PORTFOLIO 4

KEVIN DUONG N9934731

PART A - Week 10

1) Student Sam has access to two types of internet connection: a 4G phone connection, and a WiFi connection. The WiFi has unlimited data but is very unstable. The 4G connection is very stable but has limited data available. It is Friday night and Sam must submit an MZB126 assignment, there is a 90% chance that Sam will fail to submit their assignment if the WiFi isn't working and a 95% chance that Sam will fail to submit their assignment if the 4G service is out of data. Given that the probability of the WiFi not working 30% and the probability his 4G service is out of data is 25%:

a) What is the probability that Sam will have at least one way to submit their assignment, i.e. either the WiFi or the 4G connection will be working?

1 let A = Wifi not working

B = 4G data not working

C = Sam failing to submit assignment

$P(C \mid A) = 0.90 = 90\%$

$P(C \mid B) = 0.95 = 95\%$

$P(A) = 0.30 = 30\%$

$P(B) = 0.25 = 25\%$

a) $P(\text{one working internet}) = 1 - P(A \cap B)$

$= 1 - 0.3 \times 0.25$

$= 0.925$

$= 0.93$

$\boxed{P(\text{@ least 1 working internet}) = 93\%}$

b) If Sam failed to submit an assignment, what is the probability that the WiFi connection was not working, given the probability of Sam not submitting an assignment is 0.38?

Given $P(c) = 0.38$,

P(Sam failing due to Wifi not working)

$$= P(A \mid c)$$

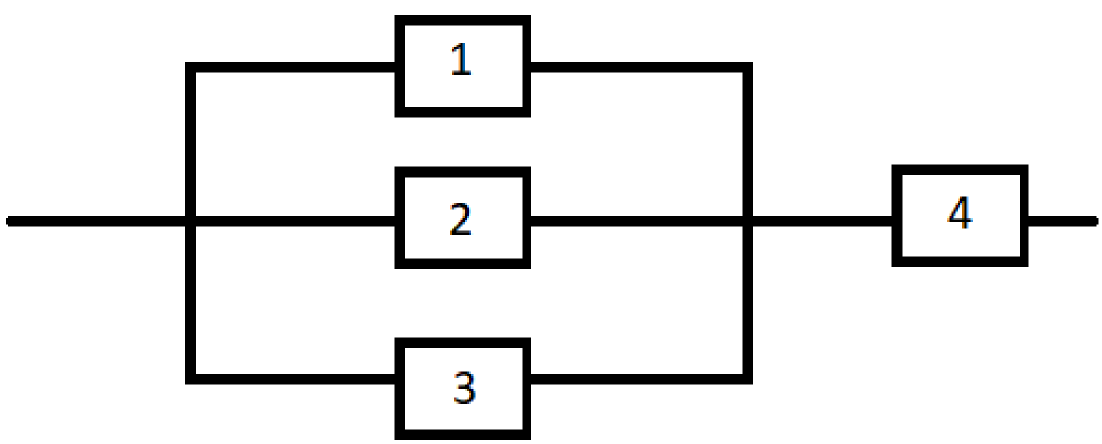
$$= \frac{P(A \cap c)}{P(c)}$$

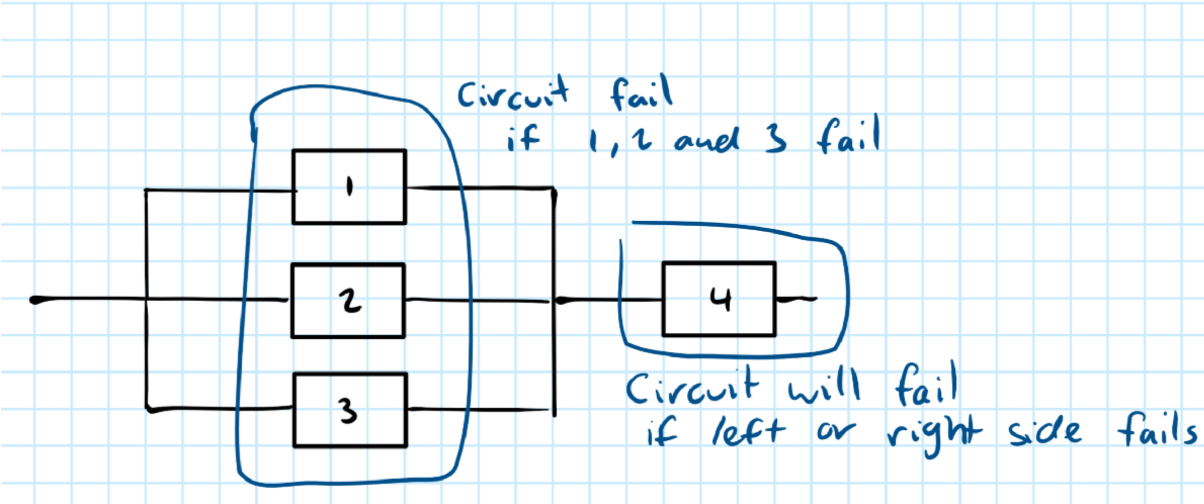
$$= \frac{P(c \mid A) \times P(A)}{0.38}$$

