ENGI 4421 Final Examination

2021 Spring

1. The random quantity X follows a normal distribution with population mean $\mu = 140.0$ and population standard deviation $\sigma = 5.6$.

(a) Evaluate $\mathrm{P}[135 \le X < 145]$, correct to 2 significant figures. [4]

(b) A random sample of size $n = 4$ is taken from this population. [6]
Evaluate $\mathrm{P}[135 \le \bar{X} < 145]$, correct to 2 significant figures.

---

2. The cumulative distribution function for a **continuous** random quantity X is given by

$$F(x) = \begin{cases} 0 & (x < 1) \\ (x-1)^2(5-2x) & (1 \le x \le 2) \\ 1 & (x > 2) \end{cases}$$

(a) Find $\mathrm{P}[X \le 1.5]$ [3]

(b) Find the probability density function $f(x)$ for X. [4]

[You may quote $(x-1)^2(5-2x) = -2x^3 + 9x^2 - 12x + 5$]

(c) Find the population variance $\mathrm{V}[X] = \sigma^2$ [6]

(d) Find the median $\tilde{\mu}$ [2]

---

3. A pair of random quantities X, Y has the joint probability mass function $p(x, y)$ given by the table below.

$p(x, y)$		$y = -1$	$y = 0$	$y = 1$	
x	-1	.0400	.0400	.0200	
	0	.1400	.1225	.0875	
	1	.1000	.0875	.0625	
	2	.1200	.1000	.0800	

(a) Complete the table to display the marginal probability functions $p_X(x)$ and $p_Y(y)$. [2]

(b) Find $\text{P}[\,X < 0 \mid Y < 0\,]$ (the probability that X is negative given that Y is negative). [3]

(c) Find the covariance of X and Y, Cov[X, Y]. [5]

(d) Are the random quantities X, Y independent? Why or why not? [3]

---

4. Is the following set of 200 observations of X consistent (at a 5% level of significance) with X having been drawn from a uniform distribution? [14]
(o_i is the number of observations for each value of x)

x_i:	1	2	3	4	5	6	7	8
o_i:	24	20	33	24	34	27	13	25

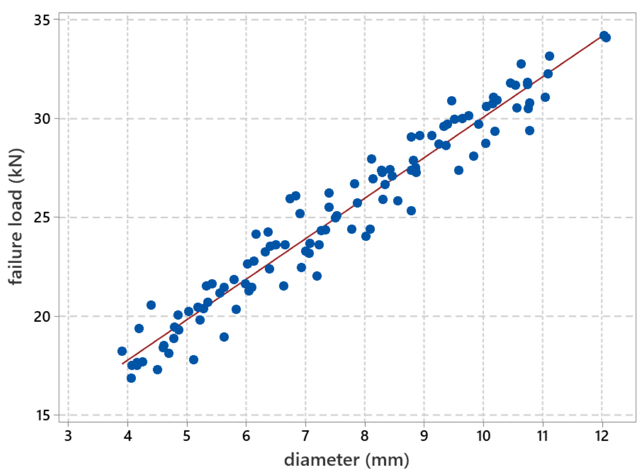
---

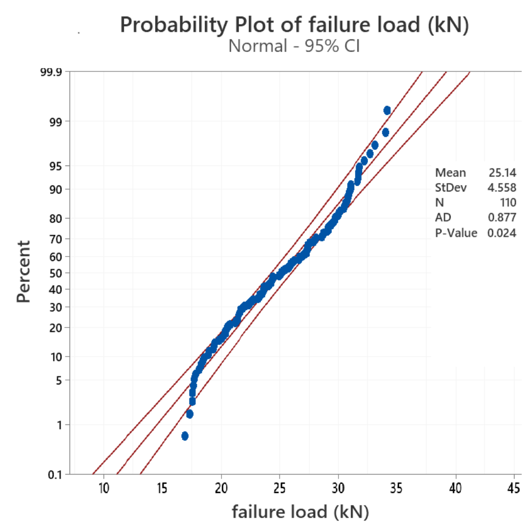
5. The average volumetric flow Q of water through a test pipe is known to be a random quantity that follows a Normal distribution. It is believed that the population mean flow is 1000 L/s and the strength of that belief is represented by the standard deviation $\sigma_o = 5\text{ L/s}$. After a maintenance cycle, a random sample of 81 new measurements of water flow through the pipe has a mean flow rate of 1005.5 L/s with a sample standard deviation of 22.5 L/s.

