1. Tristan claims that he has the ability to roll a die in such a way that 6 will almost always occur. You test him by giving him a fair die, which he proceeds to roll five times, getting a 6 each time.

   (a) (2 points) If Tristan has no control over the die, how many outcomes are possible with five rolls of the die?

   (b) (2 points) What is the probability that Tristan gets five 6's in five rolls of the die?

2. A carton contains 12 eggs, 3 of which are cracked. If we randomly select 5 eggs for hard boiling, what is the probability of the following:

   (a) (2 points) None of the cracked eggs are selected.

   (b) (2 points) All of the cracked eggs are selected.

   (c) (1 point) At least one of the cracked eggs is selected.

3. The following table classifies the number of extroverts and introverts among a group from four different countries:

	Extrovert (E)	Introvert (I)	Total
Canada (C)	120	180	300
Finland (F)	80	200	280
Japan (J)	100	140	240
Australia (A)	100	80	180
Total	400	600	1000

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

   (a) (1 point) Find the probability that the randomly selected person is introverted.

   (b) (2 points) Find the conditional probability $P(E|C)$.

   (c) (2 points) Find the probability that the randomly selected person is introverted and from Australia.

   (d) (3 points) Find the probability that the randomly selected person is extroverted, given they are not from Japan.

   (e) (2 points) Are the events E and C independent? Justify your answer mathematically.

4. At one CÉGEP, the breakdown among social science students is as follows: 48% of all social science students are male. Among males in the social science program, 65% are not in the psychology profile. Among social science students in the psychology profile, there is an equal number of males and females.

   (a) (3 points) What proportion of the social science students at this CÉGEP are not in the psychology profile?

   (b) (4 points) Danika is a social science student at this CÉGEP. What is the probability that **she** is in the psychology profile?

5. Consider the random variable X to be the number of texts sent by 14-year-old Simon during a 30-minute interval after school one afternoon.

Number of texts (x)	0	1	2	3	4
Probability $(P(x))$	0.13	0.25	0.26	0.23	?

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

(b) (2 points) Calculate the expected number of texts.

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

(d) (2 points) What is the probability that Simon will send more than the expected number of texts during this half hour?

6. An old professor has found that over the years, the class marks for his Psychology 101 course are normally distributed with a mean of 71 and a standard deviation of 5.5.

(a) (4 points) Only students with grades in the top 5% of the class will be considered for a summer stage at the local hospital. Audrey is hoping to be one of them. What class mark will she need to earn?

(b) (4 points) This semester, there are 25 students enrolled in his class. What is the probability that they will have an average class mark that is greater than 74?

7. In Québec, 45% of the population is bilingual. Suppose that at a peewee hockey match, there are 60 spectators cheering the players on. Let r be the number of these spectators who are bilingual.

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

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

(c) (2 points) Is it reasonable to use the normal distribution to approximate the distribution of bilingual spectators at this hockey match? Justify your answer mathematically.

(d) (4 points) If appropriate, use a normal approximation to find the probability that at least 30 but less than 35 of these spectators are bilingual.

8. (5 points) In Toronto, approximately 11% of anglophones can also speak French. Consider a situation where a Parisian tourist is lost in Toronto and speaks no English. If he encounters a group of 6 anglophones waiting at a bus stop, what is the probability that no more than two of them will be able to help him find his way?

9. (5 points) Many studies have attempted to show a link between playing violent video games and aggressive behavior. Suppose the following experiment is conducted. Each test subject plays either a violent or a non-violent game for 20 minutes, and is then asked to complete a few short stories by describing what happens next. The number of aggressive responses suggested in completing the stories are recorded for each test subject. The data obtained are normally distributed and summarized in the following table.

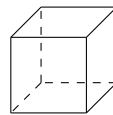|	non-violent game	violent game
sample size	51	101
sample average	6.2	7.5
sample standard deviation	2.142	1.706

Construct a 99% confidence interval for the difference in the true average number of aggressive responses between those playing non-violent and violent video games.

10. The Necker cube is an optical illusion: when we draw a cube in two dimensions with no visual cues, it can be interpreted spatially in two distinct ways.
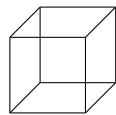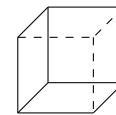
Necker cube   Interpretation A   Interpretation B

After surveying a random sample of 80 people with university degrees, it was found that 67 of them would naturally associate the Necker cube with Interpretation A.

(a) (5 points) Construct a 93.5% confidence interval for the true proportion.

(b) (1 point) Interpret in words your result from part (a), making sure to properly identify the targeted population in this study.

(c) (3 points) How many additional people would need to be surveyed in order to estimate the true proportion within 0.02 at the 95% confidence level?