$$= \frac{0.9 \times 0.3}{0.38} = 0.7105$$

$\therefore$ P(Wifi not working and Sam's failed submission) = 71 %

2) Calculate the probability that the following circuit fails to function if the probability that a component works is and each component working is independent of the others.





@ each component, failure probability = $1 - p$

$\rightarrow$ 1, 2, 3 fails if $(1-p)^3$. Left side pass @ $p = 1-(1-p)^3$

$\rightarrow$ P (Circuit fails) = P(Left fails) $\times$ P (Right passes)
+ P (Left passes) $\times$ P (Right fails)

$$\boxed{\therefore \; P(\text{circuit fails}) = (1-p)^3 \times p + (1-(1-p)^3) \times (1-p)}$$

3) A mechanical component's life expectancy is modelled by the cdf $f(t) = \dfrac{4}{t^5}$ for $t \geq 1$ and $f(t) = 0$ for $t < 1$.

a) What is the probability the machine will last at least 2 years before failing.

$$P(t \geq 2) = 1 - P(t < 2)$$

$$= 1 - \int_1^2 f(t)\,dt$$

$$= 1 - \left(\int_1^2 \frac{4}{t^5}\,dt\right)$$

$$4\int \frac{1}{t^5}\,dt = 4\int t^{-5}\,dt$$

$$= 4 \times \frac{t^{-5+1}}{-5+1} = 4 \times \frac{t^{-4}}{-4}$$

$$= \left[-\frac{1}{t^4}\right]_1^2 = -\frac{1}{16} - -1 = \frac{15}{16}$$

$$= 1 - \frac{15}{16} = \frac{1}{16} = 0.0625 \approx 0.06$$

$\therefore$ P (Machine lasting at least 2 years before failing) = 6 %

b) calculate the probability that a component's time to failure is between 1.2 and 1.8 years.

$$P(1.2 \leq t \leq 1.8) = \int_{1.2}^{1.8} f(t)\,dt$$

$$= \int_{1.2}^{1.8} \frac{4}{t^5}\,dt$$

$$= \left[-\frac{1}{t^4}\right]_{1.2}^{1.8}$$

$$= -\frac{625}{6561} - -\frac{625}{1296}$$

$$= 0.3870 \approx 0.39$$

$\therefore$ P(Component failing between 1.2 and 1.8 years) = 39 %

c) Calculate the median life expectancy of the mechanical component. (The median life expectancy is when there is a 50% the component is still working and 50% chance it has failed).

c) let median $= M$,

$P(1 \le t \le M) = 0.5$

$= \int_1^M f(t)\, dt$

$= \int_1^M \frac{4}{t^5}\, dt$

$\left[-\frac{1}{t^4}\right]_1^M = 0.5$

Solve for M

$-\frac{1}{M^4} - (-1) = 0.5$

$1 - \frac{1}{M^4} = 0.5$

$\Rightarrow \frac{1}{M^4} = \frac{1}{2}$

$\Rightarrow \sqrt{M^4} = \sqrt{2}$

$\sqrt{M^2} = \sqrt{\sqrt{2}}$

$\boxed{\therefore \text{ Median } M = 1.189 \text{ years}}$

d) Calculate the mean life expectancy and the variance of the life expentancy.

