Ve401 Probabilistic Methods in Engineering

Suggested Solutions to the Sample Midterm Exam Problems

The following exercises have been compiled from past first midterm exams of Ve401. A first midterm will usually consist of 4-7 such exercises to be completed in 100 minutes. In the actual exam, necessary tables of values of distributions will be provided. You may use all tables in Appendix A of the textbook to solve the sample exercises.

Exercise 1.
A company produces widgets in three factories, A, B and C. Factory A produces 20% of the widgets, factory B produces 45% of the widgets and factory C produces the remaining 35%. Of all widgets produced, 5% fail tolerance. Of those that fail tolerance, 25% were produced in factory A, 35% were produced in factory B and 40% were produced in factory C. In each factory, what percentage of the widgets produced fails tolerance?
(3 Marks)

Solution. We have

$$P[A] = 0.2, \qquad P[B] = 0.45, \qquad P[C] = 0.35, \qquad P[\text{fail}] = 0.05$$
$$P[A \mid \text{fail}] = 0.25, \qquad P[B \mid \text{fail}] = 0.35, \qquad P[C \mid \text{fail}] = 0.40.$$

(1 Mark) Hence,

$$P[\text{fail} \mid A] = \frac{P[A \mid \text{fail}] \cdot P[\text{fail}]}{P[A]} = 0.0625$$
$$P[\text{fail} \mid B] = \frac{P[B \mid \text{fail}] \cdot P[\text{fail}]}{P[B]} = 0.0039$$
$$P[\text{fail} \mid C] = \frac{P[C \mid \text{fail}] \cdot P[\text{fail}]}{P[C]} = 0.0057$$

(2 Marks)

Exercise 2.
A coin is tossed 28800 times. Assuming that the coin is fair (i.e., heads and tails each have a 50% chance of coming up), use the normal approximation to the binomial distribution to give the probability that between 14380 and 14399 heads come up.
(4 Marks)

Solution. We want to approximate

$$\mathcal{P} = \sum_{k=14380}^{14399} \binom{28800}{k} 0.5^k 0.5^{28800-k}.$$

We use the normal approximation, so that

$$\mu = np = 28800 \cdot 0.5 = 14400$$

and

$$\sigma^2 = np(1-p) = 0.5 \cdot 28800 \cdot (1 - 0.5) = 7200.$$

(1 Mark) Note that $\sigma = 60\sqrt{2}$. Taking into account the half-unit correction, **(1 Mark)** we have ($Z = (X - 14400)/(60\sqrt{2})$)

$$\begin{aligned}
\mathcal{P} &\approx P[14379.5 \le X \le 14399.5] \\
&= P\left[\frac{14379.5 - 14400}{60\sqrt{2}} \le Z \le \frac{14399.5 - 14400}{60\sqrt{2}}\right] \\
&= P\left[-0.24 \le Z \le -0.0059\right] \\
&= P\left[0 \le Z \le 0.24\right] - P\left[0 \le Z \le 0.0059\right] \\
&= 0.0948 - 0.0000 = 0.0948 = 9.48\%.
\end{aligned}$$

(2 Marks)

Exercise 3.
Consider an experiment with three possible outcomes P, Q, R. The probability of outcome P is $p \in (0,1)$, that of Q is $q \in (0, 1-p)$ and that of R is $r = 1-p-q$. We perform $n \in \mathbb{N} \setminus \{0\}$ independent and identical experments of this type. Let X be the random variable denoting the number of outcomes P and Y the random variable denoting the number of outcomes Q.

i) What is the probability density function f_{XY} of the joint random variable (X,Y)? Explain how you obtain f_{XY} and verify that it is indeed a density function.

ii) From the definition of a marginal density function, calculate $f_X(x)$. Could you have obtained this result more easily?

iii) Are X and Y independent?

iv) Calculate $\mathrm{E}[X]$.

v) In quality control, a widget is either approved, rejected or marked for further testing. Of all widgets, 70% are accepted, 20% are rejected and 10% are marked for further testing. 10 widgets are tested and of those, 5 are approved. What is the probability that of the remaining 5, 3 are rejected?

