Sample Final Exam B – Part A

1. The five-number summary of the heights (in inches) of the players on a basketball team is calculated to be:

   72 76 78 79 81

   The distribution of heights for the team is:

   (A) skewed to the left and so the mean is greater than the median.
   (B) skewed to the right and so the mean is greater than the median.
   (C) skewed to the left and so the median is greater than the mean.
   (D) skewed to the right and so the median is greater than the mean.
   (E) approximately symmetric and so the mean and median are approximately equal.

2. The scores for a sample of golfers competing in a tournament are ordered and are shown below:

80 81 82 86 86 87 87 87 88 89 89 91 94 96

If we constructed an outlier boxplot for these data, the lines coming out of the box (the whiskers) would extend to the values:

(A) 82 and 91
(B) 86 and 89
(C) 81.5 and 93.5
(D) 83.5 and 90.5
(E) 80 and 96

3. The weights (in pounds) of a sample of 14 pugs, 29 bulldogs and 16 border collies are measured. Consider the summary statistics shown below:

Breed	n	min	Q1	med	Q3	max	mean	std. dev.
Pug	14	12.1	13.7	15.1	17.2	20.9	16.0	2.1
Bulldog	29	45.8	55.7	61.3	64.0	67.8	59.2	5.3
Border Collie	16	27.0	31.3	34.6	36.9	41.7	34.7	3.0

What is the mean weight of all 59 dogs in the sample combined?

(A) 40.1 (B) 42.3 (C) 38.7 (D) 34.9 (E) 36.6

4. A housecleaner would like to know the most effective method for removing streaks on windows. She would like to compare the effect of spraying the windows with Windex or Glass Plus cleaner and wiping them with either paper towel or a cotton cloth. She will try each combination three times, wiping the windows (all equally dirty) for 30 seconds each time.

What is/are the factors(s) in this experiment?

(A) type of window

(B) streaks removed

(C) type of cleaner and type of material used to wipe the windows

(D) Windex, Glass Plus, paper towel, cotton cloth

(E) Windex/paper towel, Windex/cotton cloth, Glass Plus/paper towel, Glass Plus/cotton cloth

5. In 2006, Major League Baseball set up the Joint Drug Prevention and Treatment Program in an effort to eliminate the widespread use of performance enhancing drugs such as anabolic steroids from the game. Random tests are done on players to make sure that they are not on such drugs. Testers go out to each of the 30 Major League Baseball teams and take a simple random sample of 5 of the players on the team. The testers then collect a urine and blood sample from the players that have been selected. If a player tests positive for performance enhancing drugs, he is suspended for 50 games.

   The resulting sample of 150 players is a:

   (A) simple random sample.
   (B) stratified random sample.
   (C) systematic random sample.
   (D) multistage sample.
   (E) voluntary response sample.

6. A random variable X follows a normal distribution with standard deviation σ. Ten statisticians would like to estimate the mean μ of the population. They will take separate random samples of n individuals and they will each calculate a 90% confidence interval for μ. What is the probability that exactly eight of the statisticians' confidence intervals contain the value of μ?

(A) 0.1937

(B) 0.2614

(C) 0.3483

(D) depends on the value of σ

(E) depends on the sample size n

7. An archer is shooting at a circular target. It is known that each arrow fired by the archer will score 0, 1, 3, 5, or 10 points (independently of everything else) with the following probabilities:

Points	0	1	3	5	10
Probability	0.1	0.1	0.3	0.2	0.3

If the archer shoots two arrows at the target, what is the probability that his total score is at least 15 points?

(A) 0.06 (B) 0.09 (C) 0.12 (D) 0.15 (E) 0.21

The next **two** questions (**8** and **9**) refer to the following:

Suppose we have the following facts about customers buying alcohol at the Liquor Mart:

- 45% buy wine (W).
- 37% buy beer (B).
- 14% buy vodka (V).
- 20% buy wine and beer.
- 6% buy wine and vodka.
- 46% buy beer or vodka.
- 2% buy wine and beer and vodka.

