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.

3. 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

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

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

The next **two** questions (**4** and **5**) refer to the following:

Can the sodium content in a hamburger be used to predict the number of calories? Sodium Content (in mg) and Calories for hamburgers from several fast food chains were measured. The data are plotted on a scatterplot and it is apparent that a linear relationship is a realistic assumption. The equation of the least squares regression line is calculated to be $\hat{y} = 320 + 1.4x$.

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

(A) When Sodium Content increases by 1.4 mg, we predict Calories to increase by 1.

(B) When Calories increase by 1.4, we predict Sodium Content to increase by 1 mg.

(C) When Sodium Content increases by 1 mg, we predict Calories to increase by 320.

(D) When Calories increase by 1, we predict Sodium Content to increase by 1.4 mg.

(E) When Sodium Content increases by 1 mg, we predict Calories to increase by 1.4.

5. It is reported that 67% of the variation in Calories can be accounted for by its regression on Sodium Content. What is the value of the correlation between Sodium Content and Calories?

(A) 0.933 (B) 0.819 (C) 0.449 (D) 0.670 (E) 0.622

6. Which of the following statements contains an obvious error?

(A) "The standard deviation of the ages of five friends was calculated to be $s = 0$."

(B) "A variable X follows a normal distribution with mean $\mu = -5$ and standard deviation $\sigma = 7$."

(C) "The females in a sample were shorter than the males, and the correlation between height and gender was calculated to be $r = -0.56$."

(D) "Even though a simple random sample was selected, two thirds of the individuals in the sample were female."

(E) "The experiment was double blind, and so even the doctor had no idea whether the patient was receiving the actual medication or a placebo."

7. Two events A and B are independent. If $P(A) = 0.32$ and $P(A \cup B) = 0.64$, then what is $P(B)$?

(A) 0.32 (B) 0.38 (C) 0.43 (D) 0.47 (E) 0.51

8. There are three games scheduled in the National Hockey League one night. The games, and the probabilities of each team winning their respective game are shown below.

	Visitor		Home
Game 1:	Dallas (0.3)	vs.	Los Angeles (0.7)
Game 2:	Boston (0.6)	vs.	Carolina (0.4)
Game 3:	Edmonton (0.2)	vs.	Detroit (0.8)

What is the probability that exactly two of the home teams win their games?

(A) 0.348 (B) 0.398 (C) 0.428 (D) 0.488 (E) 0.538

The next **three** questions (**9**, **10**, and **11**) refer to the following:

A machine automatically fills bottles with dish soap. The amount of soap per bottle follows a normal distribution with mean 828 ml and standard deviation 4 ml.

9. What proportion of bottles contain between 823 and 834 ml of soap?

(A) 0.8276 (B) 0.8643 (C) 0.8095 (D) 0.8837 (E) 0.8408

10. What is the probability that a random sample of 10 bottles of dish soap contain a total amount greater than 8.3 litres?

(A) 0.3085 (B) 0.1539 (C) 0.9429 (D) 0.2296 (E) 0.0571

11. What amount should be placed on the label of the bottles so that only 4% of bottles contain less than that amount?

(A) 820 ml (B) 821 ml (C) 822 ml (D) 828 ml (E) 829 ml

12. 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

13. A soft drink company is having a promotion. When customers buy a bottle of pop, they can look under the cap to see if they have won a prize (a free bottle of pop, a TV, a car, etc.) According to the company, 16% of the bottles are winners. If you buy nine bottles of pop, what is the probability that less than two of them are winners?

(A) 0.6047 (B) 0.4891 (C) 0.2720 (D) 0.5652 (E) 0.3569

14. Suppose it is known that 20% of all Canadian adults are smokers. If you take a random sample of 200 Canadian adults, what is the probability that less than 17% of them are smokers?

(A) 0.2451 (B) 0.1446 (C) 0.0901 (D) 0.0516 (E) 0.1292

15. The number of undergraduate students at the University of Winnipeg is approximately 7,000, while the University of Manitoba has approximately 21,000 undergraduate students. Suppose that, at each university, a simple random sample of 3% of the undergraduate students is selected and the following question is asked: "Do you approve of the provincial government's decision to lift the tuition freeze?" Suppose that, within each university, approximately 20% of undergraduate students favour this decision. What can be said about the sampling variability associated with the two sample proportions?

(A) The sample proportion from U of W has less sampling variability than that from U of M.

(B) The sample proportion from U of W has more sampling variability that that from U of M.

(C) The sample proportion from U of W has approximately the same sampling variability as that from U of M.

(D) It is impossible to make any statements about the sampling variability of the two sample proportions without taking many samples.

(E) It is impossible to make any statements about the sampling variability of the two sample proportions because the population sizes are different.

16. 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)

17. We would like to estimate the true mean time (in minutes) it takes to play a Major League Baseball game. We measure the times for a simple random sample of 30 games and we calculate a 95% confidence interval for μ to be (160, 180), i.e., the length of the interval is 20. Suppose we had instead selected a simple random sample of 60 games and calculated a 95% confidence interval for μ. What would be the length of this interval?

(A) 5.00 (B) 7.07 (C) 10.00 (D) 14.14 (E) 28.28

The next **two** questions (**18** and **19**) refer to the following:

We would like to estimate the true mean amount of time Canadian teens spend on the internet per day. We calculate that, in order to estimate μ to within $\pm$ 10 minutes with 90% confidence, we require a sample of 100 Canadian teens.

