Student name:

MATH 120: Elementary Statistics **Final exam**
Section 1 **May 6, 2019**

Instructions:

- This is a regular "closed-book" test, and is to be taken without the use of notes, books, or any reference materials other than those provided with this test.
- Collaboration or group work is not permitted.
- Cell-phone usage, in any form, is prohibited for the entire duration of the test. This also applies to any restroom breaks taken during the test.
- The time limit for taking this test is 2 hours from the scheduled start time.
- This test adds up to 50 points.

Part I

Give short answers to each question as instructed.

I.1. [4 pts.] Is there a relationship between the size of sand grains on a beach and its slope? Scatterplot of data from a sample of 20 beaches around the world suggests there may be a linear association. Based on these data, the following linear regression model was constructed to predict grain size (in mm.) from the beach slope (in degrees):

$$\widehat{\text{median sand diameter}} = 0.16 + 0.053 \text{ (beach slope)}$$

(a) Identify the explanatory variable and the response variable, including their units.
(b) Interpret the meaning of the slope (with units) in this application context.

I.2. [4 pts.] The following statistics summarize the age distribution (in years) of a 10-person team working on a project:

Median	Mean	IQR	SD
27.8	29	3.5	3.9

The oldest person, whose age is 34 years, leaves the team and is replaced by someone who is 36 years old. Compute, if possible, the new values of the above summary statistics. Show reasoning.

I.3. [4 pts.] Explain the difference between each of the following pairs of statistical terms:
(a) Sampling bias vs. sampling error.
(b) Response bias vs. non-response bias.

I.4. [4 pts.] A nationwide study on CEO compensation at publicly-traded companies surveyed 30 executives from Fortune 500 companies, and 54 from other smaller companies. They found the average annual compensation was \$9.2 million for the Fortune 500 sample, and \$7.7 million for the other group. We want to carry out a hypothesis test to determine whether this is evidence of a statistically significant difference.
(a) Write appropriate hypotheses, being sure to clarify what your parameters and subscripts denote.

(b) Suppose the P-value for this test turns out to be 0.15. Will a 90% confidence interval for estimating the difference in mean compensation contain 0? Give reasons.

I.5. [12 pts.] Each of the following questions requires a word, phrase, or numerical value as the answer. No reasoning or justification is needed.

i. A survey organization conducted telephone interviews in which 1,248 randomly selected adults in the United States were asked to respond to the question:

"At the present time do you think television commercials are an effective way to promote a new product?"

Identify the following as precisely as possible

* The population: _____

* The parameter(s): _____

ii. Given the probabilities $P(A) = 0.4$, $P(A \text{ and } B) = 0.4$, $P(A \text{ or } B) = 0.4$, find

* $P(A \mid B)$: _____

* Are A and B disjoint events? (Yes, or No): _____

iii. A survey of employee job satisfaction at a large corporation reported the correlations shown in the table. The variables are: YS=years of service; SL=salary; PR=promotion rate; and JS=job satisfaction.

	YS	**SL**	**PR**	**JS**
YS	1			
SL	0.23	1		
PR	0.58	0.74	1	
JS	−0.79	0.82	0.88	1

Assuming the conditions necessary for interpreting correlations are met, are the following true or false:

* Higher promotion rates are associated with longer years of service: _____

* Longer years of service are associated with greater job satisfaction: _____

iv. The monthly EPS (earnings per share) in dollars, over a 20-month period for a corporation are given below in ascending order:

Month	1	2	3	4	5	6	7	8	9	10	11	12
EPS	-2.4	-2.1	-1.7	-1.6	-0.2	-0.1	1.2	1.2	2.9	3.2	3.6	3.8

Month	13	14	15	16	17	18	19	20
EPS	4.0	4.1	4.1	4.2	4.4	4.7	4.8	5.0

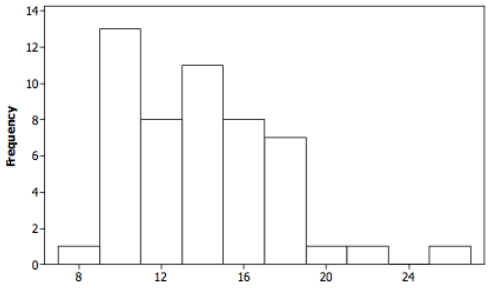
Find the 5-number summary:

v. For the histogram shown, what summary statistics would best describe the center and spread?

_____

Will the mean be larger, or smaller, than the median?

_____



vi. A meteorology organization carried out a hypothesis test to investigate whether there is a difference in true mean rainfall between two cities ($H_0 : \mu_1 = \mu_2$; $H_A : \mu_1 \neq \mu_2$). They found evidence of a significant difference, using $\alpha = 0.05$.

* What does this tell about their P-value? (i.e., how large, or small, might its value be)

_____

* Suppose the significance level for the test is changed to $\alpha = 0.1$. Could that change the conclusion/inference from the test?

_____

Part II

Give complete and detailed solutions to each question. Grading will primarily be based on correct steps, reasons, relevant sketches, and clarity.

II.1 [7 pts.] A team of scientists, researching the consequences of vitamin B_{12} deficiency, tracked a group of 57 adults with B_{12} deficiency for 7 years. At the end of this period they found 14 people in their sample exhibited symptoms of major depression.

(a) What kind of study was this? (survey? observation? experiment? prospective? retrospective? etc.)

(b) What is the nature and scope of conclusions this study can reach regarding B_{12} deficiency and depression? Explain.

(c) Use a confidence interval to estimate the true rate of depression among those with B_{12} deficiency, based on data from this sample. Be sure to include all steps and state an appropriate conclusion.

II.2 [7 pts.] The 2-way table shows the distribution of rank of university faculty at all U.S. medical schools by sex in 2016 (Source: Association of American Medical Colleges):

	Faculty Rank			
	Assistant professor	Associate professor	Professor	**Total**
Female	36,498	12,589	8,708	57,795
Male	43,561	21,991	28,634	94,186
Total	80,059	34,580	37,342	151,981

(a) Find the probability that a random person from this sample is not an assistant professor.

(b) What is the probability that of 3 randomly selected persons none is an assistant professor?

(c) Find the probability that someone who is a professor is female.

(d) Using probabilities, determine whether faculty rank and sex are independent.

[*Be sure to show all steps and justify their use.*]

II.3 [8 pts.] Does race matter when applying for National Institutes of Health grants? A study (reported in *Science,* August 2011) found that of 58,148 applications submitted by white researchers, 15,700 were funded by the NIH. Additionally, 198 of 1164 applications submitted by black researchers were funded. Is this evidence that the chance of funding is different for white and black researchers? Carry out a hypothesis test and state your conclusion. Show all steps, including: hypotheses; conditions check; sampling distribution model, with sketch; all calculations; and inference, with clear indication of what significance level you're using.