(3+(2+1)+1+1+2 Marks)

Solution. i) The probability of obtaining any single arrangement with $X = x$ and $Y = y$ is given by $p^x q^y (1-p-q)^{n-x-y}$. **(1/2 Mark)** There are $\frac{n!}{x!y!(n-x-y)!}$ such arrangements of indistinguishable objects, **(1/2 Mark)** so

$$f_{XY}(x,y) = P[X = x \text{ and } Y = y] = \frac{n!}{x!y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y}$$

(1 Mark) By the multinomial theorem,

$$\sum_{x,y} f_{XY}(x,y) = \sum_{x+y+z=n} \frac{n!}{x!y!z!} p^x q^y (1-p-q)^z = (p+q+1-p-q)^n = 1.$$

(1/2 Mark) Trivially, $f_{XY} \geq 0$, so the two properties of a density function are satisfied. **(1/2 Mark)**

ii) The marginal distribution is given by

$$\begin{aligned} f_X(x) = \sum_y f_{XY}(x,y) &= \frac{n!p^x}{x!} \sum_{y=0}^{n} \frac{1}{y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y} \\ &= \frac{n!}{x!(n-x)!} p^x \sum_{y=0}^{n} \frac{(n-x)!}{y!(n-x-y)!} q^y (1-p-q)^{n-x-y} \\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}. \end{aligned}$$

(2 Marks) Alternatively, we could have considered P a "success' and R and Q "failures," thereby transforming the experiment into a Bernoulli experiment and used the binomial distribution to obtain

$$P[X = x] = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}.$$

(1 Mark)

iii) Since

$$\begin{aligned} f_X(x) f_Y(y) &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \frac{n!}{y!(n-y)!} q^y (1-q)^{n-y} \\ &\neq \frac{n!}{x!y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y} = f_{XY}(x,y), \end{aligned}$$

X and Y are not independent. **(1 Mark)**

iv) Using the calculations of ii), above, we see that

$$\begin{aligned}\mathrm{E}[X] &= \sum_{x,y} x \cdot f_{XY}(x,y) = \sum_x x \sum_y f_{XY}(x,y) = \sum_x x f_X(x) \\ &= \sum_x x \binom{n}{x} p^x (1-p)^{n-x}\end{aligned}$$

This sum is just the expectation of a binomial random variable with parameters p and n, which is we know to be np. Thus

$$\mathrm{E}[X] = np.$$

However, we can also obtain this result more easily. As noted in answer ii) above, we can consider the experiment as a Bernoulli experiment, where one either obtains P or not (i.e., R or Q). The probability of obtaining P is p, so the expected value is $\mathrm{E}[X] = np$. **(1 Mark)**

v) We have

$$\begin{aligned}f_{Y|x}(y) &= \frac{f_{XY}(x,y)}{f_X(x)} = \frac{n!}{x!y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y} \frac{x!(n-x)!}{n! p^x (1-p)^{n-x}} \\ &= \frac{(n-x)!}{y!(n-x-y)!} \left(\frac{q}{1-p-q}\right)^y \left(1 - \frac{q}{1-p}\right)^{n-x}\end{aligned}$$

(1 Mark) Given $n = 10$ and $x = 5$, $P[Y = 3]$ is obtained as

$$f_{Y|5}(3) = \frac{5!}{3!2!} \left(\frac{0.2}{0.1}\right)^3 \left(1 - \frac{0.2}{0.3}\right)^5 = \frac{5!}{3!2!} 2^3/3^5 = 0.33$$

(1 Mark)

Exercise 4.
A fair coin is tossed repeatedly.

i) What is the probability that the 5^{th} head occurs on the 10^{th} toss?

ii) Suppose that we know that the 10^{th} head occurs on the 25^{th} toss. Find the probability density function of the toss number of the 5^{th} head.

(1+3+2 Marks)

Solution.

i) The probability is given by the negative binomial distribution with parameters $p = 0.5$ and $r = 5$. The probability is

$$P[X = 10] = \binom{10-1}{5-1} \cdot \frac{1}{2^{10}} = 0.123.$$

ii) The tenth head occurs on the 25^{th} toss, so nine heads occur in the first 24 tosses. There are $\binom{24}{9}$ ways of selecting the toss numbers for these nine heads. If the 5^{th} head occurs on the x^{th} toss, then 4 heads must have occurred in the first $x-1$ tosses and 4 heads must occur in the following $24-x$ tosses. Here $5 \le x \le 20$. It follows that $f_X \colon \Omega \to \mathbb{R}$ given by

$$f_X(x) = P[X = x] = P[5^{\text{th}} \text{ head on } x^{\text{th}} \text{ toss} \mid 10^{\text{th}} \text{ head on } 25^{\text{th}} \text{ toss}] = \frac{\binom{x-1}{4}\binom{24-x}{4}}{\binom{24}{9}}$$

is the density of the random variable $X \colon S \to \Omega = \{5, 6, \dots, 20\}$, where X is the toss number of the 5^{th} head and S is the sample space for the experiment.

Exercise 5.

