Final Exam Review

December 1, 2025 4:36 PM

THE COVER PAGE OF YOUR EXAM WILL BE THE FOLLOWING. NOTE THE LENGTH (120 MIN) AND THE INSTRUCTIONS.

THESE PROBLEMS ARE EXAM-STYLE. HOWEVER, THE ACTUAL EXAM WILL LIKELY BE SOMEWHAT LONGER.

British Columbia Institute of Technology

MATH 3042 – Final Exam

Program:	Computer Systems Technology	
Course Name:	Applied Probability and Statistics for CST	
Course Number:	Math 3042	
Date:	December 9, 2025	
Time Allotted:	120 min	
Exam Pages:	X	(including this page)
Total Marks:	Y	(40% weight for the course)

Instructions

1) Do not open the exam or write anything on these pages before you are told to begin.

2) You may use a scientific calculator with statistical functions. No other devices are allowed.

3) If your answer is a probability, round it to four digits after the decimal point. Otherwise, round to three significant digits.

4) A formula sheet is provided separately. No other notes or written materials are allowed.

5) No communication of any sort is allowed with other students or any other person besides your instructor or other exam invigilator.

6) All answers are to be written clearly in this examination booklet.

1

Q1. Correlation Test + Linear Regression

A chemical company performed an experiment to study the effect of extraction time on the efficiency of an extraction operation, and obtained the data in the table to the right.

The researchers also ran the following R commands (with output as shown)

X Y

Extraction Time (minutes)	Extraction Efficiency (%)
27	57
45	64
38	78
19	46
35	62
38	70
19	52
49	77
15	57
31	68

```
> time <- c(27, 45, 38, 19, 35, 38, 19, 49, 15, 31)
> efficiency <- c(57, 64, 78, 46, 62, 70, 52, 77, 57, 68)
> favstats(time)
 min Q1 median Q3 max mean    sd  n missing
  15 21     33 38  49 31.6 11.50 10       0
> favstats(efficiency)
 min Q1 median   Q3 max mean    sd  n missing
  46 57     63 69.5  78 63.1 10.43 10       0
> cor(time, efficiency)
[1] 0.8072386
```

X (favstats(time)), Y (favstats(efficiency)), r (cor output)

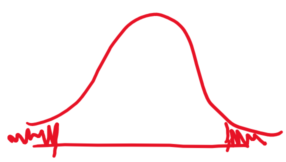
a. Is there a statistically significant linear correlation between extraction time and extraction efficiency? Perform an appropriate hypothesis test.

(1) Claim : pop. cor $\rho \neq 0$.

(2) Hypothesis $H_0: \rho = 0$
$H_1: \rho \neq 0$ [two tail]

(3) test statistic: $t = \dfrac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \dfrac{0.8072386}{\sqrt{\frac{1-0.8072386^2}{10-2}}} = 3.868$

(4) rejection region: reject if $t < -2.3060$
or $t > 2.3060$



(5) decision : Reject H_0.

(6) Conclusion: There is a non-zero correlation at the population level at the 5% significance level

⑥ Conclusion: There is a non-zero correlation at the population level at the 5% significance level.

b. Estimate the extraction efficiency when the extraction time is 25 min.

Use $x = 25$ min.

$\hat{y} = a + bx$

slope $b = r \cdot \frac{S_y}{S_x} = 0.80724 \times \frac{10.43}{11.50} = 0.7321$

y-int $a = \bar{y} - b \cdot \bar{x} = 63.1 - 0.7321 \times 31.6 = 39.96$

regression line: $\hat{y} = 39.96 + 0.7321X$

prediction: $\hat{y} = 39.96 + 0.7321 \times 25 = \boxed{58.3\ \%}$

Q2. Confidence Intervals for μ

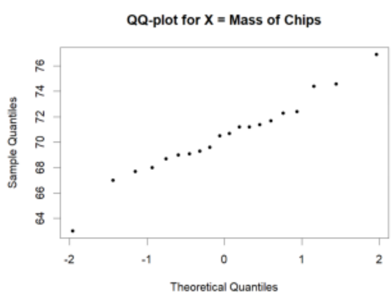
A potato chip product claims that a bag contains 66 g of chips. One interpretation of this is the claim that the *mean* mass of a bag of chips is at least 66 g.