11. The resting heart rate (in beats per minutes, or bpm) for five people is measured before training for a half-marathon. Several months later, the day following the race, the resting heart is again measured for the same five people. The raw data is shown below.

	1	2	3	4	5
resting heart rate before training	92	74	61	78	68
resting heart rate after race	80	70	63	61	59

Test the hypothesis that training for a half-marathon reduces the resting heart rate (which is a good thing!), using a 2% significance level. Assume that the differences in the resting heart rate are normally distributed across the population.

(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) (1 point) State your conclusion.

   (e) (1 point) Is there sufficient evidence to claim that the training reduces the resting heart rate?

   (f) (1 point) Would your conclusion change if you test the hypothesis that training for a half-marathon **increases** the resting heart rate? Explain.

12. A short quiz consisting of 10 simple arithmetic questions was administered to two groups of students. For the first group of 25 students, the test was given in optimal conditions, in a quiet room. The average score for those students was 8.7 correct answers. For a different group of 20 students, the test was given in more stressful conditions: the invigilator was constantly reminding the students how much time was left and telling them to hurry up. Their average score turned out to be 7.3 correct answers.

From a previous large study with the same questions, it was found that the scores were normally distributed, with a population standard deviation of 1.2 in optimal conditions and 2.6 in stressful conditions.

Test the hypothesis that there is a difference in the performance of students submitted to optimal conditions and those submitted to more stressful conditions, using a 5% significance level.

   (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. In 2014, a student poll at the University of Toronto revealed that 46% of students get 6 hours of sleep or less, 35% get 6.5 or 7 hours of sleep, and 19% get 7.5 or 8 hours of sleep per night on average. Suppose a similar poll is conducted with 80 John Abbott students; the results are tabulated below.

hours of sleep per night	6 or less	6.5 or 7	7.5 or 8
number of students	27	32	21

Perform a test using a 5% significance level to show whether the distribution of sleep hours among John Abbott students matches with the distribution obtained from the University of Toronto students.

   (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.

ANSWERS

1. (a) $6^5 = 7776$

   (b) $\dfrac{1}{7776} = 0.0001286$

2. (a) $\dfrac{C_{9,5}}{C_{12,5}} = 0.1591$

   (b) $\dfrac{C_{3,3} \cdot C_{9,2}}{C_{12,5}} = 0.04545$

   (c) $1 - \dfrac{C_{9,5}}{C_{12,5}} = 0.8409$

3. (a) $P(I) = \dfrac{600}{1000} = 0.6$

   (b) $P(E|C) = \dfrac{120}{300} = 0.4$

   (c) $P(I \text{ and } A) = \dfrac{80}{1000} = 0.08$

   (d) $P(E|J^C) = \dfrac{300}{760} = 0.395$

   (e) $P(E) = 0.4$ is equal to $P(E|C)$, therefore the two events are independent.

4. (a) 0.664

   (b) $P(\text{psych}|\text{female}) = \dfrac{168}{520} = 0.323$

5. (a) 0.13

   (b) 1.98

   (c) 1.2327

   (d) 0.62

6. (a) 80

   (b) 0.0032

7. (a) 27

   (b) 3.854

   (c) Yes. $np = 27 > 5$ and $nq = 33 > 5$.

   (d) 0.2322

8. 0.9794

9. $C.I. = [-2.2230, -0.3770]$

10. (a) $C.I. = [0.7612, 0.9138]$

   (b) We are 93.5% confident that the true proportion of people with university degrees who naturally associate the Necker cube with Interpretation A is between 76.12% and 91.38%.

   (c) 1228 additional people

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

   (b) $t^* = 2.446$

   (c) $0.025 < P\text{-value} < 0.050$

   (d) Fail to Reject H_0

   (e) No.

   (f) No. The p-value would be much larger and so we would still fail to reject the null hypothesis.

12. (a) $H_0 : \mu_1 = \mu_2; H_a : \mu_1 \neq \mu_2$

   (b) $z^* = 2.23$

   (c) $P\text{-value} = 0.0258$

   (d) Reject H_0

   There is sufficient evidence to claim that there is a difference in the performance of students submitted to optimal conditions and those submitted to more stressful conditions.

13. (a) H_0 : The sleep hours distribution among J.A.C. students matches the distribution obtained from the University of Toronto.
   H_a:The sleep hours distribution among J.A.C. students does not match the distribution obtained from the University of Toronto.

   (b) $\chi^{2,*} = 5.3944$

   (c) $0.050 < P\text{-value} < 0.100$

   (d) Fail to Reject H_0

   There is insufficient evidence to claim that the distributions do not match.