Suppose that a hotel has 200 equally-sized rooms. From experience, it is known that 10% of travellers who have reserved a room do not turn up. Use the normal approximation to the binomial distribution to answer the following questions.

i) Suppose that the hotel has received reservations for 215 rooms. What is the probability that there will enough rooms for all travelers who turn up?

ii) Suppose that the hotel has received reservations for 215 rooms. What is the probability that more than 190 rooms (but of course not more than 200) will be occupied?

iii) At what point should the hotel stop accepting reservations if management wants to be at least 99% certain to have enough rooms for all travelers who show up?

(2+2+2 Marks)

Solution. We use the normal approximation to the binomial distribution with parameter $p = 0.1$ and n depending on the situation.

i) We want to find $P[T \leq 200]$, where T denotes the number of travellers who actually turn up. There are $n = 215$ travellers who each have a probability of $p = 0.9$ of turning up. Thus, T follows a binomial distribution with these parameters. Since $nq = 21.5 > 5$ we can approximate this distribution using a normal random variable X with mean $\mu = np = 193.5$ and variance $\sigma^2 = npq = 19.35$. We denote by Z the standard normal random variable. Taking into account the half-unit correction,

$$P[T \leq 200] \approx P[X \leq 200.5] = P\left[Z \leq \frac{200.5 - 193.5}{4.399}\right] = P\left[Z \leq 1.59\right] = 0.94$$

so there is a 94% chance that there will be enough rooms.

ii) As above, we calculate $P[T \geq 190]$. Taking into account the half-unit correction,

$$P[T \geq 191] \approx P[X \geq 190.5] = P\left[Z \geq \frac{190.5 - 193.5}{4.399}\right] = P\left[Z \geq -0.68\right] = 0.752$$

so there is a 75% chance that more than 190 rooms will be occupied.

iii) We need to solve

$$0.99 \geq P[T \leq 200] \approx P[X \leq 200 + 0.5]$$

where X is a normal random variable with mean $0.9R$ and standard deviation $0.3\sqrt{R}$, Rbeing the number of reservations accepted. Hence we need to calculate

$$0.99 \geq P\left[Z \leq \frac{200.5 - 0.9R}{0.3\sqrt{R}}\right]$$

From the table we see that this implies

$$2.33 \leq \frac{200.5 - 0.9R}{0.3\sqrt{R}}.$$

There is only one solution to this equation, sind the right hand side is monotonically decreasing in R. To find the point of equality, we solve $0.2097R \leq (200.5 - 0.9R)^2 = 0.81R^2 + 360.9R + (200.5)^2$. We find the roots of

$$0 = R^2 - 446.16R + 49630,$$

yielding $R = 223.08 \pm 11.6$. The smaller of the roots is 211.48, so the hotel should not accept more than 211 reservations.

Exercise 6.

The target value for the thickness of a machined cylinder is 8 cm. The upper specification limit is 8.2 cm and the lower specification limit is 7.9 cm. A machine produces cylinders that have a mean thickness of 8.1 cm and a standard deviation of 0.1 cm. Assume that the thickness of the cylinders produced by this machine follows a normal distribution.

i) What is the probability that the thickness of a randomly selected cylinder is within specification?

ii) What is the probability that the thickness of a randomly selected cylinder exceeds the target value?

iii) Find a bound $L > 0$ such that the thickness of 90% of all cylinders lies within $(8.1 \pm L)$ cm.

(2+2+2 Marks)

Solution.

i) Denote the random variable "thickess" by T. Then the variable $Z = \frac{T-8.1}{0.1}$ is standard normal and

$$\begin{aligned} P[7.9 \leq T \leq 8.2] &= P\left[\frac{7.9-8.1}{0.1} \leq \frac{T-8.1}{0.1} \leq \frac{8.2-8.1}{0.1}\right] \\ &= P\left[\frac{7.9-8.1}{0.1} \leq Z \leq \frac{8.2-8.1}{0.1}\right] \\ &= P\left[-2 \leq Z \leq 1\right] \\ &= P[0 \leq Z \leq 1] + P[0 \leq Z \leq 2] \\ &= 0.3413 + 0.4772 = 0.8185 \approx 82\%. \end{aligned}$$

(2 Marks)

ii) We have

$$\begin{aligned} P[T \geq 8] &= P\left[\frac{T-8.1}{0.1} \geq \frac{8-8.1}{0.1}\right] \\ &= P\left[Z \geq -1\right] \\ &= P[0 \leq Z \leq 1] + P[0 \leq Z] \\ &= 0.3413 + 0.5 = 0.8413 \approx 84\%. \end{aligned}$$

