

The page is to be used as scrap paper only.

PART A - MULTIPLE CHOICE QUESTIONS

Clearly indicate your answer (A)-(E) in the "Answer" spot. Each correct answer is worth 2 marks.

1. A farmer is recording the weights of 12 pumpkins harvested from their field. The weights (in pounds) of the pumpkins are as follow:
10.1, 12.5, 14.4, 10.5, 9.6, 8.4, 14.8, 12.8, 12.7, 12.8, 10.0, 20.1
- i) Classify the variable "weight":
(A) Quantitative/Discrete
(B) Quantitative/Continuous
(C) Qualitative/Continuous
(D) Qualitative/Discrete
(E) Quantitative/Normal
- Answer: _____
- (ii) State the mode and range of the data.
(A) Mode = 12.6 Range = 11.7
(B) Mode = 12.8 Range = 10
(C) Mode = 12.6 Range = 8.4
(D) Mode = 12.8 Range = 11.7
(E) Mode = 12.7 Range = 18.1
- Answer: _____
- (iii) State the median and the mean of the data.
(A) Median = 12.6 Mean = 12.10
(B) Median = 11.6 Mean = 12.10
(C) Median = 12.6 Mean = 12.39
(D) Median = 11.6 Mean = 12.39
(E) Median = 12.1 Mean = 12.60
- Answer: _____
- (iv) Calculate the interquartile range of this data. Then chose the *closest* of the five values below:
(A) 5.75
(B) 4.75
(C) 3.75
(D) 2.75
(E) 1.75
- Answer: _____
- (v) Using the interquartile range method, that is, fences with values $Q1-(1.5)(IQR)$ and $Q3+(1.5)(IQR)$, which of the following value(s) can be considered outliers?
(A) $x=20.1$.
(B) $x=8.4$
(C) $x=8.4, 9.6$
(D) $x=8.4, 20.1$
(E) There are no outliers as all data is between the fences.
- Answer: _____

The page is to be used as scrap paper only.

2. In 2023, there were two by-elections in the city of Regina: in Coronation Park and in Walsh Acres. The following table lists the votes cast in both areas for the New Democratic Party, the Saskatchewan Party, and Other Parties. Note that two sections of the table are blank.

	New Democratic Party	Saskatchewan Party	Other Parties	Total
Regina Coronation Park	2173	1155	446	3774
Regina Walsh Acres	2535	1842	???	4643
Total	4708	2997	???	8417

- (i) How many votes were cast for Other Parties in the Regina Walsh Acres election?
 (A) 712
 (B) 446
 (C) 262
 (D) 266
 (E) 978
 Answer: _____
- (ii) A random voter is selected. What is the probability that they voted in Regina Coronation Park OR that they voted for the New Democratic Party? Round to four decimals.
 (A) 0.2582
 (B) 0.5593
 (C) 0.5758
 (D) 0.7496
 (E) 0.4616
 Answer: _____
- (iii) A random voter is selected. What is the probability that they voted in Regina Coronation Park GIVEN THAT they voted for the New Democratic Party? Round to four decimals.
 (A) 0.2582
 (B) 0.5593
 (C) 0.5758
 (D) 0.7496
 (E) 0.4616
 Answer: _____
- (iv) Are events A: "voted in Regina Coronation Park" and B: "voted for the New Democratic Party" independent?
 (A) Yes, because $P(A)=P(A|B)$
 (B) Yes, because $P(A)=P(B|A)$
 (C) Yes, because $P(A)=P(B)$
 (D) No, because $P(A)\neq P(A|B)$
 (E) No, because $P(A|B)\neq 0$
 Answer: _____
- (v) If we selected **two** random voters from the above 8417 voters, what is the probability that at least one voted for the Saskatchewan party?
 (A) 0.3561
 (B) 0.5853
 (C) 0.6439
 (D) 0.7122
 (E) 0.1268
 Answer: _____

The page is to be used as scrap paper only.

