Sample Final Exam A – Part A

1. The mean salary of the nine employees of a small business is \$52,000 per year. One employee, whose salary was \$68,000, was fired. Two new employees, who will each be paid a salary of \$40,000, are hired. What is the new mean annual salary of the employees of the business?

   (A) \$44,000 (B) \$45,000 (C) \$46,000 (D) \$47,000 (E) \$48,000

2. A consumer group is testing a certain brand of light bulb. The lifetimes (in hours) for a sample of 12 light bulbs (the time until the bulbs burn out) are shown below:

$$179 \quad 294 \quad 400^{+} \quad 289 \quad 93 \quad 101 \quad 372 \quad 400^{+} \quad 153 \quad 400^{+} \quad 245 \quad 340$$

The study lasted 400 hours. The time for a bulb that was still working at the end of the 400 hours is recorded as "400^{+}". The median lifetime for bulbs in this sample is:

(A) 245 hours.

(B) 291.5 hours.

(C) 236.5 hours.

(D) 317 hours.

(E) impossible to calculate because we don't know the exact lifetime of three of the bulbs.

The next **two** questions (**3** to **4**) refer to the following:

We would like to determine whether a man's shoe size X can be used to predict his height Y. The shoe sizes and heights of a random sample of eight men are shown below:

Shoe Size	11	10	9.5	12	11	11.5	10.5	10
Height (inches)	69	70	67	74	72	70	71	68

The correlation between Shoe Size and Height is calculated to be $r = 0.78$, and the equation of the least-squares regression line is calculated to be $\hat{y} = 50 + 2x$.

3. What is the correct interpretation of the slope of the least squares regression line?

(A) When a man's shoe size increases by one, we predict his height to increase by two inches.

(B) When a man's height increases by two inches, we predict his shoe size to increase by one.

(C) When a man's shoe size increases by two, we predict his height to increase by one inch.

(D) When a man's height increases by one inch, we predict his shoe size to increase by two.

(E) When a man's shoe size increases by one, his height will always increase by exactly two inches.

4. Which of the following statements is **false**?

   (A) The predicted height of a man with a size 11 shoe is 72 inches.

   (B) It would not be appropriate to use this regression line to predict the height of a man with a size 8 shoe.

   (C) About 78% of the variation in height is accounted for by its regression on shoe size.

   (D) The high correlation between shoe size and height does not indicate a causal relationship.

   (E) It would not be appropriate to use this regression line to predict the height of a woman from her shoe size.

5. A bowler plays one game every day for one month. His scores are ordered and are shown below:

139	142	144	149	156	166	171	178	179	179
181	182	183	185	185	189	190	190	190	191
193	195	200	202	207	212	220	227	228	235

The five-number summary for this data set is as follows:

139	178	187	200	235

If we construct a modified (outlier) boxplot for the bowlers scores, how many scores would be labeled as potential outliers?

(A) 3 (B) 4 (C) 5 (D) 6 (E) 7

6. The faculty association at a university would like to conduct a survey of its full-time professors. They anticipate that professors in different academic areas may respond differently, so they decide to take a stratified random sample. Which of the following sampling procedures will produce a stratified random sample?

   (A) Select a simple random sample of five academic departments and survey all professors in those departments.

   (B) Select a simple random sample of three faculties. Within each of the selected faculties, select a simple random sample of departments, and survey all professors in those departments.

   (C) Select a simple random sample of five academic departments. Within each of the selected departments, select a simple random sample of four professors and give them the survey.

   (D) Select a simple random sample of four professors in each academic area and give them the survey.

   (E) Print the survey in the faculty association's monthly newsletter and ask professors to log onto a website to respond.

7. Randomization is used in a randomized block design to:

(A) select the individuals to participate in the experiment.
(B) select which treatments will be compared.
(C) place the individuals into blocks.
(D) assign the treatments to the individuals within the blocks.
(E) all of the above.

8. A coffee store owner knows that a customer gets a coffee in 35% of the store's transactions. If we take a random sample of 300 customers, what is the probability that less than 32% of them buy a coffee?

(A) 0.1292 (B) 0.1379 (C) 0.1469 (D) 0.1587 (E) 0.1685

9. Suppose it is known that diastolic blood pressures (measured in mm of mercury) of patients visiting a clinic follow a normal distribution with mean 67 and standard deviation 6. What is the probability that the mean diastolic blood pressure of a sample of 20 patients is between 66.8 and 67.9?

(A) 0.4099 (B) 0.3082 (C) 0.3743 (D) 0.1890 (E) 0.5217

The next **two** questions (**10** and **11**) refer to the following:

Weights of pears sold at a supermarket follow a normal distribution with mean 195 grams and standard deviation 40 grams.