(2 Marks)

iii) We want to find L such that

$$0.05 = P[T > 8.1 + L] = P\left[\frac{T-8.1}{0.1} > 10L\right] = 1 - P[Z \leq 10L] = 0.5 - P[0 \leq Z \leq 10L]$$

(1 Mark) From the table, $10L = 1.645$ (interpolating between 1.64 and 1.65), so $L = 0.1645$. Thus the thickness of 90% of all cylinders lies within (8.1 ± 0.16) cm. **(1 Mark)**

Exercise 7.
Consider the continuous random variable X with density

$$f_X(x) = \frac{c}{e^{-x} + e^{x}} \qquad \text{for } x \in \mathbb{R}.$$

i) Determine the constant $c \in \mathbb{R}$.

ii) Find $P[X \leq 1]$.

iii) Find the density of the random variable X^2.

(1+1+2 Marks)

Solution.

i) The density must satisfy $\int_{\mathbb{R}} f_X(x)\,dx = 1$. Since

$$\begin{aligned}\int_{\mathbb{R}} f_X(x)\,dx &= c\int_{-\infty}^{\infty} \frac{dx}{e^{-x}+e^{x}}\\ &= c\int_{-\infty}^{\infty} \frac{e^{x}\,dx}{1+e^{2x}}\\ &= c\int_{0}^{\infty} \frac{dy}{1+y^2}\\ &= c\arctan y\big|_0^{\infty} = c\frac{\pi}{2}\end{aligned}$$

we obtain $c = 2/\pi$. **(1 Mark)**

ii) The probability is given by

$$\begin{aligned}P[X \leq 1] &= \frac{2}{\pi}\int_{-\infty}^{1} \frac{dx}{e^{-x}+e^{x}}\\ &= \frac{2}{\pi}\int_{-\infty}^{1} \frac{e^{x}\,dx}{1+e^{2x}}\\ &= \frac{2}{\pi}\int_{0}^{e} \frac{dy}{1+y^2}\\ &= \frac{2}{\pi}\arctan y\big|_0^{e} = \frac{2}{\pi}\arctan e.\end{aligned}$$

(1 Mark)

iii) Note that the function $\varphi\colon \mathbb{R} \to \mathbb{R}_+ \cup \{0\}$, $\varphi(x) = x^2$ is not bijective, so we can't simply apply the theorem for transforming random variables from the lecture. Let $y > 0$. Then, using the fact that f_X is even,

$$\begin{aligned}F_Y(y) = P[Y \leq y] = P[X^2 \leq y] = P[-\sqrt{y} \leq X \leq \sqrt{y}] = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx\\ = 2\int_0^{\sqrt{y}} f_X(x)\,dx.\end{aligned}$$

(1/2 Mark) It follows that

$$f_Y(y) = F_Y'(y) = 2f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = \frac{2}{\pi\sqrt{y}}\frac{1}{e^{-\sqrt{y}} + e^{\sqrt{y}}}$$

for $y > 0$. **(1 Mark)** For $y \leq 0$ we have

$$F_Y(y) = P[Y \leq y] = P[X^2 \leq y] = 0,$$

so $f_Y(y) = 0$ for $y \leq 0$. **(1/2 Mark)**

(2 Marks)

Exercise 8.
The distribution function of the speed (modulus of the velocity) V of a gas molecule is described by the Maxwell-Boltzmann law

$$f_V(v) = \begin{cases} \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{3/2} v^2 e^{-\frac{m}{kT}v^2/2} & v > 0 \\ 0 & v \leq 0 \end{cases}$$

where $m > 0$ is the mass of the molecule, $T > 0$ is its temperature and $k > 0$ is the Boltzmann constant.

i) Find the mean and variance of V.

ii) Find the mean of the kinetic energy $E = mV^2/2$.

iii) Find the probability density f_E of E.

(3 + 2 + 2 Marks)

Solution. i) The mean of V is given by

$$\begin{aligned} \mathrm{E}[V] &= \int_{\mathbb{R}} v f_V(v)\, dv \\ &= \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty \frac{m}{kT} v^3 e^{-\frac{m}{kT}v^2/2}\, dv \\ &= 2\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty v e^{-\frac{m}{kT}v^2/2}\, dv \\ &= 2\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{-1/2} [-e^{-\frac{m}{kT}v^2/2}]_0^\infty \\ &= 2\left(\frac{2kT}{m\pi}\right)^{1/2}. \end{aligned}$$