8. What is the probability that a randomly selected customer buys vodka but **not** beer?

(A) 0.06 (B) 0.07 (C) 0.08 (D) 0.09 (E) 0.10

9. What is the probability that a randomly selected customer buys exactly two of the three types of alcohol?

(A) 0.23 (B) 0.25 (C) 0.27 (D) 0.29 (E) 0.31

The next **four** questions (**10** to **13**) refer to the following:

The **time** it takes students to write the final exam in a large course follows a left-skewed distribution with mean 88 minutes and standard deviation 23 minutes. **Scores** on the exam follow a normal distribution with mean 71 and standard deviation 9.

10. What is the 91^{st} percentile of the distribution of exam **scores**?

(A) 81 (B) 83 (C) 85 (D) 87 (E) 89

11. We take a simple random sample of four students. What is the probability that their mean **score** on the exam is greater than 73?

(A) 0.3300
(B) 0.5556
(C) 0.6700
(D) 0.4444
(E) impossible to determine with the information given

12. We take a random sample of four students. What is the probability that the mean **time** it takes them to write the exam is greater than 95 minutes?

(A) 0.1739

(B) 0.4325

(C) 0.8261

(D) 0.5675

(E) impossible to determine with the information given

13. We will take a random sample of 100 students and calculate the average **time** $\bar{X}$ it takes them to write the exam. The distribution of $\bar{X}$ is:

(A) approximately normal with mean 88 and standard deviation 23.

(B) skewed to the left with mean 88 and standard deviation 23.

(C) approximately normal with mean 88 and standard deviation 2.3.

(D) skewed to the left with mean 88 and standard deviation 2.3.

(E) approximately normal with mean 88 and standard deviation 0.23.

14. It is known that 47% of students at a large university are male. If we take a random sample of 200 students at the university, what is the approximate probability that less than half of them are male?

(A) 0.7291 (B) 0.8023 (C) 0.7852 (D) 0.8508 (E) 0.7517

15. We would like to estimate the true mean size (in acres) of all farms in a U.S. state. It is calculated that, in order to estimate the true mean size to within 2 acres with 95% confidence, a sample of 90 farms is required. What sample size would be required to estimate the true mean size to within 3 acres with 95% confidence?

(A) 40 (B) 60 (C) 74 (D) 135 (E) 203

16. We would like to estimate the true mean contents μ of a bottle of a certain brand of cough syrup. We measure the contents of a random sample of 35 bottles and we calculate a mean of 252 ml. It is known that the contents of a bottle of cough syrup follow a normal distribution with standard deviation 1.5 ml. An 82% confidence interval for μ is:

(A) $2.52 \pm 1.48 \left(\frac{1.5}{\sqrt{35}}\right)$

(B) $2.52 \pm 1.34 \left(\frac{1.5}{\sqrt{35}}\right)$

(C) $2.52 \pm 0.82 \left(\frac{1.5}{\sqrt{35}}\right)$

(D) $2.52 \pm 0.92 \left(\frac{1.5}{\sqrt{35}}\right)$

(E) $2.52 \pm 1.28 \left(\frac{1.5}{\sqrt{35}}\right)$

17. A random variable X follows a normal distribution with known standard deviation σ. We would like to construct a confidence interval for the true mean μ of the distribution of X. For which of the following combinations of sample size and confidence level would the confidence interval be the **narrowest**?

(A) 96% confidence level with $n = 25$

(B) 96% confidence level with $n = 100$

(C) 98% confidence level with $n = 25$

(D) 98% confidence level with $n = 100$

(E) depends on the value of σ.

18. A random variable X follows a normal distribution with standard deviation 5. We take a random sample of 100 individuals from the population and calculate a confidence interval for μ to be (40.973, 43.027). What is the confidence level for this interval?