3. It is estimated that 80% of patients will recover from a stomach infection without the need for medical intervention.
- (i) Suppose ten people contract a stomach infection. What is the probability that exactly eight will recover without need for medical intervention?
(A) 0.8
(B) 0.64.
(C) 0.1678
(D) 0.00067
(E) 0.3020
Answer: _____
- (ii) Suppose ten people contract a stomach infection. What is the probability that at least one will recover without need for medical intervention?
(A) 0.9999
(B) 0.8926
(C) 0.6447
(D) 0.1600
(E) 0.1074
Answer: _____
- (iii) What is the mean and standard deviation of people recovering from a stomach infection in a random sample of 250 individuals?
(A) mean = 200 standard deviation = 2.519
(B) mean = 200 standard deviation = 6.325
(C) mean = 250 standard deviation = 2.519
(D) mean = 250 standard deviation = 4.040
(E) mean = 50 standard deviation = 2.519
Answer: _____
- (iv) Use the normal approximation to the binomial (with continuity correction) to estimate the probability that out of a random sample of 250 individuals, more than 210 will recover without medical intervention.
(A) 0.9515
(B) 0.9129
(C) 0.8742
(D) 0.1258
(E) 0.0485
Answer: _____
4. A multiple choice examination has five questions, each with five possible answers, only one of which is correct. A student didn't prepare for the exam and decides to make random and independent guesses for each of the five questions. The passing grade for the exam is 80%, i.e. the student must get at least four of the five questions correct in order to pass. What is the probability that the student passes this exam?
(A) 0.3277
(B) 0.1600
(C) 0.0400
(D) 0.0067
(E) 0.0031
Answer: _____

The page is to be used as scrap paper only.

5. The weight of Bengal Cats is normally distributed with mean 5.20 kg and standard deviation 0.78 kg.

(i) What is the probability that a randomly selected Bengal Cat weighs less than 5.00 kg?

- (A) 0.2564
- (B) 0.3974
- (C) 0.7436
- (D) 0.6026
- (E) 0.2013

Answer: _____

(ii) What is the probability that a randomly selected Bengal Cat weighs between 5.00 kg and 6.00 kg?

- (A) 0.2564
- (B) 0.4001
- (C) 0.8921
- (D) 0.4511
- (E) 0.3407

Answer: _____

(iii) How much would a Bengal Cat weigh if it was considered to be in the 90th percentile in weight?

- (A) 6.5 kg
- (B) 6.0 kg
- (C) 6.4 kg
- (D) 6.2 kg
- (E) 5.9 kg

Answer: _____

(iv) What is the standard error of the sample mean for Bengal Cat weight, given samples of size $n=60$?

- (A) 0.101
- (B) 0.013
- (C) 0.087
- (D) 0.671
- (E) 0.125

Answer: _____

(v) If we selected a sample of 60 Bengal Cats, what is the probability that their mean weight would be more than 5.25 kg?

- (A) 0.5
- (B) 0.6915
- (C) 0.1666
- (D) 0.3224
- (E) 0.3085

Answer: _____

The page is to be used as scrap paper only.

6. A psychology study was conducted to investigate the relationship between the amount of time spent meditating each day (in minutes) and reported levels of stress (on a scale of 1 to 10). The data collected from participants is as follows:

x: time spent	10	15	20	25	30	35	40	45
y: stress level	7	6	6	5	4	3	2	3

Note that $\sum x = 220$ $\sum x^2 = 7100$ $\sum y = 36$ and $\sum y^2 = 184$

- (i) Calculate the value of $\sum xy$.
- (A) 845
 (B) 1050
 (C) 22
 (D) 145
 (E) 7920
- Answer: _____
- (ii) Calculate the value of r , the correlation coefficient.
- (A) 0.95
 (B) -0.95
 (C) -0.15
 (D) 0.87
 (E) -0.87
- Answer: _____
- (iii) Based on the r -value, what best describes the relationship between the time spent meditating and stress level?
- (A) They have a strong and positive correlation.
 (B) They have a weak and positive correlation.
 (C) They have a weak and negative correlation.
 (D) They are not correlated at all.
 (E) They have a strong and negative correlation.
- Answer: _____
- (iv) Calculate the slope of the regression line.
- (A) 0.1381
 (B) -0.1381
 (C) 8.2978
 (D) -8.2978
 (E) 6.5900
- Answer: _____
- (v) If a participant meditates for 50 minutes per day, what do you predict their reported stress level would be (choose the closest value on a scale of one to ten)?
- (A) 4
 (B) 3
 (C) 2
 (D) 1
 (E) 5
- Answer: _____

The page is to be used as scrap paper only.

