop. of int → 10 rand. selected pieces of beef

• $n \geq 30$, no $n = 10$ OR pop. is normal (don't know!)

Sample mean differences



The hist. of sample diff. is symm w/o outliers. Hard to judge w/ only 10, but seems reasonable to assume pop. is normal. (t-test is robust)

1 pt.

μ_d = mean diff (A − B) in amt. of E. coli that would be detected by the 2 methods

$H_0: \mu_d = 0$

$H_a: \mu_d \neq 0$

1 pt.

$\bar{x}_d = .29$

$s_d = .63$

$n = 10$

* $df = 10 - 1 = 9$

$$t = \frac{.29 - 0}{.63/\sqrt{10}} = 1.46$$

$2 \cdot P(t > 1.46) = .18$ ← p-value

OR

$2(.10 < p < .05)$

$.20 < p < .10$

1 pt.

With a p-value of .18, this is not sign at the .10 level, so I fail to reject H_0.

Therefore, there is not enough evid. based on this sample to say there's a sign. diff. in the mean amt. of E. Coli detected by the 2 methods

each part is correct or incorrect
(no partial)
score = # of parts correct.

[redacted]

5. Researchers want to determine whether drivers are significantly more distracted while driving when using a cell phone than when talking to a passenger in the car. In a study involving 48 people, 24 people were randomly assigned to drive in a driving simulator while using a cell phone. The remaining 24 were assigned to drive in the driving simulator while talking to a passenger in the simulator. Part of the driving simulation for both groups involved asking drivers to exit the freeway at a particular exit. In the study, 7 of the 24 cell phone users missed the exit, while 2 of the 24 talking to a passenger missed the exit.

(a) Would this study be classified as an experiment or an observational study? Provide an explanation to support your answer.

(b) State the null and alternative hypotheses of interest to the researchers.

(c) One test of significance that you might consider using to answer the researchers' question is a two-sample z-test. State the conditions required for this test to be appropriate. Then comment on whether each condition is met.

(d) Using an advanced statistical method for small samples to test the hypotheses in part (b), the researchers report a p-value of 0.0683. Interpret, in everyday language, what this p-value measures in the context of this study and state what conclusion should be made based on this p-value.

5)

a) experiment → impose the treatment

randomly assigned the people to talk on cell or to passenger.

- random assign
- control →
- replication

b)

$H_0: p_{cell} = p_{pass}$

$H_a: p_{cell} > p_{pass}$

p_{cell} = prop. of all drivers who would be dist. w/ cell phone.

24
24

$24\left(\frac{9}{48}\right) \geq 5$

$4.5 \geq 5$

$24\left(1-\frac{9}{48}\right) \geq 5$

$\hat{p}_c = \frac{7+2}{24+24} = \frac{9}{48}$

c) $n_1 \hat{p}_1 \geq 5$ $\quad n_1(1-\hat{p}_1) \geq 5$

$n_2 \hat{p}_2 \geq 5$ $\quad n_2(1-\hat{p}_2) \geq 5$

$24\left(\frac{2}{24}\right) \geq 5$

$2 \geq 5$

fails n is too small

(obs. study)

• 2 independent SRS

OR

★ random assignment (exp) met = says r.a. cell/pass.

c) $24\left(\frac{7}{24}\right) \geq \cancel{5}$ 10

$7 \geq 5$

$24\left(\frac{2}{24}\right) \geq 5$

$\cancel{2 \geq 5}$

$24\left(1-\frac{7}{24}\right) \geq 5$

$17 \geq 5$

$24\left(1-\frac{2}{24}\right) \geq 5$

$22 \geq 5$

-fails, n is too small

$\hat{p} = \frac{7+2}{24+24} = \frac{9}{48} = .1875$

$n_1 = n_2 = 24$

$24\left(\frac{9}{48}\right) \geq 5$ $\quad 24\left(1-\frac{9}{48}\right) \geq 5$

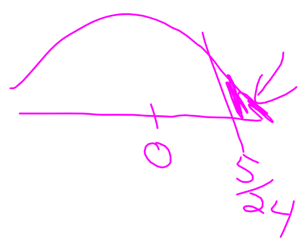
$4.5 \geq 5$ $\quad \geq 5$

2 ind SRS or r.a.

<u>int. p-value</u>

prob. of obtaining sample results as or more extreme, if H_0 is true.

$H_0: p_c = p_p$



6) a) slope = 1.080

slope → predicted ~~increase~~ decrease in y for each 1 unit inc. in x.

