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.

Sample Final A Answers

Part A

1. E
2. B
3. B
4. E
5. B

6. C
7. D
8. D
9. A
10. E

11. B
12. B
13. D
14. B
15. B

16. C
17. D
18. E
19. C
20. E

21. D
22. D
23. B
24. E
25. D

26. A
27. A

Part B

1. (a) (i) storage time and temperature
   (ii) 1 year/5°C, 1 year/10°C, 3 years/5°C, 3 years/10°C, 5 years/5°C, 5 years/10°C
   (iii) taste of wine
   (iv) type (colour) of wine

   (c) Yes

2. (a) No

   (b) 0.7486

   (c) 0.0853

   (d) approximate

3. (a) (43.17, 47.43)

   (c) $t = -2.59$, P-value between 0.01 and 0.02, reject H_0

4. $z = 1.60$, P-value $= 0.1096$, fail to reject H_0

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.

Sample Final B Answers

Part A

1. C
2. A
3. B
4. C
5. B

6. A
7. E
8. D
9. B
10. B

11. A
12. E
13. C
14. B
15. A

16. B
17. B
18. C
19. D
20. B

21. D
22. C
23. C
24. C
25. D

26. A
27. E

Part B

1. (b) 2.1

   (c) 71.55%

2. (b) No

   (c) 0.58

3. (b) skewed to the right

   (c) $z = -1.86$, P-value $= 0.0628$, fail to reject H_0

4. (a) $t = -2.63$, P-value between 0.01 and 0.02, fail to reject H_0