(A) 90% (B) 95% (C) 96% (D) 98% (E) 99%

19. We would like to estimate the true mean size μ (in square feet) of all two-bedroom apartments in Winnipeg. A random sample of 14 two-bedroom apartments in the city is selected, and the mean and standard deviation of the sizes of these apartments are calculated to be 1000 square feet and 200 square feet, respectively. Assuming the sizes of two-bedroom apartments in the city follow a normal distribution, a 99% confidence interval for μ is:

(A) (862, 1138)
(B) (845, 1155)
(C) (858, 1142)
(D) (839, 1161)
(E) (876, 1124)

20. Packages of frozen peas are supposed to have a mean weight of 10 oz. The manufacturer wishes to detect if the mean is either too low (which is illegal) or too high (which reduces profit). Experience shows that the weights have a normal distribution with standard deviation 0.26 oz. The mean weight of a random sample of 18 bags is found to be 9.84 oz. We conduct a hypothesis test at the 5% level of significance to test the manufacturer's concern. The P-value for the appropriate hypothesis test is:

(A) 0.0045 (B) 0.0090 (C) 0.0125 (D) 0.0250 (E) 0.0375

21. We would like to conduct a hypothesis test to determine whether the true mean pH level in a lake is less than 7.0. Lake pH levels are known to follow a normal distribution. We take 5 water samples from random locations in the lake. For these samples, the mean pH level is 6.73 and the standard deviation is 0.4. What is the P-value of the appropriate test of significance?

(A) between 0.01 and 0.02

(B) between 0.025 and 0.05

(C) between 0.05 and 0.10

(D) between 0.10 and 0.15

(E) between 0.15 and 0.20

22. A man accused of committing a crime is taking a polygraph (lie detector) test. The polygraph is essentially testing the hypotheses

$$H_0\text{: The man is telling the truth. vs. } H_a\text{: The man is lying.}$$

Suppose we use a 5% level of significance. Based on the man's responses to the questions asked, the polygraph determines a P-value of 0.08. We conclude that:

(A) There is insufficient evidence that the man is telling the truth.

(B) There is sufficient evidence that the man is telling the truth.

(C) There is insufficient evidence that the man is lying.

(D) The probability that the man is lying is 0.33.

(E) The probability that the man is telling the truth is 0.33.

23. We would like to construct a 90% confidence interval for the true mean blood calcium of all healthy pregnant young women. A clinic measures the blood calcium of a sample of 50 healthy pregnant young women. The mean and standard deviation of these 50 measurements are, respectively, $\bar{x} = 9.8$ and $s = 0.5$. Calcium levels for healthy pregnant young women are known to follow a normal distribution. What is the standard error of the sample mean $\bar{X}$?

(A) 0.5000 (B) 0.1423 (C) 0.0707 (D) 0.1185 (E) 0.0100

24. We would like to test whether the true mean amount of money spent by gamblers at a casino differs from \$100; that is, we want to test $H_0 : \mu = 100$ vs. $H_a : \mu \neq 100$. We select a random sample of 50 gamblers and calculate a 98% confidence interval to be (\$99, \$114). Which of the following is true?

(A) We would reject H_0 at the $\alpha = 0.02$ level of significance.

(B) We would not reject H_0 at the $\alpha = 0.04$ level of significance.

(C) We would not reject H_0 at the $\alpha = 0.02$ level of significance.

(D) We would reject H_0 at the $\alpha = 0.01$ level of significance.

(E) A confidence interval cannot be used to conduct this test.

25. A manufacturer of electronic components will send a large shipment to a retailer only if there is significant evidence that less than 5% of the components in the shipment are defective. The manufacturer tests a random sample of 300 components and finds that 9 of them are defective. The test statistic for the appropriate test of $H_0 : p = 0.05$ vs. $H_a : p < 0.05$ is:

(A) $z = \dfrac{0.03 - 0.05}{\sqrt{\dfrac{(0.03)(0.97)}{300}}}$

(B) $z = \dfrac{0.05 - 0.03}{\sqrt{\dfrac{(0.05)(0.95)}{300}}}$

(C) $z = \dfrac{0.03 - 0.05}{\sqrt{\dfrac{(0.03)(0.05)}{300}}}$

(D) $z = \dfrac{0.03 - 0.05}{\sqrt{\dfrac{(0.05)(0.95)}{300}}}$

(E) $z = \dfrac{0.05 - 0.03}{\sqrt{\dfrac{(0.03)(0.97)}{300}}}$

26. We would like to estimate the proportion p of Manitoban's who are bilingual (fluent in two languages). We take a random sample of 200 Manitobans and find that 42 of them are bilingual. What is the margin of error for a 95% confidence interval for p?