To test this claim, you sample $n = 20$ bags of chips. The measurements are:

$X =$

72.4	76.9	70.7	63.0	68.7	71.4	67.0	71.2	74.6	67.7
72.3	74.4	70.5	69.6	68.0	69.1	71.2	69.3	69.0	71.7

A QQ-plot indicates that the distribution is normal.



a. Calculate the sample statistics:

$\bar{X} =$ 70.435 g
$s =$ 3.038 g
$n =$ 20

b. Determine the 95% confidence intervals for μ, the mean mass of a chip bag.

$1 - \alpha = 0.95$
$\alpha = 0.05$

$t_{\alpha/2} = 2.0930$ [t-table: $df = n - 1 = 19$, $\alpha = 0.05$]

$\alpha = 0.05$

$t_{\alpha/2} = 2.0930$ [t-table: $df = n-1 = 19$, $\alpha = 0.05$]

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 2.0930 \times \frac{3.0389}{\sqrt{20}} = 1.422 \text{ g}$$

Conf. interval: $\bar{X} - E = \boxed{69.0 \text{ g}}$

to $\bar{X} + E = \boxed{71.9 \text{ g}}$

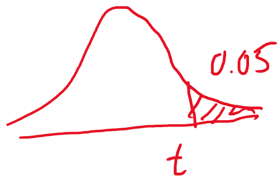
c. Test the hypothesis that the mean mass, μ, is greater than 66 g. Use $\alpha = 5\%$.

(1) Claim: $\mu > 66$ g

(2) Hypotheses: $H_0: \mu = 66$
$H_1: \mu > 66$ (right tail)

(3) Test statistic: $t = \dfrac{\bar{X} - \mu}{s/\sqrt{n}} = \dfrac{70.435 - 66}{3.038/\sqrt{20}} = 6.53$

(4) Rejection Region:



Critical $t = 1.7291$

[using $df = 19$, two-tail $\alpha = 0.1$]

(5) Decision: Reject H_0 since $6.53 > 1.7291$.

(6) Conclusion: There is sufficient evidence at the 5% significance level to support the claim that the mean mass of a bag is greater than 66 g.

$n = 50$
$x = 7$

$\hat{p} = \frac{x}{n} = \frac{7}{50} = 0.14$

Q3. Confidence Interval for p

A random sample of 50 BCIT students has 7 left-handed and 43 right-handed students.

a. Test the claim at the 5% significance level that 11% of BCIT students are left-handed.

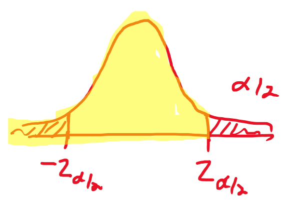
(1) Claim: Population proportion $p = 0.11$

(2) Hypotheses: $H_0: p = 0.11$
$H_1: p \neq 0.11$

(3) Test Statistic: $z = \dfrac{\hat{p} - p}{\sqrt{pq/n}} = \dfrac{0.14 - 0.11}{\sqrt{\frac{0.11 \times 0.89}{50}}} = 0.678$

(4) Rejection Region:

$z_{\alpha/2} = 1.96$ [z-table]



(5) Decision: Fail to Reject H_0

(6) Conclusion:

(5) Decision: Fail to Reject H_0

(6) Conclusion:

$P(Z < Z_{\alpha/2}) = 0.9750$

There is insufficient evidence to reject the claim that 11% of BCIT students are left-handed. [The claim is good.]

b. If your hypothesis test led to the wrong conclusion, what kind of error was it (Type 1 or Type 2)? How could you have reduced the probability of making such a mistake?

Type 2 [Fail to Reject False Null]

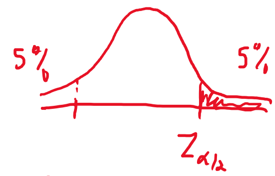
Can reduce by increasing α and/or increasing the sample size, n.

c. Determine the 90% confidence interval for p, the population proportion of left-handed students at BCIT.

$1-\alpha = 0.90$

$\alpha = 0.1$

$Z_{\alpha/2} = 1.645$ [1.64 or 1.65 is OK]



$E = Z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}\hat{q}}{n}} = 1.645 \times \sqrt{\dfrac{0.14 \times 0.86}{50}} = 0.0807$

Conf int: $\hat{p} - E = 0.059$

to $\hat{p} + E = 0.221$

The proportion of left-handed students is between 0.059 (5.9%) and 0.221 (22.1%) with 90% confidence.

4

Q4. Central Limit Theorem

Let X = the volume of liquid in a bottle of Gatorade. Assume that X has:

- mean: $\mu = 595$ mL
- std dev: $\sigma = 3$ mL

a. What is the probability that a bottle of Gatorade contains more than 600 mL?