$\frac{1.080 \text{ perc. dist}}{1 \text{ actual dist.}}$

- perceived dist. is predicted to inc. 1.080 ft. for each addt'l foot in actual dist.

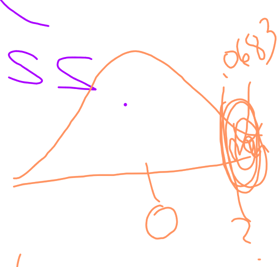
b) If no dist. between objects → it makes sense to have a pred. perceived dist. of 0.

d) non-contact $\hat{y} = 1.05x$

contact $\hat{y} = 1.17x$

e)

p-value of .0683, this is not sign. at the .05 level, So I fail to reject H_0. Not enough evid. to prove that cell phones are more dist. than pass



p-value of .0683 is the prob. of getting a sample result as extreme if the H_0 is true.

[redacted]

STATISTICS
SECTION II
Part B
Question 6

Spend about 25 minutes on this part of the exam.
Percent of Section II grade—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. A study was designed to explore subjects' ability to judge the distance between two objects placed in a dimly lit room. The researcher suspected that the subjects would generally overestimate the distance between the objects in the room and that this overestimation would increase the farther apart the objects were.

The two objects were placed at random locations in the room before a subject estimated the distance (in feet) between those two objects. After each subject estimated the distance, the locations of the objects were rerandomized before the next subject viewed the room.

After data were collected for 40 subjects, two linear models were fit in an attempt to describe the relationship between the subjects' perceived distances (y) and the actual distance, in feet, between the two objects.

$$\text{Model 1: } \hat{y} = 0.238 + 1.080 \times (\text{actual distance})$$

The standard errors of the estimated coefficients for Model 1 are 0.260 and 0.118, respectively.

$$\text{Model 2: } \hat{y} = 1.102 \times (\text{actual distance})$$

The standard error of the estimated coefficient for Model 2 is 0.393.

(a) Provide an interpretation in context for the estimated slope in Model 1.

(b) Explain why the researcher might prefer Model 2 to Model 1 in this context.

(c) Using Model 2, test the researcher's hypothesis that in dim light participants overestimate the distance, with the overestimate increasing as the actual distance increases. (Assume appropriate conditions for inference are met.)

The researchers also wanted to explore whether the performance on this task differed between subjects who wear contact lenses and subjects who do not wear contact lenses. A new variable was created to indicate whether or not a subject wears contact lenses. The data for this variable were coded numerically (1 = contact wearer, 0 = noncontact wearer), and this new variable, named "contact," was included in the following model.

$$\text{Model 3: } \hat{y} = 1.05 \times (\text{actual distance}) - 0.12 \times (\text{contact}) \times (\text{actual distance})$$

The standard errors of the estimated coefficients for Model 3 are 0.357 and 0.032, respectively.

[redacted]

[redacted]

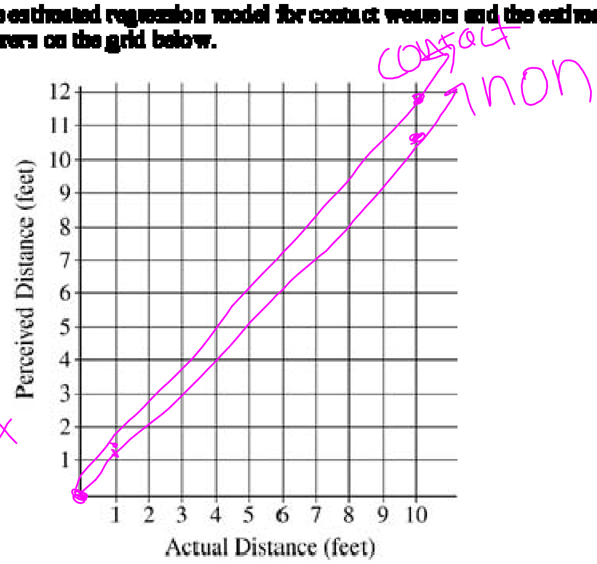
(d) Using Model 3, sketch the estimated regression model for contact wearers and the estimated regression model for noncontact wearers on the grid below.

non

$\hat{y} = 1.05x$

$\hat{y} = 1.05x + .12x$

$\hat{y} = 1.17x$



(e) In the context of this study, provide an interpretation of the estimated coefficients for Model 3.

STOP

END OF EXAM