(1 Mark) Furthermore,

$$\begin{aligned} \mathrm{E}[V^2] &= \int_{\mathbb{R}} v^2 f_V(v)\, dv \\ &= \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty \frac{m}{kT} v^4 e^{-\frac{m}{kT}v^2/2}\, dv \\ &= 3\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty v^2 e^{-\frac{m}{kT}v^2/2}\, dv \\ &= 3\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{-1/2} \int_0^\infty e^{-\frac{m}{kT}v^2/2}\, dv. \end{aligned}$$

Setting $w = \sqrt{m/(kT)}v$, we have

$$\mathrm{E}[V^2] = 6\left(\frac{m}{kT}\right)^{-1} \underbrace{\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-w^2/2}\, dw}_{=1/2} = \frac{3kT}{m}.$$

(1 Mark) It follows that

$$\operatorname{Var} V = \mathrm{E}[V^2] - \mathrm{E}[V]^2 = \frac{kT}{m}\left(3 - \frac{8}{\pi}\right)$$

(1 Mark)

ii) The expectation value of the kinetic energy is given by

$$\mathrm{E}[E] = \frac{m}{2}\mathrm{E}[V^2] = \frac{m}{2}\frac{3kT}{m} = \frac{3}{2}kT.$$

(2 Marks)

iii) Note that the function $\varphi\colon \mathbb{R} \to \mathbb{R}_+ \cup \{0\}$, $\varphi(v) = \frac{m}{2}v^2$ is not bijective, so we can't simply apply the theorem for transforming random variables from the lecture. Let $\varepsilon > 0$. Then

$$F_E(\varepsilon) = P[E \le \varepsilon] = P\left[\frac{m}{2}V^2 \le \varepsilon\right] = P\left[-\sqrt{2\varepsilon/m} \le V \le \sqrt{2\varepsilon/m}\right] = \int_{-\sqrt{2\varepsilon/m}}^{\sqrt{2\varepsilon/m}} f_V(v)\,dv$$
$$= \int_0^{\sqrt{2\varepsilon/m}} f_V(v)\,dx.$$

(1/2 Mark) It follows that

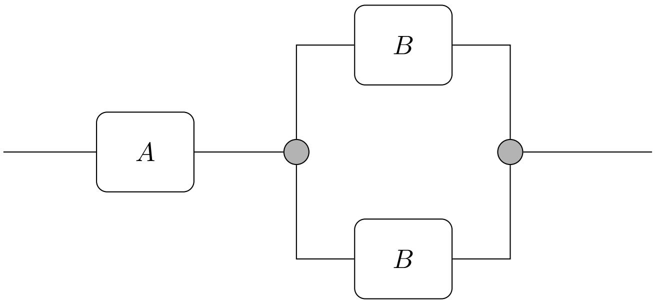
$$\begin{aligned}
f_E(\varepsilon) = F_E'(\varepsilon) &= f_V(\sqrt{2\varepsilon/m}) \cdot \frac{1}{\sqrt{2m\varepsilon}} \\
&= \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{3/2} \frac{2\varepsilon}{m} e^{-\frac{\varepsilon}{kT}} \cdot \frac{1}{\sqrt{2m\varepsilon}} \\
&= \frac{2}{\sqrt{\pi}} (kT)^{-3/2} \sqrt{\varepsilon} e^{-\frac{\varepsilon}{kT}}
\end{aligned}$$

for $\varepsilon > 0$. **(1 Mark)** For $\varepsilon \le 0$ we have

$$F_E(\varepsilon) = P[E \le \varepsilon] = P\left[\frac{m}{2}V^2 \le \varepsilon\right] \le P\left[\frac{m}{2}V^2 \le 0\right] = 0,$$

so $f_E(\varepsilon) = 0$ for $\varepsilon \le 0$. **(1/2 Mark)**

Exercise 9.
Consider the following system of components:



The system will fail if either component A or both components marked B fail. The components A and B have failure densities

$$f_A(t) = \frac{1}{100}e^{-t/100}, \qquad f_B(t) = \frac{1}{50}e^{-t/50}, \qquad t \geq 0,$$

respectively.

i) Find the reliability functions $R_A(t)$ and $R_B(t)$ of component A and each component B.

ii) Find the reliability function $R(t)$ of the system.

iii) What is the expected time of failure of each individual component?

iv) What is the expected time of failure of the system?