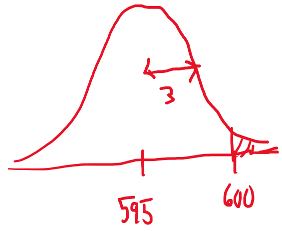
- std dev: $\sigma = 3$ mL

a. What is the probability that a bottle of Gatorade contains more than 600 mL?

i. Assuming X is normal.

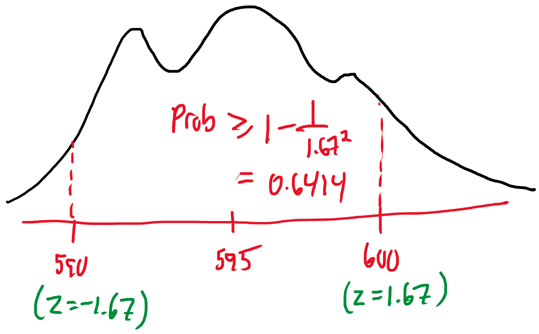
$$P(X>600) = P(Z>1.67)$$
$$= 1-0.9525$$
$$= \boxed{0.0475}$$

$$Z = \frac{X-\mu}{\sigma} = \frac{600-595}{3} = 1.67$$



ii. Without assuming X is normal. *Hint: Chebyshev.*

Chebyshev says



Therefore

$$P(X>600) \leq 1-0.6414 = 0.3586$$

[The probability is at most 0.3586.]

b. If you pour 100 bottles of Gatorade into a larger container, what is the probability that the resulting volume is greater than 59.6 L?

$n = 100$

$\mu_{\bar{X}} = \mu = 595$

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{100}} = 0.3$

$\bar{X}$ is normal since $n>30$

If 100 bottles add up to over 60L, then the average is $\bar{X} > \frac{60 \times 1000\text{ mL}}{100} = 600$ mL

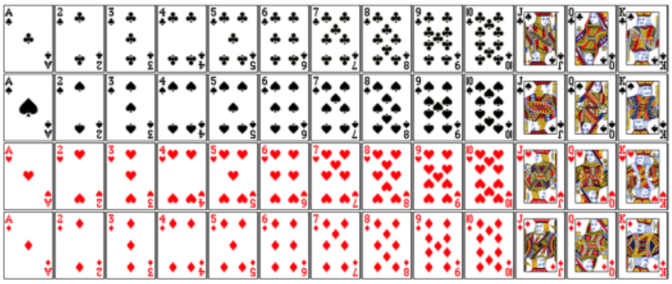
$$P(\bar{X}>600) = P(Z>16.7) = 0^+$$

[beyond last value in z-table]

$$Z = \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{600-595}{0.3}$$
$$= 16.7$$

There is practically no probability that the total is greater than 60L.

Q5. General Probability



$N = 52$

a. Suppose you select four cards from a standard deck. What is the probability that all four cards have a different *value* (e.g., A, 2, 3, 4 have different value)?

$n = 4$ (no replacement)

$$P(\text{all have different value}) = \frac{52}{52} \times \frac{48}{51} \times \frac{44}{50} \times \frac{40}{49}$$

$$= \boxed{0.6761}$$

(This question is redundant, sorry.)

b. Suppose you select *six cards*. What is the probability that all six cards have different value?

$$P(\text{all six different}) = \frac{52 \times 48 \times 44 \times 40 \times 36 \times 32}{52 \times 51 \times 50 \times 49 \times 48 \times 47}$$

$$= \boxed{0.3452}$$

c) Select $n = 6$.
what is prob. that half are "Face Card" (J, Q, K) given that the first card is a king?

$P(\text{Half Face} \mid 1^{st} \text{ king})$

$$P(\text{Half Face} \mid 1^{st} \text{ King})$$

$$= \frac{C(11,2) \times C(40,3)}{C(51,5)}$$

$$= 0.2313$$

6

Q6. Discrete – Geometric, Binomial, Poisson

The number of flaws in cables produced by a certain manufacturer follows a Poisson distribution with mean 0.2 flaws per meter. Cables are 4m long and are considered to be acceptable if they contain at most one flaw.