PART B - FULL ANSWER QUESTIONS

To receive full credit for correct answers it is necessary to show all work in the space provided.

7. Suppose we wish to estimate a population mean. If we want the sample estimate to be within 2.30 units of the true population mean, 9 out of 10 times, how large a sample should we select. We will assume that the population standard deviation is known to be 12.10. [4 marks]
8. To estimate the normal body temperature of a healthy human, a random sample of 130 healthy patients is selected. Their data yielded a sample mean of 36.79°C with standard deviation 0.38°C . Construct a 95% confidence interval for the true healthy human body temperature. [6 marks]

The page is to be used as scrap paper only.

9. A recent study on children's health claims that 17% of Canadian children are classified as obese. To test this claim, a sample of 100 randomly selected Canadian children is taken. In this sample, 8 children are classified as obese. Is this sufficient evidence to conclude that the claimed 17% rate is too high? Perform the appropriate hypothesis test, use $\alpha=0.05$. [10 marks]

The page is to be used as scrap paper only.

10. A biologist believes that the average number of bird species in a particular forest has decreased compared to historical averages, which have typically been 10 species per location. To test this hypothesis, the biologist conducts a study and records the number of bird species observed in 10 randomly selected locations within the forest. The data collected is as follows:

Number of bird species observed: 12, 10, 8, 9, 7, 8, 6, 7, 8, 5

Note: $\sum x = 80$ and $\sum x^2 = 676$

At a significance level of $\alpha=0.05$, can we conclude that the biologist is correct?

[10 marks]

The page is to be used as scrap paper only.

11. To compare the lifespan of two different brands of 6W LED light bulbs, a sample of 100 bulbs is tested from each brand. The resulting data is shown below:

	Brand A	Brand B
Sample Size	100	100
Mean Lifespan (hours)	49,532	50,887
Standard Deviation (hours)	2,352	1,983

Use the given data to construct a 90% confidence interval for the difference in lifespan (in hours) for these two brands of light bulb. [10 marks]

The page is to be used as scrap paper only.

12. An IT-Support centre has five different support technicians for phone-in questions. The centre manager wants to determine if there is any significant difference in the speed with which these five technicians handle support calls. To do so, Technicians #1, #3 and #4 were observed for four support calls each. Technicians #2 and #5 were observed for five support calls each. For each observed support call, the time (in minutes) was recorded in the table below. [10 marks]

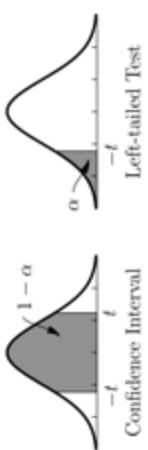
Technician #1	Technician #2	Technician #3	Technician #4	Technician #5
12.5	11.1	15.2	18.1	13.5
13.1	19.1	20.7	11.1	16.7
12.7	12.3	20.6	12.4	15.4
16.5	12.6	21.0	12.6	14.9
	12.9			15.2

- a) Complete the ANOVA table (fill in the three missing degrees of freedom, the two missing Mean Squares, and the F-Score).

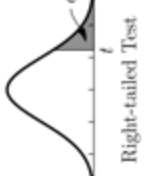
Source	d.f.	Sum of Squares	Mean Squares	F
Treatment		101.63		
Error		107.83		
Total		209.46		

- b) At the $\alpha=0.05$ level of significance, how would you respond to the claim that there is no difference in mean times for these five technicians? Clearly support your answer.