(2+2+2+2 Marks)

Solution. i) We have

$$P[t \leq T] = \int_0^T f_A(t)\,dt = \frac{1}{100}\int_0^T e^{-t/100}\,dt = 1 - e^{-T/100}$$
$$R_A(T) = 1 - P[t \leq T] = e^{-T/100}$$
$$R_B(T) = e^{-T/50}$$

(2 Marks)

ii) The reliability R_2 of the subsystem containing the two B components is given by

$$1 - R_2(t) = (1 - R_B(t))^2 = 1 - 2e^{-t/50} + e^{-t/25}$$

so $R_2(t) = 2e^{-t/50} - e^{-t/25}$. **(1 Mark)** The reliability of the entire system is given by

$$R(t) = R_A(t)R_2(T) = 2e^{-3t/100} - e^{-t/20}.$$

(1 Mark)

iii) The expected time of failure for component A is $\mathrm{E}[A] = 100$, for component B $\mathrm{E}[B] = 50$, since their time-to-failures follow exponential distributions with parameters $\beta = 10$ and 5, respectively. **(2 Marks)**

iv) The failure density for the system is

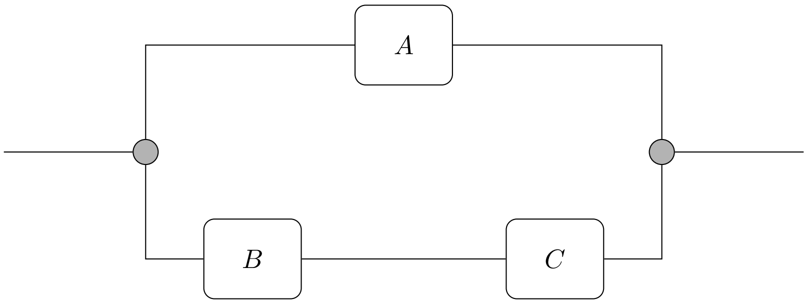
$$f(t) = -R'(t) = \frac{3}{50}e^{-3t/100} - \frac{1}{20}e^{-t/20}$$

(1 Mark) The expected time of failure is

$$2\frac{3}{100}\int_0^\infty te^{-3t/100}\,dt - \frac{1}{20}\int_0^\infty te^{-t/20}\,dt = 2\frac{100}{3} - 20 = \frac{140}{3} = 46.67.$$

(1 Mark)

Exercise 10.
Consider the following system of components:



The components A, B and C have failure densities

$$f_A(t) = \frac{t}{50}e^{-t^2/100}, \qquad f_B(t) = \frac{1}{40}e^{-t/40}, \qquad f_C(t) = \frac{1}{25}e^{-t/25}, \qquad t \geq 0,$$

respectively.

i) Find the reliability functions $R_A(t)$, $R_B(t)$ and $R_C(t)$.

ii) Find the reliability function $R(t)$ of the system.

iii) What is the expected time of failure of each component? Evaluate all integrals explicitly; there should be no gamma functions or other integrals in your answer.

iv) What is the expected time of failure of the system? Evaluate all integrals explicitly; there should be no gamma functions or other integrals in your answer.

(2+2+2+3 Marks)

Solution.

i) We have $P[t \leq T] = \int_0^T f(t)\,dt$ and $R(T) = 1 - P[t \leq T]$, so we obtain

$$\begin{aligned}
R_A(T) &= 1 - \int_0^T f_A(t)\,dt = 1 - \int_0^T \frac{t}{50}e^{-t^2/100}\,dt \\
&= 1 - [-e^{-t^2/100}]_0^T = 1 - (-e^{-T^2/100} + 1) = e^{-T^2/100}, && \textbf{(1 Mark)} \\
R_B(T) &= 1 - \int_0^T f_B(t)\,dt = 1 - \int_0^T \frac{1}{40}e^{-t/40}\,dt = e^{-T/40} && \textbf{(1/2 Mark)} \\
R_C(T) &= e^{-T/25} && \textbf{(1/2 Mark)}
\end{aligned}$$

ii) The reliability R_2 of the subsystem containing the components B and C is given by

$$R_2(t) = R_B(t)R_C(T) = e^{-t/40}e^{-t/25} = e^{-13t/200}$$

and the reliability $R(t)$ of the entire system is given by

$$\begin{aligned}
1 - R(t) &= (1 - R_A(t))(1 - R_2(t)) = (1 - e^{-t^2/100})(1 - e^{-13t/200}) \\
&= 1 - e^{-13t/200} - e^{-t^2/100} + e^{-(2t^2+13t)/200}
\end{aligned}$$

so

$$R(t) = e^{-13t/200} + e^{-t^2/100} - e^{-(2t^2+13t)/200}$$

(2 Marks)

iii) Component A has the failure density of a gamma random variable with $\beta = 2$ and $\alpha = 1/100$. Thus the expected time of failure is

$$\begin{aligned}
\mathrm{E}[X_A] &= \alpha^{-1/\beta}\Gamma(1+1/\beta) = 10\Gamma(3/2) \\
&= 10\int_0^\infty z^{1/2}e^{-z}\,dz \\
&= 20\int_0^\infty w^2 e^{-w^2}\,dw \\
&= 10\int_0^\infty w(-2w)e^{-w^2}\,dw \\
&= 10\int_0^\infty e^{-w^2}\,dw \\
&= \frac{10}{\sqrt{2}}\int_0^\infty e^{-z^2/2}\,dz \\
&= 5\sqrt{\pi}
\end{aligned}$$