(A) 0.0565 (B) 0.0353 (C) 0.0693 (D) 0.0288 (E) 0.0404

27. It is commonly believed that about 10% of the population is left-handed. One researcher believes that the actual proportion is lower than this. He takes a simple random sample of 200 individuals and finds that 16 of them are left-handed. What is the P-value for the appropriate test of significance to test the researcher's suspicion?

(A) 0.0572 (B) 0.0885 (C) 0.1271 (D) 0.1492 (E) 0.1736

Sample Final Exam B – Part B

1. The professor of a mechanical engineering course would like to determine whether a students midterm score can be used to predict his or her final exam score. The midterm score (out of 50) and the final exam score (out of 100) are recorded for a sample of students. The data are plotted on a scatterplot linear relationship is apparent. The professor calculates the means and standard deviations to be $\bar{x} = 31.0$, $\bar{y} = 61.4$, $s_x = 6.5$ and $s_y = 14.6$. The equation of the least squares regression line is calculated to be $\hat{y} = 2.5 + 1.9x$.

   (a) Interpret the meaning of the slope of the least squares regression line.

   (b) One student scored 36 on the midterm and 73 on the final exam. What is the residual for this student?

   (c) What percentage of the variation in final exam score can be accounted for by its regression on midterm score?

2. (a) Explain the difference between two events being mutually exclusive and two events being independent.

   (b) Event A has probability 0.4 to occur and Event B has probability 0.7 to occur. Are A and B mutually exclusive (disjoint)? Explain.

   (c) Event A has probability 0.4 to occur and Event B has probability 0.7 to occur. If A and B are independent, what is the probability P(A or B^c)? (Hint: If A and B are independent, it can be shown that A and B^c are independent as well.)

3. One measure of the quality of education provided by a university is the number of students per class. The president of a large university would like to estimate the true mean size of all third-year classes at the university. A random sample of 30 third-year classes is selected. The number of students in each of the classes are ordered and shown below:

10	11	14	14	15	17	18	20	22	22
22	23	26	26	27	28	31	33	34	34
36	42	44	49	50	55	62	68	77	82

From these data, the sample mean is calculated to be 33.73. The population standard deviation of class sizes is known to be 18.50.

(a) Before even looking at the data, we know that the distribution of class sizes could not possibly be normal. Explain why.

(b) Construct a stemplot for these data. What is the shape of the data distribution?

(c) It is clear that class size does not follow a normal distribution. Explain why it is still appropriate to use inference methods which rely on the assumption of normality.

(d) Conduct an appropriate hypothesis test, at the 1% level of significance, to determine whether there is evidence that the true mean third-year class size at this university differs from the national average of 40. Show all of the steps, including the statement of hypotheses, the calculation of the appropriate test statistic and P-value, and a carefully-worded conclusion.

4. Sixteen people volunteered to be part of an experiment. All 16 people were Caucasian, between the ages of 25 and 35, and were supplied with nice clothes. Eight of the people were male and eight were female. The question of interest in this experiment was whether females receive faster service at restaurants than males. Each of the eight male participants was randomly assigned a restaurant, and each of the eight females was randomly assigned to one of these same eight restaurants. One Friday night, all 16 people went out to eat, each one alone. The male and female assigned to the same restaurant would arrive within five minutes of each other, with the order determined by flipping a coin (male first or female first). Each person then ordered a similar drink and a similar meal. The time (in minutes) until the food arrived at the table was recorded. Some information that may be helpful is shown below:

Females	Males	Difference (d = F−M)
mean = 17.5	mean = 20.2	mean =−2.7
std. dev. = 3.7	std. dev. = 4.8	std. dev. = 2.9

(a) Conduct an appropriate hypothesis test, at the 1% level of significance. Show all of the steps, including the statement of hypotheses, the calculation of the appropriate test statistic and P-value, and a carefully-worded conclusion. Assume the appropriate normality conditions are satisfied.

(b) Interpret the P-value of the test to someone with little or no background in statistics.