10. Which of the following is closest to the third quartile of weights of all pears sold at the supermarket?

(A) 225 grams
(B) 215 grams
(C) 218 grams
(D) 236 grams
(E) 222 grams

11. A random sample of five pears is selected. What is the probability that the total weight of the pears is greater than one kilogram (i.e. 1,000 grams)?

(A) 0.2795 (B) 0.3897 (C) 0.2451 (D) 0.3264 (E) 0.4129

12. A backpacking party carries five emergency flares, each of which will light with a probability of 0.93. What is the probability that exactly four of the flares will light?

(A) 0.1271 (B) 0.0524 (C) 0.2947 (D) 0.2618 (E) 0.1835

13. Lumber intended for building houses and other structures must be monitored for strength. A random sample of 25 specimens of Southern Pine is selected, and the mean strength is calculated to be 3700 pounds per square inch. Strengths are known to follow a normal distribution with standard deviation 500 pounds per square inch. An 85% confidence interval for the true mean strength of Southern Pine is:

(A) (3615, 3785)

(B) (3671, 3729)

(C) (3556, 3844)

(D) (3544, 3856)

(E) (3596, 3804)

14. A 99% confidence interval for the mean time μ spent in the shower (in minutes) by adult males is calculated to be 6 ± 3. Which of the following provides a correct interpretation of this interval?

(A) There is a 99% probability that μ lies in this interval.

(B) In 99% of all samples of the same size, $\bar{x}$ will be between 3 and 9.

(C) We are 99% confident that a randomly selected adult male will spend between 3 and 9 minutes in the shower.

(D) In repeated sampling, 99% of similarly constructed intervals will contain $\bar{x}$.

(E) In repeated sampling, 99% of similarly constructed intervals will contain μ.

15. 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.

The next **two** questions (**16** to **17**) refer to the following:

A university administrator would like to estimate the true mean salary of all male professors at the university. She calculates that, in order to estimate the true mean to within \$500 with 90% confidence, she will require a random sample of 80 male professors.

16. What sample size would be required in order to estimate the true mean salary of all male professors at the university to within \$250 with 90% confidence?

(A) 20 (B) 40 (C) 80 (D) 160 (E) 320

17. Suppose the administrator would also like to estimate the true mean salary of all female professors at the university. It is known that the university has twice as many male professors as female professors, and that the standard deviation of salaries is equal for males and females. What sample size would be required to estimate the true mean salary of all female professors at the university to within \$500 with 90% confidence?

(A) 20 (B) 40 (C) 80 (D) 160 (E) 320

18. A statistical test of significance is designed to:

(A) prove that the null hypothesis is true.
(B) find the probability that the alternative hypothesis is true.
(C) find the probability that the null hypothesis is true.
(D) assess the strength of the evidence in favour of the null hypothesis.
(E) assess the strength of the evidence in favour of the alternative hypothesis.

19. Prior to distributing a large shipment of bottled water, a beverage company would like to determine whether there is evidence that the true mean fill volume of all bottles differs from 600 ml, which is the amount stated on the labels. Fill volumes are known to follow a normal distribution with standard deviation 2.0 ml. A random sample of 25 bottles is selected. The sample has a mean of 598.8 ml and a standard deviation of 3.0 ml. What is the value of the test statistic for testing the appropriate test of significance?

(A) $t = -0.50$ (B) $z = -2.00$ (C) $t = -2.00$ (D) $z = -3.00$ (E) $t = -3.00$

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 0.2 oz. The mean weight of a random sample of 25 bags is found to be 9.7 oz. The appropriate null and alternative hypotheses for testing the manufacturer's concern are:

(A) $H_0 : \mu = 10$ vs. $H_a : \mu < 10$

(B) $H_0 : \mu = 9.7$ vs. $H_a : \mu \neq 9.7$

(C) $H_0 : \bar{X} = 10$ vs. $H_a : \bar{X} \neq 10$

(D) $H_0 : \mu = 9.7$ vs. $H_a : \mu < 9.7$

(E) $H_0 : \mu = 10$ vs. $H_a : \mu \neq 10$

21. We would like to test whether the true mean IQ of all adult Canadians is less than 110. Suppose that IQs of adult Canadians follow an approximate normal distribution with standard deviation 17. A sample of 30 adult Canadians has a sample mean IQ of 108. What is the P-value for the appropriate test of $H_0 : \mu = 110$ vs. $H_a : \mu < 110$?

(A) 0.6444 (B) 0.2090 (C) 0.3556 (D) 0.2611 (E) 0.3156

22. We would like to determine whether the true mean systolic blood pressure μ of healthy adults differs from 120. We obtain a sample of healthy adults and conduct an appropriate hypothesis test, which results in a P-value of 0.021. Which of the following statements is true?

I. A 96% confidence interval for μ would contain the value 120.
II. A 98% confidence interval for μ would contain the value 120.
III A 99% confidence interval for μ would not contain the value 120.

