1. An experiment will attempt to determine if a teacher discriminates against certain colours of clothing. In a group of 20 students, 15 are wearing a white shirt while the other 5 are wearing a pink shirt. Without knowing the purpose of the experiment, a teacher is told to give 8 cookies to 8 different students from this group.

   (a) (1 point) Suppose the teacher selects a sample of 8 students randomly for this experiment. How many ways can this be done?

   (b) (2 points) Calculate the probability of obtaining a sample representative of the group, that is one with 6 white and 2 pink shirts.

   (c) (2 points) Calculate the probability of obtaining a sample with equal proportions, that is one with 4 white and 4 pink shirts.

   (d) (3 points) If only 4 cookies are to be given to 4 different students, what is the probability that the teacher, choosing students randomly, will give his first two cookies to pink shirts, followed by two cookies to white shirts?

2. According to *Statistics Canada*, the following table classifies the number of young people who died in Canada during the year 2015 from the three leading causes of death: accidental, suicide and cancer.

	Accidental (A)	Suicide (S)	Cancer (C)	Total
Age 15-24 (X)	706	518	150	1374
Age 25-34 (Y)	940	675	408	2023
Total	1646	1193	558	3397

Suppose that one person is selected at random from this group.

   (a) (2 points) Find the conditional probabilities $P(X|S)$ and $P(S|X)$.

   (b) (2 points) Find the probability that the randomly selected person was from the age group 25-34 and did not die of cancer.

   (c) (2 points) Using this data, give an example of two events that are mutually exclusive and justify.

   (d) (2 points) Are the events X and S independent? Justify your answer mathematically.

3. Tromp University has two campuses: a downtown campus and a regional campus. The composition of the student body varies between the two locations. At the downtown campus, the population of students consists of 70% Canadian and 6% American citizens, the rest being other international students. At the regional campus, the composition is 81% Canadian and 10% American citizens. In addition, it is known that 90% of all students attend the downtown campus.

   (a) (4 points) Given that Carlos is a student at Tromp University, and that he is neither Canadian nor American, what is the probability that he attends his classes at the downtown campus?

   (b) (2 points) Find the percentage of Canadian students at Tramp University.

4. Consider the random variable X to be the number of minutes a patient spends in the waiting room before being called by a psychiatrist for a consultation, relative to the scheduled appointment. A negative number means that the patient is called *before* their scheduled appointment.

Relative wait time in minutes (x)	-5	0	5	15
Probability $(P(x))$	0.05	0.30	0.62	?

  (a) (1 point) Find the missing probability.

  (b) (2 points) Calculate the expected waiting time.

  (c) (3 points) Calculate the standard deviation of this random variable.

  (d) (2 points) What is the probability of waiting at most 5 minutes, given that you are not called before your scheduled appointment?

5. At one metro station, suppose only 35% of people will climb up the stairs instead of using the escalator.

  (a) (2 points) In a random sample of 10 independent people, calculate the probability that exactly 1 of them will use the **stairs**.

  (b) (4 points) In a random sample of 10 independent people, calculate the probability that at most 8 of them will use the **escalator**.

6. Bobsung's new BX smartphone has an average battery life of 15 hours with a standard deviation of 2.9 hours. Assume that the battery life of a phone is normally distributed.

  (a) (3 points) Suppose a new BX smartphone is chosen at random. What is the probability that its battery life is between 12 hours and 20 hours?

  (b) (3 points) Which battery life delimits the longest 10% of battery lifetimes of all the BX smartphones?

  (c) (4 points) An electronics store receives a shipment of 25 BX smartphones. What is the probability that their average battery life is longer than 16 hours?

7. A Loonie Line game consists of giving a loonie to spin a wheel of fortune, for the chance of winning various delicious baked goods. The wheel has a 4% probability of falling on the grand prize, a full plate of brownies. Suppose that 180 students spin the wheel during the week.
Let r be the number of grand prize winners.

  (a) (1 point) What is the expected value of r?

  (b) (1 point) What is the standard deviation of r?

  (c) (4 points) Use a normal approximation to find the probability that more than 15 students win the grand prize.

8. Élections Québec is undertaking a study among Quebec CEGEP students who are eligible to vote in the coming provincial elections. The government body wants to estimate p, the probability that an eligible student will vote in the coming provincial elections.

  (a) (3 points) What is the minimum sample size needed to estimate p to within 3% of its true value, with a confidence level of 92%?

(b) (2 points) By how many could the minimum sample size be reduced if Élections Québec uses the proportion of students who voted in the last election, $\hat{p}=0.32$, as a preliminary estimate for p?

9. A team of psychologists is developing a risk-aversion quantification tool as a behavioural predictor. The tool gives a numerical score based on various tests. It is assumed that the scores are normally distributed in the general population.
   Preliminary results from 20 participants show an average risk-aversion score of 47 with a sample standard deviation of 10.7.

   (a) (5 points) Construct a 99% confidence interval for the true mean risk-aversion score in the population.

   (b) (1 point) Interpret your confidence interval from part (a).

   (c) (2 points) How would the width of the confidence interval in part (a) be affected if we knew the standard deviation of risk-aversion scores for the population was equal to that obtained from the sample? Briefly explain.