d) Mean life :

$$E(t) = \int_1^\infty t \times f(t)\, dt$$

$$= \int_1^\infty t \times \frac{4}{t^5}\, dt$$

$$= 4 \int t \times \frac{1}{t^5}\, dt$$

$$= 4 \int \frac{1}{t^4}\, dt = 4 \int t^{-4}\, dt$$

$$= 4 \times \frac{t^{-4+1}}{-4+1}$$

$$= \left[\frac{-4}{3t^3} \right]_1^\infty$$

$$= -\frac{4}{3} \left[0 - 1 \right] = \frac{4}{3} \text{ years}$$

$$\boxed{\text{Mean} = E(t) = \frac{4}{3}}$$

Variance :

$$E(t^2) = \int_1^\infty t^2 \times f(t)\,dt$$

$$= \int_1^\infty t^2 \times \frac{4}{t^5}\,dt$$

$$= 4\int t^{-3}\,dt$$

$$= 4 \times \frac{t^{-3+1}}{-3+1}$$

$$= \left[-\frac{2}{t^2}\right]_0^\infty$$

$$= -2\left[0 - 1\right] = 2$$

$$\boxed{E(t^2) = 2}$$

$$\therefore \text{Variance} = E(t^2) - (E(t))^2$$

$$= 2 - \left(\frac{4}{3}\right)^2$$

$$= \frac{2}{9} \text{ years}$$

$$\boxed{\text{Mean} = \frac{4}{3} \text{ years}, \quad \text{Variance} = \frac{2}{9} \text{ years}}$$

PART B - Week 11

After acing their MZB126 exams, Sam scored an internship, helping a company write a bid to expand the shoulder of the Dalby highway (a 300km section of road, stretching from Dalby to st George). Sam is to calculate several probabilities below. For each task, Sam must:

i) identify the distribution,

ii) calculate the parameters of the distribution,

iii) calculate the probability by hand,

iv) and then confirm the probability in MATLAB.

1) As part of expanding the highway, several mail boxes must be relocated. The local council has announced that one mail box can be expected every 25km. What is the probability that Sam observes 15 mail boxes while out counting them?

Make λ the expected no. of mail boxes in the 300 km section of road.

So $\lambda = \frac{300}{25} = 12$

- 300 km section of road
- 25 km for every mailbox seen

let x be the number of mailboxes in a 300 km section of road.

$x \sim P(12)$, $P(X = x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$, where $x = 0, 1, 2, 3 \ldots$

where $x = 15$,

$$P(X = 15) = \frac{e^{-12} \times 12^{15}}{15!} = 0.072391$$

$\boxed{\text{Probability that Sam observes 15 mailboxes is } 7.2\%}$

```
%Poisson Distribution
a = 15;
lambda = 12;

Probability = poisspdf(15,12)
```

```
Probability = 0.0724
```

```
lowerP = poisscdf(a,lambda)
```

```
lowerP = 0.8444
```

```
upperQ = poisscdf(a,lambda, 'upper')
```

```
upperQ = 0.1556
```

2) Fortunately, not all mail boxes have to be relocated. There is a 20% chance that the mail box is already far enough back from the road. To get a maximum possible cost, what is the probability all 15 mail boxes have to be relocated?

let x be no. of relocated mailboxes:

Given that there is 20% of not relocating.

Probability of relocating is $P = 1 - 0.80 = 0.20$

$\therefore P = 0.80$

$X \sim$ Binomial given $n = 15$, $P = 0.80$

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ where } x = 0, 1, 2, 3 \ldots$$

$$P(X = 15) = \binom{15}{15} (0.80)^{15} (0.20)^{15-15}$$

$$= \frac{15!}{(15-15)!\,15!} (0.80)^{15} (0.20)^0$$

$$= 0.035184$$

$\boxed{\text{Probability that 15 mail boxes have to be relocated is } 3.5\%}$

```
%Binomial Distribution
a = 15;
N = 15;
p = 0.8;

Probability = binopdf(a,N,p)
```

```
Probability = 0.0352
```

3) Data on the cost of relocating letter boxes on a previous job are listed in the data file attached. Given that the costs follow a normal distribution, calculate the probability that the cost/letter box on the current job will exceed $100/letterbox.