(A) I only
(B) II only
(C) III only
(D) I and II only
(E) II and III only

23. A random variable X is known to follow a normal distribution with unknown mean μ and unknown standard deviation σ. We would like to test whether the true mean differs from 75. We take a random sample of 10 individuals and we calculate a test statistic of $t = 2.79$. The P-value of the test is:

(A) between 0.005 and 0.01.

(B) between 0.01 and 0.02.

(C) between 0.02 and 0.04.

(D) between 0.04 and 0.05.

(E) between 0.05 and 0.10.

24. The stemplot below displays the heights (in cm) of all 26 players on the Winnipeg Jets hockey team:

17	8
18	0 1 1 2 3
18	5 7 7 8 8 8 9 9
19	0 0 1 1 1 2 3 3 3
19	6 6 7

We would like to conduct a hypothesis test to determine whether the true mean height of all Winnipeg Jets players differs from 190 cm. Which of the following statements is true?

(A) We should use Z as the test statistic because the distribution of heights appears to be approximately normal.

(B) We should use T as the test statistic because the population standard deviation is not known.

(C) We should use Z as the test statistic because the Central Limit Theorem tells us that $\bar{X}$ will have an approximate normal distribution when the sample size is large.

(D) We should use T as the test statistic because there is no way to know whether the distribution of heights is normal.

(E) A hypothesis test is unnecessary in this situation.

25. A statistician conducted a test of $H_0 : \mu = 1$ vs.$H_a : \mu > 1$ for the mean μ of some population. Based on the gathered data, the statistician concluded that H_0 could be rejected at the 1% level of significance. Using the same data, which of the following statements must be **true**?

    I. A test of $H_0 : \mu = 1$ vs. $H_a : \mu > 1$ at the 10% level of significance would also lead to rejecting H_0
    II. A test of $H_0 : \mu = 0$ vs. $H_a : \mu > 0$ at the 1% level of significance would also lead to rejecting H_0
    III. A test of $H_0 : \mu = 1$ vs. $H_a : \mu \neq 1$ at the 1% level of significance would also lead to rejecting H_0

    (A) I only
    (B) I and II only
    (C) I and III only
    (D) II and III only
    (E) I, II and III

26. We would like to construct a 95% confidence interval to estimate the true proportion of all voters who plan to support the New Democratic Party in an upcoming provincial election. What sample size is required in order to estimate this proportion to within 0.04 with 95% confidence?

(A) 601 (B) 801 (C) 1001 (D) 1201 (E) 1401

27. Suppose it is known that 83% of motorists wear a seatbelt while driving. The police stop a random sample of 200 drivers. What is the probability that more than 80% of them are wearing a seat belt?

(A) 0.8023 (B) 0.8212 (C) 0.8554 (D) 0.8708 (E) 0.8997

Sample Final Exam A – Part B

1. Hat # 1 contains four gold coins, one silver coin and five copper coins. Hat # 2 contains three gold coins, six silver coins and one copper coin.

   (a) You will randomly select one coin from each of the two hats. The outcome of interest is the colour of each of the selected coins. Give the complete sample space of possible outcomes, and calculate the probability of each outcome.

   (b) What is the probability that the two selected coins are the same colour?

   (c) Let X be the number of gold coins that are selected. Find the probability distribution of X.

2. The average body temperature of healthy adults is commonly thought to be 37.0°C. A physician believes that the true mean is actually less than this. She measures the body temperatures of a random sample of 27 healthy adult patients. From these data, the sample mean is calculated to be 36.85°C and the sample standard deviation is calculated to be 0.55°C. It is known that body temperatures of healthy adults follow a normal distribution.

   (a) Conduct an appropriate hypothesis test, at the 1% level of significance, to test the physician's suspicion. 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.

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

3. A drug company manufactures antacid that is known to be successful in providing relief for 70% of people with heartburn. The company tests a new formula on a simple random sample of 150 people with heartburn. 114 of the subjects who try the new antacid report feeling some relief.

   Conduct an appropriate hypothesis test, at the 10% level of significance, to determine if there is a difference in effectiveness between the old formula and the new formula.

4. The owner of an ice cream truck would like to know if the temperature affects the amount of ice cream he sells. The following table displays the temperature (in °C) and ice cream sales (in $) for a sample of days from the past summer.

Day	1	2	3	4	5	6	Mean	Std. Dev.
Temp.	23	31	18	24	28	20	24.00	4.86
Sales	290	485	200	395	340	150	310.00	123.98

The correlation between Temperature and Sales is calculated to be $r = 0.88$ and the equation of the least squares regression line is calculated to be $\hat{y} = -232.03 + 22.58x$.

(a) Use the grid below to make a scatterplot of the data. Comment on the relationship between Temperature and Sales.

(b) Interpret the slope of the least squares regression line **in this example**.

(c) Calculate the residual for Day 5.

(d) What fraction of the variation in Sales can be accounted for by its regression on Temperature?