1. A bin contains 50 tokens, all identical except for their colour. There are 25 red, 15 yellow, and 10 green tokens. You randomly draw 12 tokens from the bin. Find the probability of the following events.

[2 pts]   (a) None of the tokens in the draw is green.

[2 pts]   (b) At least one of the tokens in the draw is green.

[2 pts]   (c) All the tokens in the draw are the same colour.

2. Five candidates, three women and two men, are to be interviewed separately for a job. The interviews are to be scheduled in consecutive order on the same day.

[2 pts]   (a) How many ways can the interviews be scheduled?

[3 pts]   (b) How many ways can the interviews be scheduled if the first is to be with a man and the last with a woman?

3. The route used by a certain motorist crosses railroad tracks on both the way to work and on the way home. The probability that she must stop for a train on the way to work is .05, and the analogous probability for on the way home is .08. Also, the probability that she must stop in both directions is .01. What is the probability that she must stop

[2 pts]   (a) At least once during the round trip.

[2 pts]   (b) On the way home, given that she had to stop on the way to work.

[3 pts]   (c) On the way to work, but not on the way home.

[2 pts] 4. Which statement about a pair events, both having a nonzero chance of occurring, is true?

☐ If the pair A and B are independent, then they are disjoint.

☐ If the pair A and B are disjoint, then they are independent.

☐ If the pair A and B are disjoint, then they are dependent.

☐ If the pair A and B are disjoint, then so are the pair A' and B'.

[6 pts] 5. Only 1 in 500 adults is afflicted with a certain disease for which a diagnostic test has been developed. The test is such that when an individual actually has the disease, a positive result will occur 99% of the time, whereas an individual without the disease will show a positive result only 2% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease?

6. A small department at a college has 5 faculty members. Let X denote the number of faculty members from this department who are absent on any random day. The pmf of X is given in the following table.

x	0	1	2	3	4	5
$p(x)$	.820	.100	.050	.010	.015	.005

Calculate the following values.

[2 pts]   (a) The expected value of the number of faculty members absent on any random day.

[2 pts]   (b) The probability that at least two faculty members are not absent on any random day.

7. Systolic blood pressure readings of individuals are thought to be related to weight. The following data was collected from a sample of eight men by measuring their weight (x) in pounds and systolic blood pressure (y).

x	164	172	168	205	184	178	226	255
y	130	125	140	160	162	152	165	158

Some summary statistics of data:
$$\sum x = 1\,552, \quad \sum x^2 = 308\,370, \quad \sum y = 1\,192, \quad \sum y^2 = 179\,262, \quad \sum xy = 233\,584.$$

[4 pts]   (a) What is the value of the correlation coefficient r (with x as the explanatory variable)?

[2 pts]   (b) Describe the correlation between between the x and y variables.

[4 pts]   (c) What is the line of least squares?

[2 pts]   (d) What is residual associated with the third observation?

[2 pts]   8. In Ontario as a whole, over the last year, 56% of all babies born were boys. A government agency selects 100 birth records from this period, and finds that 53% of these babies were boys. We are interested only in the province of Ontario. Are the values 56% and 53% parameters or statistics? Select your answer.

☐ 56% is a statistic and 53% is a parameter.

☐ 56% and 53% are both parameters.

☐ 56% and 53% are both statistics.

☐ 56% is a parameter and 53% is a statistic.

9. The amount the first-time homebuyer spends in Quebec has a normal distribution with a mean of \$225 000 and a standard deviation of \$5 800.

[2 pts]   (a) What is the probability that a first-time homebuyer in Quebec will spend between \$225 000 and \$230 000?

[2 pts]   (b) Suppose we consider a random sample of 20 first-time homebuyers in Quebec. What is the probability that the average amount they spent on their homes is between \$224 000 and \$227 000.

[6 pts] 10. In order to claim a prize a potential winner in a certain contest must correctly answer a skill testing question (STQ). Out of all participants in the contest 15% are unable to answer the STQ correctly. In a random sample of 500 potential winners, what is the approximate probability that between 400 and 450 will answer the STQ correctly?

11. Suppose that X is a normally distributed random variable with mean $\mu = 350$ and standard deviation $\sigma = 40$.

[4 pts]   (a) Find the 85th percentile of the distribution of X.

[4 pts]   (b) If $\overline{X}$ = mean of a random sample of size 16, find $P(|\overline{X} - 365| \le 5)$.

[5 pts] 12. A political party wants to conduct a survey to find the level of support for their new leader. From the survey they wish to calculate a 98% confidence interval for the proportion of the population who would vote for the party's leader. Without any prior estimates for the leader's support what sample size should be used to obtain an interval with a (maximal) margin of error less than or equal to 3 percentage points.

