1. French Tarot is played with a deck of 78 cards: the same 52 cards as a standard deck, plus 4 Knights (one in each suit), a separate 21-card trump suit (numbered 1 to 21) and a single card known as the Fool (playing the mandolin).

[2]   (a) Drawing 1 card from a French Tarot deck, find the probability that the card is a Knight or the Fool.

[2]   (b) Drawing 2 cards from a French Tarot deck, without replacement, find the probability that the first card is a Knight and the second card is the Fool.

2. Pizzeria da Luigi offers a choice of 3 different crusts, 2 bases, 5 cheeses and 18 toppings including the controversial pineapple.

[2]   (a) How many possible different pizzas can be made by choosing one crust, one base, 2 different cheeses and 3 different toppings?

[3]   (b) If you choose the three toppings randomly, what is the probability that you end up with a pizza with pineapple on it?

[3]   (c) If 10 toppings out of 18 are vegetables and you choose all three toppings randomly, what is the probability to end up with a pizza with at least one vegetable topping?

3. The table below describes the intention to get vaccinated for Covid-19 among Canadians and Americans. (Source: Leger, April 12 2021)

	Canada (C)	USA (U)	Total
Yes (Y)	1203	611	1814
No (N)	180	240	420
Don't know (D)	121	151	272
Total	1504	1002	2506

Find the probability that a randomly selected person from this group:

[2]   (a) intends to get vaccinated.

[2]   (b) intends to get vaccinated given that the person lives in Canada.

[2]   (c) lives in the USA or Doesn't know.

[2]   (d) Are the events Yes (Y) and Canada (C) independent? Justify your answer mathematically.

[6] **4.** According to a Leger survey on April 12, 70% of Quebecers were satisfied with the measures put in place to fight the pandemic by their provincial government, and 60% of Ontarians were dissatisfied with theirs. In this survey, 38% of responders live in Quebec. Given that a responder was dissatisfied at the time of the survey, what is the probability that the responder lives in Ontario?

5. Let the random variable X be the age of a student at the college's gardening club. Below is the probability distribution for X:

x	16	17	18	19	20	21
$P(x)$	0.05	0.11	0.21	0.26	?	0.15

[2]   (a) Find the missing probability.

[2]   (b) Calculate the expected value.

[3]   (c) Calculate the standard deviation.

[2]   (d) What is the probability that a randomly selected student is younger than the expected age?

[3]   (e) If we consider a random sample of 10 independently selected students from the gardening club, what is the probability that exactly two of them will be 18 years old?

[6]   **6.** In the United States, a Household Pulse Survey reveals that an astonishing 42.7% of adults aged 18-29 are reporting symptoms of anxiety in March 2021 (this percentage was a lot lower before the pandemic).

Consider a random sample of 200 adults aged 18-29 in the United States. Use a normal approximation to calculate the probability that more than 75 but less than 100 people are reporting anxiety symptoms in this sample.

7. SkipTheEats is a new food delivery service that is so slow, it almost forces you to skip the meal entirely. Suppose the delivery times are normally distributed with a mean of 150 minutes and a standard deviation of 33 minutes.

[4]   (a) You are hoping that the next delivery will be among the quickest 15% deliveries of SkipTheEats. If this is true, what is the maximum time you are going to wait for your food?

[4]   (b) Suppose that next week you will make 7 orders using SkipTheEats. What is the probability that the average delivery will exceed 2 hours?

8. A mobile phone company in Montreal wants to determine the current percentage of customers aged 70+ who use text messaging on their cell phones. After surveying a random sample of 105 people aged 70+, it was found that 67 of them use text messaging on their cell phones.

[5]   (a) Construct a 95% confidence interval for the true proportion.

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

[3]   (c) How many **additional people** would need to be surveyed in order to estimate the true proportion within 0.02 at the 90% confidence level?

9. A new drug is used to treat Covid-19. The following data represent the remission times of Covid-19 (in days) for a random sample of 21 patients using this new drug:

$$10 \quad 7 \quad 32 \quad 23 \quad 22 \quad 6 \quad 16 \quad 34 \quad 32 \quad 25 \quad 11$$
$$20 \quad 19 \quad 6 \quad 17 \quad 35 \quad 6 \quad 13 \quad 9 \quad 6 \quad 10$$

The sample mean is 17.1 days, with a sample standard deviation of 10 days. Assume the distribution of remission times is approximately normally distributed.