(a) Construct a classical 95% confidence interval estimate for the true mean flow μ. [5]
(b) Construct a Bayesian 95% confidence interval estimate for the true mean flow μ. [5]
(c) Is there sufficient evidence to conclude that the true mean flow μ is no longer 1000 L/s? [2]
(d) Provide a *brief* reason for the different widths of your two confidence intervals. [2]

---

6. For 110 observations of cable sheath thickness (x), the failure load (y) of each cable is recorded. A scatterplot and a normal probability plot of the observations are presented here:





(a) State ***two*** reasons why it is appropriate to use the simple linear regression model. [2]
(b) Given the summary statistics [3]

$$n = 110 \qquad \sum x = 835.0 \qquad \sum y = 2765.5$$
$$\sum x^2 = 6851.4 \qquad \sum xy = 22041.7 \qquad \sum y^2 = 71793.7$$
$$nS_{xx} = 56429.0 \qquad nS_{xy} = 115394.5 \qquad nS_{yy} = 249316.8$$

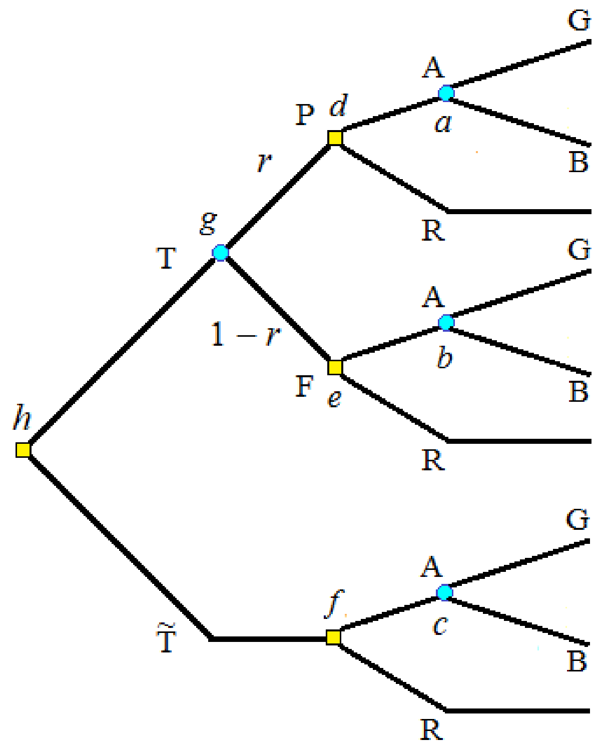
find the equation of the regression line $y = \hat{\beta}_0 + \hat{\beta}_1 x$.

(c) Construct the ANOVA table from these data. [4]
(d) To the nearest one per cent, what proportion of the variation in Y is explained by the linear regression? [2]
(e) Conduct an hypothesis test to determine whether or not there is a useful linear relationship between Y and x. [3]
(f) Find the 95% prediction interval for a future observation of Y at $x = 10.0$. [5]

7. A contractor knows that 70% of all boxes of pressure valves are "good" (that is, contain no defective valves; event 'G'), while the other 30% of boxes are "bad" (that is, contain at least one defective valve; event 'B').

A test is available, costing \$10 per box, to determine whether a box contains any defective valves. However, the test is not perfect. A "good" box has a probability of 0.75 to pass the test. A "bad" box has a probability of 0.05 to pass the test.

Accepting a good box earns \$100.
Accepting a bad box causes a loss of \$150.
Rejecting a box causes a loss of \$25 (regardless of whether it is good or bad).
The situation is illustrated in this partially completed decision tree diagram.



'A' represents a decision to accept a box,
'R' represents a decision to reject a box,
'T' represents a decision to invest in the test,
'P' represents the event passing the test and
'F' represents the event failing the test.

(a) Show that the probability that a box is good, given that it passed the test, is $\frac{35}{36}$. [3]

(b) Complete the decision tree by finding the expected values *a*, *b*, ..., *h*, determine the optimum strategy that maximizes the expected gain per box for the contractor and find the exact value of that maximum expected gain. [10]

(c) By how much would the test cost have to change, to alter the optimum strategy? [2]

8. *Bonus Question* [+5]

Poiseuille's equation expresses the pressure drop p in an incompressible and Newtonian fluid in laminar flow flowing through a long cylindrical pipe of constant cross section:

$$p = \frac{8\eta QL}{\pi r^4}$$

where η is the dynamic viscosity, Q is the volumetric flow rate through the pipe, L is the length of the pipe and r is the radius of the circular cross section of the pipe.

Given the measurements

$\eta = (1.002 \pm 0.005) \times 10^{-3}\,\text{kg m}^{-1}\text{s}^{-1}$

$Q = (5.103 \pm 0.250)\,\text{m}^3\text{s}^{-1}$

$L = (2000.00 \pm 0.05)\,\text{m}$

$r = (0.285 \pm 0.001)\,\text{m}$

and assuming independence between all of these measurements,
find the uncertainty in the pressure drop p.

☜ return to the index of questions on to the solutions ☞