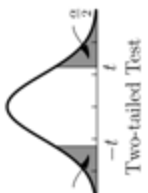
Student's *t* Distribution (Critical Values)



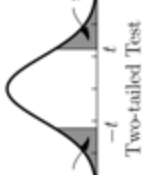
Confidence Interval



Right-tailed Test



Left-tailed Test



Two-tailed Test

df	Confidence Coefficient, 1 - α			
	0.80	0.90	0.95	0.99
	Level of Significance for One-Tailed Test, α			
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	5.841
4	1.533	2.132	2.776	4.604
5	1.476	2.015	2.571	4.032
6	1.440	1.943	2.447	3.707
7	1.415	1.895	2.365	3.499
8	1.397	1.860	2.306	3.355
9	1.383	1.833	2.262	3.250
10	1.372	1.812	2.228	3.169
11	1.363	1.796	2.201	3.106
12	1.356	1.782	2.179	3.055
13	1.350	1.771	2.160	3.012
14	1.345	1.761	2.145	2.977
15	1.341	1.753	2.131	2.947
16	1.337	1.746	2.120	2.921
17	1.333	1.740	2.110	2.898
18	1.330	1.734	2.101	2.878
19	1.328	1.729	2.093	2.861
20	1.325	1.725	2.086	2.845
21	1.323	1.721	2.080	2.831
22	1.321	1.717	2.074	2.819
23	1.319	1.714	2.069	2.807
24	1.318	1.711	2.064	2.797
25	1.316	1.708	2.060	2.787
26	1.315	1.706	2.056	2.779
27	1.314	1.703	2.052	2.771
28	1.313	1.701	2.048	2.763
29	1.311	1.699	2.045	2.756
30	1.310	1.697	2.042	2.750

df	Confidence Coefficient, 1 - α			
	0.80	0.90	0.95	0.99
	Level of Significance for One-Tailed Test, α			
1	1.282	1.645	1.960	2.326
2	1.282	1.645	1.960	2.326
3	1.282	1.645	1.960	2.326
4	1.282	1.645	1.960	2.326
5	1.282	1.645	1.960	2.326
6	1.282	1.645	1.960	2.326
7	1.282	1.645	1.960	2.326
8	1.282	1.645	1.960	2.326
9	1.282	1.645	1.960	2.326
10	1.282	1.645	1.960	2.326
11	1.282	1.645	1.960	2.326
12	1.282	1.645	1.960	2.326
13	1.282	1.645	1.960	2.326
14	1.282	1.645	1.960	2.326
15	1.282	1.645	1.960	2.326
16	1.282	1.645	1.960	2.326
17	1.282	1.645	1.960	2.326
18	1.282	1.645	1.960	2.326
19	1.282	1.645	1.960	2.326
20	1.282	1.645	1.960	2.326
21	1.282	1.645	1.960	2.326
22	1.282	1.645	1.960	2.326
23	1.282	1.645	1.960	2.326
24	1.282	1.645	1.960	2.326
25	1.282	1.645	1.960	2.326
26	1.282	1.645	1.960	2.326
27	1.282	1.645	1.960	2.326
28	1.282	1.645	1.960	2.326
29	1.282	1.645	1.960	2.326
30	1.282	1.645	1.960	2.326

The *F* Distribution (Critical Values)

df1	df2			
	2	3	4	5
2	19.000	17.000	16.000	15.520
3	10.133	9.013	8.451	8.182
4	7.709	6.591	6.134	5.946
5	6.591	5.473	5.016	4.830
6	5.965	4.847	4.390	4.203
7	5.509	4.391	3.962	3.776
8	5.146	3.935	3.534	3.476
9	4.844	3.479	3.106	3.176
10	4.588	3.023	2.678	2.876
11	4.352	2.567	2.250	2.576
12	4.137	2.111	1.822	2.276
13	3.942	1.655	1.394	1.976
14	3.765	1.200	0.966	1.676
15	3.604	0.744	0.538	1.376
16	3.456	0.288	0.110	1.076
17	3.320			0.776
18	3.194			0.476
19	3.077			0.176
20	2.968			0.076
21	2.866			0.026
22	2.770			0.010
23	2.680			0.005
24	2.595			0.003
25	2.515			0.002
26	2.439			0.001
27	2.365			0.001
28	2.293			0.001
29	2.222			0.001
30	2.152			0.001
40	1.638			0.000
50	1.362			0.000
60	1.211			0.000
70	1.114			0.000
80	1.048			0.000
90	1.000			0.000
∞	0.995			0.000