Sample Final – 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   294   400+   289   93   101   372   400+   153   400+   245   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 dont 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:

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

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

$$139 \qquad 178 \qquad 187 \qquad 200 \qquad 235$$

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. 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.

7. 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 faculty 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.

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. What is 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. The Central Limit Theorem states that:

   (A) when n gets large, the standard deviation of the sample mean $\overline{X}$ gets closer and closer to $\sigma/\sqrt{n}$.
   (B) if a random variable X has a normal distribution, then for any sample size n, the sampling distribution of $\overline{X}$ is also normal.
   (C) regardless of the population distribution of a random variable X, when n gets large, the sampling distribution of $\overline{X}$ is approximately normal.
   (D) if a random variable X follows a normal distribution, then when n gets large, the sampling distribution of $\overline{X}$ is exactly normal.
   (E) when n gets large, the sample mean $\overline{X}$ gets closer and closer to the population mean μ.

13. 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

14. 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)

15. 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.

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. 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) $t = -3.00$

19. 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. $H_a : \mu \neq 10$

20. 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

21. 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

22. 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.

The next **two** questions (**23** to **24**) 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.

23. What are the appropriate 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. $H_a : \mu_d > \mu_R - \mu_L$

24. 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

25. We would like to conduct a matched pairs t test to determine whether premium gasoline is more efficient than regular gasoline. We take a sample of ten different models of car. Each car drives for one tank on premium gasoline and one tank on regular gasoline (with the order randomly determined) and the mileage is recorded for each. Which of the following statements about the matched pairs t test are **true**?

(I)   For each car, mileage on premium gas and mileage on regular gas are independent.

(II)   For each car, mileage on premium gas and mileage on regular gas are dependent.

(III)   For any two cars, mileages on premium gas are dependent.

(IV)   We must assume that mileage on premium gas and mileage on regular gas are both normally distributed.

(V)   We must assume the difference in mileage (premium – regular) follow a normal distribution.

(A) I and IV only

(B) II and V only

(C) I, III and IV only

(D) I, III and V only

(E) II, III and V only

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

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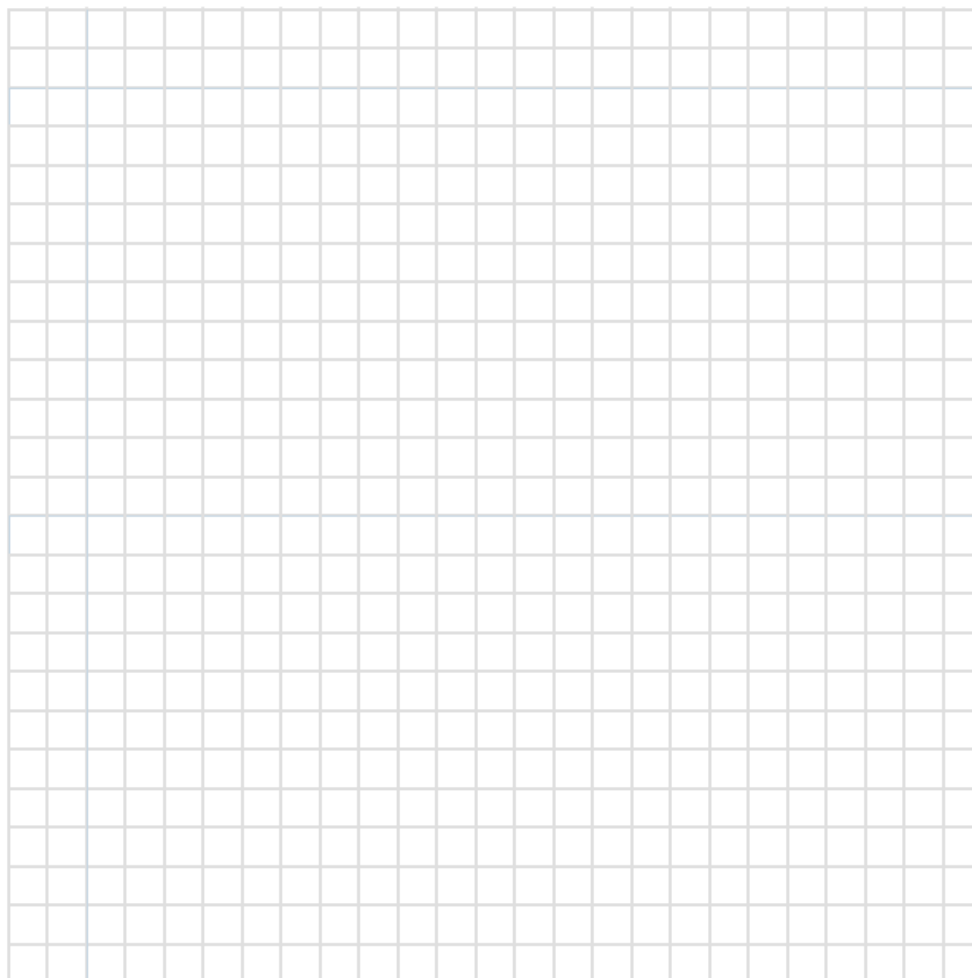
(b) Interpret the slope of the least squares regression line **in this example**.