(1 Mark) Components B and C have failure densities of an exponential random variable with $\beta = 40$ and 25, respectively, so that

$$\mathrm{E}[X_B] = \beta = 40, \qquad \mathrm{E}[X_C] = 25.$$

(1 Mark)

iv) The expected time of failure of the system is

$$\begin{aligned}
E[X] &= \int_0^\infty tf(t)\,dt = -\int_0^\infty tR'(t)\,dt = -tR(t)|_0^\infty + \int_0^\infty R(t)\,dt \\
&= -0 + 0 + \int_0^\infty e^{-13t/200} + e^{-t^2/100} - e^{-(2t^2+13t)/200}\,dt \\
&= \left[\frac{-200}{13}e^{-13t/200}\right]_0^\infty + \sqrt{50}\int_0^\infty e^{-s^2/2}\,ds - e^{169/1600}\int_0^\infty e^{-(t+13/4)^2/100}\,dt \\
&= \frac{200}{13} + \sqrt{25\pi} - e^{169/1600}\int_{13/4}^\infty e^{-s^2/100}\,ds \\
&= \underbrace{\frac{200}{13}}_{\textbf{(1 Mark)}} + \underbrace{5\sqrt{\pi}}_{\textbf{(1 Mark)}} - \sqrt{50}e^{169/1600}\int_{13/(4\sqrt{50})}^\infty e^{-t^2/2}\,dt \\
&= 15.3846 + 8.8623 - e^{169/1600}\cdot 5\sqrt{2}\cdot\frac{\sqrt{2\pi}}{\sqrt{2\pi}}\int_{0.460}^\infty e^{-t^2/2}\,dt \\
&= 24.2469 - e^{169/1600}\cdot 10\sqrt{\pi}\cdot\left(\frac{1}{2} - \frac{1}{\sqrt{2\pi}}\int_0^{0.460} e^{-t^2/2}\,dt\right) \\
&= 24.2469 - \underbrace{19.6991\cdot(0.5 - 0.1772)}_{\textbf{(1 Mark)}} \\
&= 17.89.
\end{aligned}$$

Exercise 11.
Let X be a continuous random variable with density

$$f_\theta(x) = \begin{cases} \dfrac{\theta+1}{x^{\theta+2}} & \text{for } x > 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is a parameter. Find the method-of-moments estimator for the parameter θ.
(3 Marks)

Solution. We calculate

$$E[X] = \int_1^\infty \frac{\theta+1}{x^{\theta+1}}\,dx = (\theta+1)\int_1^\infty x^{-\theta-1}\,dx = \frac{\theta+1}{-\theta}[x^{-\theta}]_1^\infty = 1 + \frac{1}{\theta}$$

so that

$$\theta = \frac{1}{\mathrm{E}[X]-1}.$$

Using $\widehat{\mathrm{E}[x]} = \overline{X}$ we see that the method-of-moments estmator for θ is

$$\widehat{\theta} = \frac{1}{\overline{X}-1}$$

Exercise 12.
Let (X, f_X) be a continuous random variable following a uniform distribution on the interval $[0, \theta]$, $\theta > 0$, i.e.,

$$f_X(x) = \begin{cases} 1/\theta & 0 \le x \le \theta \\ 0 & \text{otherwise.} \end{cases}$$

i) Find the method-of-moments estimator for θ.

ii) Find the maximum-likelihood estimator for θ.