Acceptable means : $X \le 1$

a. What is the probability that a randomly-selected cable is not acceptable?

$$\lambda = 0.2 \frac{\text{flaw}}{m} \times 4m$$

$$= 0.8 \text{ flaws}$$

$$P(\text{not Acceptable}) = 1 - P(X \le 1)$$

$$= 1 - P(X=0) - P(X=1)$$

$$= 1 - e^{-0.8} \cdot \frac{0.8^0}{0!} - e^{-0.8} \times \frac{0.8^1}{1!}$$

$$= \boxed{0.1912}$$

b. If the manufacturer ships 10 cables to a customer, what is the probability that more than 25% of the cables are not acceptable?

$$P(3 \text{ or more are unacceptable})$$

$$= 1 - (P(0) + P(1) + P(2))$$

$$= 1 - (10C0) \times (0.1912)^0 (0.8088)^{10} - (10C1) \times 0.1912 \times 0.8088^9$$

$$- (10C2) \times 0.1912^2 \times 0.8088^8$$

$$= \boxed{0.2958}$$

7

Q7. Continuous – Exponential Distribution, Normal Distribution

The average time between global pandemics is 15 yrs. Suppose that the time X until the next global pandemic can be modelled as an *exponential* random variable.

$\beta = 15$

Let's say COVID-19 started in December 2019.

Dec 2019 → Dec 2024 is 5 years.

a. In Dec 2019, what was the probability that the next global pandemic would start before December 2024?

$$P(X < 5) = F(5)$$
$$= 1 - e^{-5/15}$$
$$= \boxed{0.2835}$$

b. There has been no new global pandemic (as of Dec 2024) since COVID-19. What is the probability that the next global pandemic will occur before Dec 2029?

From now until Dec 2029 is 5 more years.

So the answer is $\boxed{0.2835}$

Isn't Dec 2019 → Dec 2029 ten years?

$1 - e^{-10/15} = 0.4866$?

Isn't Dec 2019 → Dec 2029 ten years!

So why isn't the answer $1 - e^{-10/15} = 0.4866$?

Because we are asking: $P(X<10 \mid X \geq 5)$

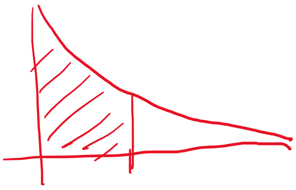
$$= \frac{P(5 \leq X < 10)}{P(X \geq 5)} = \frac{(1-e^{-10/15}) - (1-e^{-5/15})}{1-(1-e^{-5/15})}$$

$$= \frac{e^{-5/15} - e^{-10/15}}{e^{-5/15}}$$

$$= 1 - e^{-5/15} = \boxed{0.2835}$$

Exponential Variables are "memoryless".

c. Find the median value of X.



P_{50}

$$F(x) = 1 - e^{-x/\beta} = 0.50$$

$$e^{-x/\beta} = 0.5$$

$$-x/\beta = \ln(0.5)$$

$$P_{50} = x = -\beta \cdot \ln(0.5) = 15\,\text{yr} \times \ln 2$$

$$= \boxed{10.4\ \text{yr}}$$

Q8. Conditional Probability

Suppose that 27% of Canadians own a cat, 36% of Canadians own a dog, and that 21% of Canadians that own a cat also own a dog.

a. What is the probability that a randomly chosen Canadian dog owner also owns a cat?

$$P(\text{cat}) = 0.27$$

$$P(\text{dog}) = 0.36$$

$$P(\text{dog} \mid \text{cat}) = 0.21$$

$$P(\text{cat} \mid \text{dog}) = \frac{P(\text{cat} \cap \text{dog})}{P(\text{dog})} = \frac{P(\text{cat}) \cdot P(\text{dog} \mid \text{cat})}{P(\text{dog})}$$

$$= \frac{0.27 \times 0.21}{0.36}$$

$$= \frac{0.27 \times 0.21}{0.36}$$

$$= \boxed{0.1575}$$

b. In addition, suppose that 15% of Canadians that own both a dog and a cat also own a bird. What is the probability that a randomly chosen Canadian owns all three of these types of pet (cat, dog, bird)?

$$P(\text{Bird} \mid \text{Dog and Cat}) = 0.15$$

$$P(\text{Dog and Cat and Bird})$$

$$= P(\text{Dog and Cat}) \times P(\text{Bird} \mid \text{Dog and Cat})$$

$$= P(\text{Cat}) \times P(\text{Dog} \mid \text{Cat}) \times P(\text{Bird} \mid \text{Dog and Cat})$$

$$= 0.27 \times 0.21 \times 0.15$$

$$= \boxed{0.0085}$$