10. A particular type of avalanche studied in Canada had an average thickness of $\mu = 67$ cm. This type of avalanche was studied in a region of the southwest United States. A random sample of 16 such avalanche thicknesses had a mean of 73 cm and a known standard deviation $\sigma = 11.3$ cm. Assume this thickness has an approximately normal distribution. Use a 5% level of significance to test the claim that this mean avalanche thickness in this US region is different from that in Canada.

    (a) (2 points) State the null and alternative hypotheses.

    (b) (3 points) Calculate the test statistic.

    (c) (3 points) Find (or estimate) the p-value for the test.

    (d) (2 points) State and **interpret** your conclusion.

11. A dietician wishes to see if a person's cholesterol level will change if their diet is supplemented by a certain mineral. Five subjects were pretested, and then they took the mineral over a six-week period. The results (in milligrams per decilitre) follow:

Before	200	235	208	190	172
After	190	210	210	188	173

    Suppose that cholesterol levels are normally distributed. Use a 1% significance level to test the hypothesis that cholesterol levels have changed.

    (a) (2 points) State the null and alternative hypotheses.

    (b) (3 points) Calculate the test statistic.

    (c) (3 points) Find (or estimate) the p-value for the test.

    (d) (2 points) State and **interpret** your conclusion.

12. A researcher wants to determine whether there is a relationship between the gender of an individual and the amount of alcohol consumed. A sample of 68 people is selected and the following data is obtained.

	Male	Female	Total
Light/Non-Drinker	10	13	23
Moderate Drinker	9	16	25
Heavy Drinker	8	12	20
Total	27	41	68

Test the researcher's hypothesis at the 10% level of significance that one's gender does in fact influence one's alcohol consumption.

(a) (2 points) State the null and alternative hypotheses.

(b) (3 points) Calculate the test statistic.

(c) (3 points) Find (or estimate) the p-value for the test.

(d) (2 points) State and **interpret** your conclusion.

13. Dawn is comparing the fat content (in grams) of Healthy Snack trail mix with that of Overland trail bars. The fat content of each is known to be normally distributed.
A summary of her findings is given below:

	Sample Size	Sample Mean Fat Content	Sample Standard Deviation
Healthy Snack Trail Mix	12	5.3	1.2
Overland Trail Bars	10	4.2	1.6

(a) (3 points) Find an 85% confidence interval for the true average difference between the fat content of the two varieties of trail bars.

(b) (2 points) Fill in the blank by circling one of the choices that follows.

There is _____ a difference in the true mean fat content of the two snacks.

certainly   probably   probably not   certainly not   no way to determine if there is

ANSWERS

1. (a) $C_{20,8} = 125,970$

    (b) $\dfrac{C_{15,6} \cdot C_{5,2}}{C_{20,8}} = 0.3973$

    (c) $\dfrac{C_{15,4} \cdot C_{5,4}}{C_{20,8}} = 0.0542$

    (d) 0.0361

2. (a) $P(X|S) = 0.4342$ and $P(S|X) = 0.3770$

    (b) 0.4754

    (c) X and Y are mutually exclusive, since a person cannot belong in both the 15-24 and 25-34 age categories

    (d) $P(X|S) = 0.4342$ is not equal to $P(X) = 0.4045$, therefore the two events are not independent

3. (a) 0.96

    (b) 71.1%

4. (a) 0.03

    (b) 3.3

    (c) 3.5511

    (d) 0.9684

5. (a) 0.0725

    (b) 0.9140

6. (a) 0.8058

    (b) 18.712 hours

    (c) 0.0427

7. (a) 7.2

    (b) 2.629

    (c) 0.0008

8. (a) 851 students

    (b) 110 fewer than without a preliminary estimate

9. (a) $C.I. = [\bar{x} - E, \bar{x} + E] = [40.155, 53.845]$

   (b) We are 99% confident that the true mean risk-aversion score in the population is between 40.155 and 53.845.

   (c) The margin of error would be calculated with $z_c = 2.576 < t_c$, narrowing the width of the confidence interval.

10. (a) $H_0 : \mu = 67; H_a : \mu \neq 67$

    (b) $z^* = 2.12$

    (c) P-value ≈ 0.034

    (d) Reject H_0

    There is sufficient evidence to claim that the true mean avalanche thickness in this U.S. region is different from that in Canada.

11. (a) $H_0 : \mu_d = 0; H_a : \mu_d \neq 0$

    (b) $t^* = 1.356$

    (c) $0.200 < P$-value < 0.250

    (d) Fail to Reject H_0

    There is insufficient evidence to claim that the cholesterol levels have changed.

12. (a) H_0 : gender and alcohol consumption are not related.
    H_a:gender and alcohol consumption are related.

    (b) $\chi^{2,*} = 0.2808$

    (c) $0.100 < P$-value < 0.900

    (d) Fail to Reject H_0

    There is insufficient evidence to claim that gender and alcohol consumption are related.

13. (a) $(0.135, 2.065)$

    (b) There is **probably** a difference in the true mean fat content of the two snacks.