It was previously determined that, without the drug, Covid-19 patients have a mean remission time of 12.5 days. Test the hypothesis that the mean remission time from Covid-19 when using the new drug is different from 12.5 days, using a 1% significance level.

[3]   (a) State the null and alternative hypotheses.

[3]   (b) Calculate the test statistic.

[2]   (c) Find (or estimate) the p-value for the test.

[1]   (d) State your conclusion.

[2]   (e) Would your conclusion change if you use 10% significance level? Explain.

10. A botanist knows that for a certain plant, the proportion of seeds germinating is 80%. The proportion of seeds germinating for a new hybrid plant is unknown, but the botanist claims that it is also 80%. To test this claim, 400 seeds from the hybrid plant are tested, and it is found that 312 germinate. Use a 5% level of significance to test the claim that the proportion germinating for the hybrid differ from 80%.

[3]   (a) State the null and alternative hypotheses.

[3]   (b) Calculate the test statistic.

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

[2]   (d) State and **interpret** your conclusion.

11. Each night, a portion of your sleep time is called REM (rapid eye movement) sleep. Assume that REM sleep time is normally distributed for both children and adults. A random sample of 12 children (9 years old) showed that they had an average REM sleep time of 2.8 hours per night with a sample standard deviation of 0.5 hour. Another random sample of 15 adults showed that they had an average REM sleep time of 2.1 hours per night with a sample standard deviation of 0.7 hour. Does this data indicate that, on average, 9-year-old children tend to have more REM sleep than adults? Use a 1% level of significance.

[3]   (a) State the null and alternative hypotheses.

[3]   (b) Calculate the test statistic.

[2]   (c) Find (or estimate) the p-value for the test.

[2]   (d) State and **interpret** your conclusion.

[4]   (e) Let μ_1 be the average REM sleep time of all 9-year-old children and let μ_2 be the average REM sleep time of all adults. Find a 98% confidence interval for $\mu_1 - \mu_2$.

[2] **12.** Can a p-value be greater than 1? Explain.

ANSWERS:

1. (a) $\dfrac{5}{78} \approx 0.0641$

   (b) $\dfrac{4}{78} \cdot \dfrac{1}{77} \approx 0.000666$

2. (a) 48960

   (b) 0.1667

   (c) 0.9314

3. (a) $\dfrac{1814}{2506} \approx 0.7239$

   (b) $\dfrac{1203}{1504} \approx 0.7999$

   (c) $\dfrac{1123}{2506} \approx 0.4481$

   (d) not independent since $P(Y) \neq P(Y|C)$

4. 0.7654

5. (a) 0.22

   (b) 18.94

   (c) 1.3843

   (d) 0.37

   (e) 0.3011

6. 0.9005

7. (a) 115.68 minutes

   (b) 0.9920

8. (a) $(0.5462, 0.7300)$

   (b) We are 95% confident that the true percentage of all customers aged 70+ using text messaging on their phone is between 54.62% and 73%.

   (c) 1458 additional people

9. (a) $H_0 : \mu = 12.5, \quad H_a : \mu \neq 12.5$

   (b) $t^* \approx 2.1080$

   (c) $0.02 < p\text{-value} < 0.05$

   (d) fail to reject H_0

   (e) yes, the conclusion would change to *reject the null hypothesis*

10. (a) $H_0 : p = 0.8, \quad H_a : p \neq 0.8$

    (b) $z^* \approx -1.00$

    (c) $p\text{-value} \approx 0.3174$

    (d) fail to reject H_0

11. (a) $H_0 : \mu_c - \mu_a = 0, \quad H_a : \mu_c - \mu_a > 0$

    (b) $t^* \approx 3.0264$

    (c) $0.005 < p\text{-value} < 0.010$

    (d) reject H_0

    (e) $(0.0713, 1.3287)$

12. NO. The p-value is defined as the *probability* of observing an effect at least as extreme as the one we observed in our sample, and we know probabilities cannot be greater than 1. Alternatively, the p-value can be obtained by finding an *area* under a continuous distribution (such as the standard normal or the t distribution), and we know the total area under a distribution is 1.