(c) Calculate the residual for Day 5.

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

Sample Final – B
Part A

1. A consumer group surveyed the prices for white cotton twin sheet sets in five different department stores and reported the median price to be $37. We visited four of the five stores, and found the prices to be $37, $42, $37 and $35. Assuming that the consumer group is correct, what is the price of the sheet set at the store we did not visit?

   (A) $34
   (B) $35
   (C) $36
   (D) $37
   (E) We dont know – the price at the store we didnt visit could be anything.

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

$$72 \qquad 76 \qquad 78 \qquad 79 \qquad 81$$

   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.

3. The University of Manitoba uses a grade point system with a maximum possible Grade Point Average of 4.5. For each course, a student is awarded the following number of grade points per credit hour for each of the possible letter grades:

Letter Grade	Grade Points
A^+	4.5
A	4.0
B^+	3.5
B	3.0
C^+	2.5
C	2.0
D	1.0
F	0.0

One year, a student took the following courses, and received the letter grades shown:

Course	Credit Hours	Letter Grade
CIVL 2830	2	B^+
HIST 1260	3	A^+
ECE 2130	4	A
ECE 2160	5	C
COMP 2061	6	B

What is the student's GPA for the year?

(A) 3.265   (B) 3.375   (C) 3.400   (D) 3.225   (E) 3.175

4. Over a period of one month, the correlation between the daily high temperature X in Winnipeg and the daily high temperature Y in Grand Forks, North Dakota (both measured in degrees Celsius) is calculated to be 0.79. If we had measured the temperatures in degrees Fahrenheit (temperature in °F $= 32 + 1.8$ (temperature in °C)), the correlation between X and Y would be equal to:

(A) 0.79   (B) 0.44   (C) 0.89   (D) 0.68   (E) 0.72

5. We would like to examine how the age (X) of a certain model of car affects its selling price (Y). The age (in years) and price (in $) for a sample of 15 cars of the same make and model are recorded from the classified ads in the newspaper one weekend. A scatterplot shows a linear relationship between X and Y and the correlation is calculated to be $r = -0.9$. Despite the strong correlation between age and price, we cannot say that a car getting older **causes** its value to decrease. This is because of the potential effect of one or more lurking variables. Which of the following is a likely lurking variable in this case?

   (A) colour
   (B) make and model
   (C) mileage
   (D) age
   (E) all of the above

6. An experiment will be conducted to determine how the length of an icicle is affected by temperature and wind speed. Two different temperatures ($-20°$C and $-30°$C) and three different wind speeds (20 km/h, 40km/h and 60km/h) will be examined. Ten icicles will be subjected to each combination of factor levels.

   What are the treatments in this experiment?

   (A) temperature and wind speed
   (B) temperature, wind speed and icicle length
   (C) $-20°$C, $-30°$, 20 km/h, 40km/h, 60km/h
   (D) $-20°$C/20 km/h, $-30°$/20 km/h, $-20°$C/40 km/h, $-30°$/40 km/h, $-20°$C/60 km/h, $-30°$/60 km/h
   (E) temperature, wind speed and number of icicles

7. A sample of 100 University of Manitoba professors is asked whether their spouse also works in the education field. What is the population of interest?

   (A) the 100 University of Manitoba professors
   (B) all University of Manitoba professors
   (C) the spouses of the 100 University of Manitoba professors
   (D) the spouses of all University of Manitoba professors
   (E) all University of Manitoba professors whose spouses also work in the education field

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

   A random variable X is described by a semi-circular density curve with mean 0 and standard deviation 0.4, as shown below:
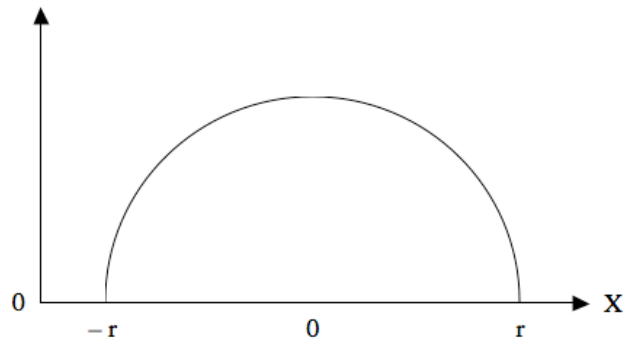
8. What must be the value of r in order for this to be a legitimate density curve? (Hint: Recall the area of the circle is equal to πr^2, where r is the radius).