[5 pts] 13. A random sample of 20 weights of newborn babies was taken at the Lakeshore hospital. The sample had a mean of 6.87 lb and a standard deviation of 1.76 lb. It is assumed that weights of newborns are normally distributed. Using this sample compute a 95% confidence interval for the mean weight of all babies born in this hospital.

[6 pts] 14. What is your favourite colour? About 24% of the population claim blue as their favourite colour. Suppose a random sample of 56 college students were surveyed and 12 of them said that blue is their favourite colour. Does this information imply that the colour preference of all college students is different (either way) from that of the general population? Define the parameter of interest and then state and test the relevant hypotheses at a significance level of $\alpha = .05$

[2 pts] 15. Suppose we intend to calculate a 99% confidence interval for a population mean. Which of the statements below best describes the meaning of the 99%?

☐ The probability that the population mean will lie between the endpoints of a 99% confidence interval calculated from one sample is 99%.

☐ If you draw a sample from this population, there is a 99% probability that the population mean will be between the highest and lowest sample values.

☐ 99% of the values in the population will be close to the population mean.

☐ Out of all the samples we could possibly draw from this population, only 1% of them will produce a 99% confidence interval that does not contain the population mean.

16. A manufacturing company claims that its new floodlights have an average lifetime of 1 200 hours. A consumer watch group is skeptical of this claim. From a simple random sample of size 36, the watch group computes a mean of 1 036 hours and standard deviation of 362 hours.

[6 pts]   (a) Define the parameter of interest and then state and test the relevant hypotheses at a significance level of 0.025.

[2 pts]   (b) If the conclusion of the test in part (a) is an error, would it be of type I or of type II?

[6 pts] 17. An economy and premium brand of rope both claim the same breaking strength of 7 000 lbs. Is the economy brand weaker? The following data was collected:

Brand	sample size	sample mean	sample sd
premium	5	7 100	50
economy	8	7 010	60

State and test the relevant hypotheses at a significance level of .025.

[6 pts] 18. The psychologist David Keirsey developed a theory of temperaments in which people are divided into four temperament categories; artisans, guardians, idealists, and rationals. A study on personality traits considered whether temperament was influenced by having siblings. This study involved a simple random sample of the population from which the data in the following table was obtained.

		Temperament				
		Artisan	Guardian	Idealist	Rational	Row total
Only child	Yes	6	5	8	10	29
	No	16	18	12	25	71
Column total		22	23	20	35	100

Does the data suggest that someone's temperament category is independent of them being an only child or not. State and test the relevant hypotheses at a significance level of $\alpha = .10$. Interpret your conclusion in the the context of the study.

Answers:

1. (a) 0.04602

    (b) 0.9540

    (c) 0.00004284

2. (a) 120

    (b) 36

3. (a) 0.12

    (b) 0.2

    (c) 0.04

4. If the pair A and B are disjoint, then they are dependent.

5. 0.09025

6. (a) 0.315

    (b) 0.98

7. (a) 0.6731

    (b) positive (moderate-large)

    (c) $86.77 + 0.32x$

    (d) -0.53

8. 56% is a parameter and 53% is a statistic.

9. (a) 0.3051

    (b) 0.7176

10. 0.9986

11. (a) 391.6

    (b) 0.1359

12. 1509

13. (6.046,7.694)

14. $p =$ proportion of students whose favourite colour is blue.
    $H_0 : p = .24$, $H_a : p \neq .24$, test statistic: $z = -0.45$, P-value $= 0.6528 > \alpha$.
    Do not reject H_0.

15. Out of all the samples we could possibly draw from this population, only 1% of them will produce a 99% confidence interval that does not contain the population mean.

16. $\mu =$ mean lifetime of the company's new floodlights.
    $H_0 : \mu = 1\,200$, $H_a : \mu < 1\,200$, test statistic: $t = -2.72$ ($df = 35$), $.005 < P$-value $< .01 < \alpha$.
    Reject H_0.

17. $\mu_1(\mu_2) =$ mean strength of the rope from the premium (economy) brand.
    $H_0 : \mu_1 - \mu_2 = 0$, $H_a : \mu_1 - \mu_2 > 0$, test statistic: $t = 2.92$ ($df = 4$), $.01 < P$-value $< .025 = \alpha$. Reject H_0.

18. H_0: Temperament and being an only child are independent. H_a : Temperament and being an only child are dependent.
    Test statistic: $\chi^2 = 1.799$, $\alpha = 0.1 < P$-value < 0.9. Fail to reject H_0.
    There is not enough evidence to support the claim that there is a relation between a person's temperament category and whether or not they are an only child.