18. What sample size would be required to estimate the true mean amount of time Canadian teens spend on the internet per day to within $\pm$ 2 minutes with 90% confidence?

(A) 500 (B) 4 (C) 224 (D) 20 (E) 2,500

19. The United States has a population ten times as large as Canada. Assuming equal standard deviations, what sample size would be required to estimate the true mean amount of time American teens spend on the internet per day to within $\pm$ 10 minutes with 90% confidence?

(A) 1,000 (B) 10 (C) 100 (D) 10,000 (E) 317

20. 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.

21. 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 printed 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 the appropriate test of significance?

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

22. 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

23. 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

24. Do snow tires help vehicles stop more quickly in winter driving conditions? A sample of 10 vehicles is outfitted with snow tires. The vehicles travel 90 km/h in winter driving conditions and apply the brakes. The sample mean stopping distance for these vehicles is 185 meters. We would like to test whether the true mean stopping distance is less than 190 meters. The test statistic is calculated to be $t = -2.80$. At the 1% level of significance, we should:

(A) fail to reject H_0, since the P-value is between 0.005 and 0.01.

(B) reject H_0, since the P-value is between 0.01 and 0.02.

(C) fail to reject H_0, since the P-value is between 0.02 and 0.04.

(D) reject H_0, since the P-value is between 0.005 and 0.01.

(E) fail to reject H_0, since the P-value is between 0.01 and 0.02.

The next **two** questions (**25** and **26**) refer to the following:

A researcher would like to determine if right-handed people react faster with their right hand than their left hand. The researcher locates 30 right-handed people who agree to participate in an experiment. Each person is asked to push a button with their right hand as quickly as they can after they hear a beep. They will do the same with their left hand. The order of the two hands for each subject is randomly determined. Data for the reaction times (in seconds) for the right and left hand are shown below, as well as data for the difference in times (d = Right − Left) for the 30 subjects.

Right	Left	d=R-L
mean = 0.17	mean = 0.26	mean = −0.09
std. dev. = 0.05	std. dev. = 0.09	std. dev. = 0.06

A hypothesis test is conducted to determine whether there is evidence that right-handed people react more quickly with their right hand than their left hand.

25. What are the hypotheses for the appropriate test of significance?

(A) $H_0 : \mu_R = \mu_L$ vs. $H_a : \mu_R > \mu_L$

(B) $H_0 : \bar{X}_R = \bar{X}_L$ vs. $H_a : \bar{X}_R < \bar{X}_L$

(C) $H_0 : \mu_d = 0$ vs. $H_a : \mu_d > 0$

(D) $H_0 : \mu_d = 0$ vs. $H_a : \mu_d < 0$

(E) $H_0 : \mu_d = \mu_R - \mu_L$ vs. $H_a : \mu_d > \mu_R - \mu_L$

26. What is the value of the test statistic for the appropriate test of significance?

(A) -8.22 (B) -6.57 (C) -4.38 (D) -7.04 (E) -5.91

27. 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

Sample Final Exam A – Part B

1. Grapes grown in a vineyard are used for making wine. The owners of the vineyard would like to conduct an experiment to examine the effect of storage time (1, 3 or 5 years) and temperature (5°C or 10°C) on the taste of their wine. They believe different types of wine (red and white) will react differently to the various treatments, and so a randomized block design is used. 60 bottles of red wine and 60 bottles of white wine are available for the experiment.

   (a) Identify each of the following in this experiment:

   (i) Factor(s)
   (ii) Treatment(s)
   (iii) Response Variable(s)
   (iv) Blocking Variable(s)

(b) Explain how the blocking should be done and how the treatments should be assigned.

(c) Suppose that one of the treatments produces much better tasting wine than the others. Can we conclude that the treatment was likely the cause?

2. The time X (in minutes) that it takes a teller to serve a customer at a bank follows some right-skewed distribution with mean 5 minutes and standard deviation 4 minutes.

   (a) Can you calculate the probability that the teller spends more than 6 minutes with the next customer? If so, do the calculation. If not, explain why.

   (b) What is the probability that the teller spends an average between 4.5 and 7 minutes with the next 30 customers?

   (c) What is the probability that the teller serves the next 30 customers in under 2 hours? (Assume there is always a customer waiting in line.)

   (d) Are the probabilities you calculated in (b) and (c) exact or approximate? Explain.

3. When the NDP formed the provincial government in 1999, the mean wait time for a particular type of surgery was 48 days. Public health officials take a sample of 30 individuals who have had the surgery in 2014 and record the number of days the patients had to wait prior to having the surgery. The mean and standard deviation of wait times for these patients are calculated to be 45.3 days and 5.7 days, respectively. Waiting times for this type of surgery are known to follow a normal distribution.

   (a) Construct a 95% confidence interval for the true mean wait time for this type of surgery in 2014.

   (b) Provide an interpretation of the confidence interval in (a).

   (c) Conduct an appropriate hypothesis test, at the 5% level of significance, to determine whether the mean wait time for this type of surgery has changed over the past 15 years. 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.

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

   (e) Could you have used the confidence interval in (a) to conduct the test in (c)? If you could, explain why, and explain what your conclusion would be, and why. If you couldn't, explain why not.

4. 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.