(A) $\sqrt{\dfrac{1}{\pi}}$ (B) $\sqrt{2\pi}$ (C) $\sqrt{\dfrac{\pi}{2}}$ (D) $\sqrt{\pi}$ (E) $\sqrt{\dfrac{2}{\pi}}$

9. We take a simple random sample of 100 individuals from the semi-circular distribution and calculate the sample mean $\bar{x}$. The sampling distribution of $\bar{X}$ is:

(A) approximately normal with mean 0 and standard deviation 0.004.

(B) semi-circular with mean 0 and standard deviation 0.04.

(C) approximately normal with mean 0 and standard deviation 0.4.

(D) semi-circular with mean 0 and standard deviation 0.004.

(E) approximately normal with mean 0 and standard deviation 0.04.

10. The scores on a university examination are normally distributed with a mean of 62 and a standard deviation of 11. If the bottom 5% of the students get failing grades on the examination, what is the lowest mark that a student can obtain and still be awarded a passing grade?

(A) 51   (B) 57   (C) 44   (D) 40   (E) 50

11. The number of strokes X taken by professional golfers on a particular hole follows a probability distribution as shown below, where k is some constant:

x	3	4	5	6	7
$P(X = x)$	0.12	$4k$	$3k$	$2k$	0.07

What proportion of golfers take at least five strokes on this hole?

(A) 0.48 (B) 0.25 (C) 0.57 (D) 0.43 (E) 0.52

12. A student driving to university must pass through seven sets of traffic lights. Suppose it is known that each of the traffic lights is red 35% of the time and that all lights function independently. What is the probability that the student will have to stop at two or more lights on her way to university?

(A) 0.6828 (B) 0.2985 (C) 0.4893 (D) 0.7662 (E) 0.5997

The next **two** questions (**13** to **14**) refer to the following:

The fill volumes per bottle for a certain brand of beer follow a normal distribution with mean 341 ml and standard deviation 3 ml.

13. If you buy a case of 24 bottles of beer, what is the probability that the bottles will contain an average between 340 and 342 ml? (Assume that the 24 bottles can be considered a simple random sample.)

(A) 0.7359 (B) 0.8968 (C) 0.6723 (D) 0.9252 (E) 0.5870

14. If you buy a twelve-pack of beer, what is the probability that the total volume of beer will exceed 4.1 liters? (Assume that the 12 bottles can be considered a simple random sample.)

(A) 0.0838 (B) 0.2206 (C) 0.1397 (D) 0.4443 (E) 0.3632

15. 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) $2.52 \pm 1.48 \left(\dfrac{1.5}{\sqrt{35}} \right)$ (B) $2.52 \pm 1.34 \left(\dfrac{1.5}{\sqrt{35}} \right)$ (C) $2.52 \pm 0.82 \left(\dfrac{1.5}{\sqrt{35}} \right)$

(D) $2.52 \pm 0.92 \left(\dfrac{1.5}{\sqrt{35}} \right)$ (E) $2.52 \pm 1.28 \left(\dfrac{1.5}{\sqrt{35}} \right)$

16. 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 widest?

(A) 97% confidence level with $n = 25$

(B) 97% confidence level with $n = 100$

(C) 99% confidence level with $n = 25$

(D) 99% confidence level with $n = 100$

(E) depends on the value of σ.

17. 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%

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

City engineers in Hamilton, Ontario would like to estimate the true mean commuting distance of all workers in the city between home and their principal place of business. They calculate that, in order to estimate this mean to within ±1 kilometer with 99% confidence, they require a sample of 120 people.

18. What sample size would be required to estimate the true mean commuting distance for all workers in Hamilton to within ±2 kilometers with 99% confidence?

(A) 30   (B) 60   (C) 85   (D) 240   (E) 480

19. The city of Montreal has a population five times greater than that of Hamilton (and so it would be logical to assume that it has five times as many workers). Suppose we wanted to estimate the true mean commuting distance between home and work for all workers in Montreal to within ±1 kilometer with 99% confidence. Assuming equal standard deviations for the two cities, we would require a sample of how many Montreal workers?