Normal Distribution

Standard Deviation

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

$\mu = 94.41051$ (mean)

$N = 30$ (No. of costs)

$$\sum_{i=1}^{30}(x_i - 94.41)^2 = 1887.17$$

$$\sigma^2 = \frac{1}{30} \times 1887.17 = 62.9057 \text{ (Variance } \sigma^2\text{)}$$

$$\sigma = \sqrt{62.91} = 7.93$$

$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{100 - 94.41}{7.93}$$

$$z = 0.7049$$

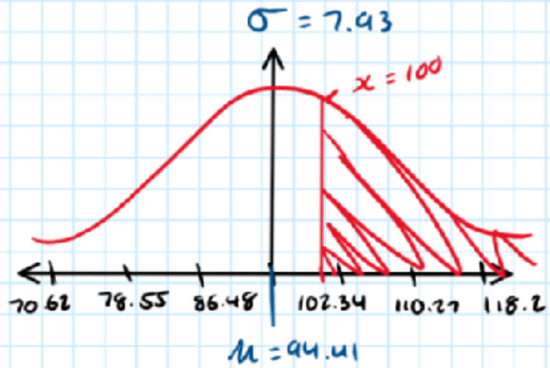
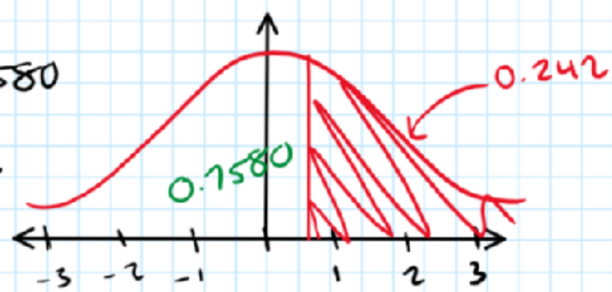


Table value = 0.7580

$\therefore\ 1 - 0.7580 = 0.242$



$$\boxed{\text{Probability that cost / letterbox on current job exceeding \$100 is 24\%}}$$

```
%Gaussian Distribution
s = 8.0669003971473;
b = 100;
mu = 94.410512263667;
sigma = 7.9313125826422;

Probability = 1 - normcdf(b,mu,sigma)-normcdf(s,mu,sigma)
```

```
Probability = 0.2405
```

PART C - Week 12

1) Write a .m file that accepts a vector of data, and calculates the sample mean and sample standard deviation.

```
% check n9934731_functions.m
```

2) Write a .m file that accepts a vector of data, and calculates a 95% CI of the mean of the data.

```
% check n9934731_functions.m
```

3) The data on the amount of whats produced at a small hydroplant are recorded in the csv file "hydro.csv" on blackboard.

a) A local community is concerned that not enough power is being produced from the plant. If the town requires 18 Megawatts to maintain its power supply, perform a hypothesis to determine the test statistic for $H_0 : \mu = 18 \text{ Megawatts}$.Write a concluding statement regarding whether or not there is evidence to reject the hypothesis that the powerplant produces enough power.

```
%Hypothesis Testing
N = 109;
SampleMean = 18.9961;
s = 2.0897;
u0 = 18;

% H0: u = 18
% HA: u does not equal 18

Ttest = (SampleMean - u0)/(s/sqrt(N))
```

```
Ttest = 4.9766
```

```
t = tinv(0.975,N-1)
```

```
t = 1.9822
```

Ttest > t. Therefore, there is strong evidence to reject the null hypothesis.

b) Compute a 95% confidence interval (either by hand or using your .m) of the average megawatts produced by the dam. State whether your confidence interval verifies the result of the hypothesis test from Q3a) . Give reasoning as to how you came up with this statement.

```
% Confidence Interval
CI = [SampleMean - t*s/sqrt(N),SampleMean + t*s/sqrt(N)]
```

```
CI =

   18.5994   19.3928
```

```
% .m file
CI_mfile = [18.5994, 19.3929]
```

```
CI_mfile =

   18.5994   19.3929
```

According to the Confidence Interval, the average megawatts produced by the dam provided by the hydro data gives a 95% confidence interval of 18.5994 and 19.3929 of the lower and upper bounds respectively. The u0 however is outside of the confidence interval, therefore verifying the hypothesis test from the previous question. This means there is now enough evidence to reject the null hypothesis of the powerplant having a mean of 18 Megawatts, suggesting that the alternative hypothesis is true and the mean does not equal to 18 Megawatts.