(A) 24   (B) 120   (C) 269   (D) 600   (E) 3000

20. The level of significance α of a hypothesis test is:

(A) the lowest P-value for which the null hypothesis will be rejected.

(B) the probability that the null hypothesis will be rejected.

(C) the lowest value of the test statistic for which the null hypothesis will be rejected.

(D) the probability that the null hypothesis is true.

(E) the highest P-value for which the null hypothesis will be rejected.

21. 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.

22. A clinic measures the blood calcium of a sample of 100 healthy pregnant young women. The mean and standard deviation of these 100 measurements are, respectively, 9.8 and 0.5. Calcium levels for healthy pregnant young women are known to follow a normal distribution with standard deviation 0.4. We would like to conduct a test of significance to determine whether the data provide evidence that the mean calcium level in the population of healthy pregnant young women is less than 10. The test statistic for the appropriate test is:

(A) $t = -5.00$ (B) $z = -4.00$ (C) $z = -5.00$ (D) $t = -4.00$ (E) $z = -0.40$

23. 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.

24. Five runners were asked to run a 10-kilometre race on each of two consecutive weeks. In one of the races, the runners wear one brand of shoe and in the other a different brand (with order randomly determined). All runners are timed and are asked to run their best in each race. The results (in minutes) are given below, with some sample calculations that may be useful:

Runner	1	2	3	4	5	Mean	Std. Dev.
Brand 1	31.2	29.3	30.5	32.2	33.1	31.26	1.47
Brand 2	32.0	29.0	30.9	32.7	33.0	31.52	1.62
Difference	−0.8	0.3	−0.4	−0.5	0.1	−0.26	0.45

We wish to conduct a hypothesis test to determine if there is evidence that running times for the two brands differ. Suppose it is known that differences in times for the two brands follow a normal distribution. The P-value for the appropriate test of significance is:

(A) between 0.40 and 0.50

(B) between 0.20 and 0.30

(C) between 0.10 and 0.20

(D) between 0.05 and 0.10

(E) between 0.04 and 0.05

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. You want to estimate the proportion of Canadians who would support a merger of the federal NDP and Liberal parties to within ± 0.035 with 92% confidence. How large a sample is required?

(A) 625   (B) 784   (C) 466   (D) 816   (E) 553

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 (8%) are left handed. What are the hypotheses for the appropriate test of significance to test the researcher's suspicion?

(A) H_0: $\hat{p} = 0.08$ vs. H_a: $\hat{p} < 0.08$

(B) H_0: $\hat{p} = 0.10$ vs. H_a: $\hat{p} < 0.10$

(C) H_0: $p = \hat{p}$ vs. H_a: $p < \hat{p}$

(D) H_0: $p = 0.10$ vs. H_a: $p < 0.10$

(E) H_0: $p = 0.08$ vs. H_a: $p < 0.08$

Part B

1. Most experts agree that depression is best treated using a combination of medication and therapy. A psychiatrist plans to conduct an experiment to examine the effects of two different types of therapy (cognitive behavioural therapy and psychodynamic therapy) and three different medications (Zoloft, Paxil and Effexor). Effects of therapy and medication are believed to differ depending on the severity of a subject's depression so the psychiatrist decides to use a randomized block design. 180 subjects with mild depression and 120 subjects with severe depression are available for the experiment. At the end of one year, the improvement of each subject's condition will be evaluated.

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

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

    (b) How is randomization used in this experiment?

    (c) If individuals receiving one of the treatments improve significantly more than others, can we conclude that the treatment may be the **cause**? Why or why not?

2. Event A has probability 0.4 to occur and Event B has probability 0.7 to occur.

    (a) Are A and B mutually exclusive (disjoint)? Explain.

    (b) If A and B are independent, what is the probability $P(A \text{ or } B^c)$?

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:

$$
\begin{array}{cccccccccc}
10 & 11 & 14 & 14 & 15 & 17 & 18 & 20 & 22 & 22 \\
22 & 23 & 26 & 26 & 27 & 28 & 31 & 33 & 34 & 34 \\
36 & 42 & 44 & 49 & 50 & 55 & 62 & 68 & 77 & 82
\end{array}
$$

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 is lower than 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.

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

Sample Final – Solutions

Question	Sample Test A	Sample Test B
1	E	E
2	B	C
3	A	D
4	C	A
5	B	C
6	D	D
7	D	B
8	B	E
9	B	E
10	E	C
11	B	E
12	C	D
13	D	B
14	C	B
15	E	B
16	E	C
17	C	C
18	D	A
19	E	B
20	D	E
21	B	C
22	C	C
23	D	C
24	A	B
25	B	D
26	A	A
27	D	D