(2+3 Marks)

Solution. i) The mean of X is given by

$$\mathrm{E}[X] = \int_0^\theta \frac{x}{\theta}\,dx = \frac{\theta}{2}.$$

Thus,

$$\theta = 2\,\mathrm{E}[X].$$

The method-of-moments estimator for θ is then

$$\hat{\theta} = 2\overline{X}.$$

ii) The likelihood function for a random sample $X_1, \ldots, X_n$ of X is given by

$$L(\theta) = \prod_{i=1}^n f_X(x_i) = \begin{cases} \frac{1}{\theta^n} & \text{if } 0 < x_1, \ldots, x_n < \theta \\ 0 & \text{otherwise} \end{cases}$$

In order for $L(\theta) > 0$ we need $\theta \ge x_i$ for $i = 1, \ldots, n$, or $\theta \ge \max_{1 \le i \le n} x_i$. The likelihood function is then maximized if $\theta = \max_{1 \le i \le n} x_i$. The maximum-likelihood estimator hence is

$$\hat{\theta} = \max_{1 \le i \le n} x_i$$

Exercise 13.
A certain type of harddrive is known to have a lifetime given by an exponential distribution with (unknown) parameter $\beta > 0$. To estimate β, n identical and independent hard drives are tested and their times of failure recorded. After time $T > 0$ the test is stopped and $n - m$ hard drives are found to be still working.

Find the maximum likelihood estimator for β.
(4 Marks)

Solution. The density of the exponential distribution is given by

$$f(t) = \begin{cases} \frac{1}{\beta} e^{-t/\beta} & t > 0, \\ 0 & t \leq 0. \end{cases}$$

(1/2 Mark) This gives

$$P[t \leq T] = \int_0^T f(t)\, dt = 1 - e^{-T_k/\beta}.$$

The probability of failing after time T is

$$P[t > T] = 1 - P[t \leq T_k] = e^{-T/\beta}$$

(1/2 Mark) Suppose that the first m hard drives fail at times $T_1, \dots, T_m < T$. The likelihood function is then

$$\begin{aligned} L(\beta) &= f_1(T_1) \cdot f_2(T_2) \dots f_m(T_m) \cdot \underbrace{e^{-T/\beta} \cdots e^{-T/\beta}}_{n-m \text{ terms}} \\ &= \frac{1}{\beta^m} e^{-(T_1 + \cdots + T_m + (n-m)T)/\beta} \end{aligned}$$

(1 Mark) To find the maximum of L, we take the logarithm,

$$\ln(L(\beta)) = -(T_1 + \cdots + T_m + (n-m)T)/\beta - m \ln(\beta)$$

(1/2 Mark) and set the derivative equal to zero,

$$\frac{T_1 + \cdots + T_m + (n-m)T}{\beta^2} - \frac{m}{\beta} = 0.$$

(1/2 Mark) Solving for β gives

$$\widehat{\beta} = \frac{T_1 + \cdots + T_m + (n-m)T}{m}.$$

(1 Mark)

Exercise 14.
A soft drink bottler is studying the internal pressure strength of 1-liter glass non-returnable bottles. A random sample of 16 bottles is tested and the pressure strengths obtained. The data (in units of psi) are shown below.

226.16	202.20	219.54	193.73	208.15	195.45	193.71	200.81
211.14	203.62	188.12	224.39	221.31	204.55	202.21	201.63

i) Draw a histogram for these data.

ii) Create a boxplot, clearly identifying any near and far outliers.

iii) Do you believe the data is normally distributed? Why or why not?

iv) Assuming normality, find a 95% confidence interval for the mean pressure strength μ and the variance σ^2.

(2+2+1 Marks)

Solution.

i)

```
Data = {226.16, 202.20, 219.54, 193.73, 208.15, 195.45, 193.71,
  200.81, 211.14, 203.62, 188.12, 224.39, 221.31, 204.55, 202.21, 201.63}
```

```
{226.16, 202.2, 219.54, 193.73, 208.15, 195.45, 193.71, 200.81,
 211.14, 203.62, 188.12, 224.39, 221.31, 204.55, 202.21, 201.63}
```

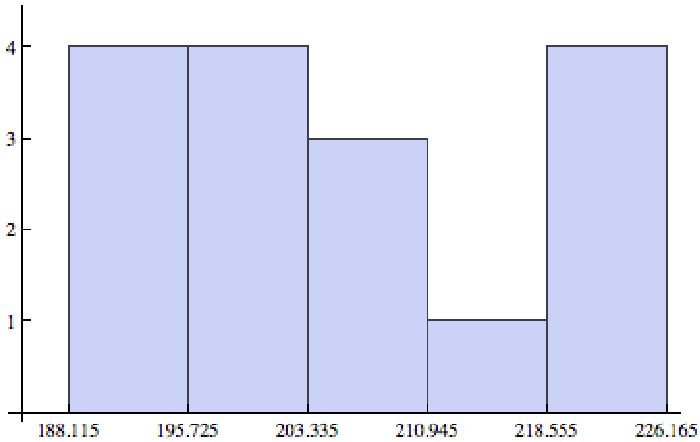
```
n = 5;
CatLength = Round[(Max[Data] - Min[Data]) / n, 0.01]
```

```
7.61
```

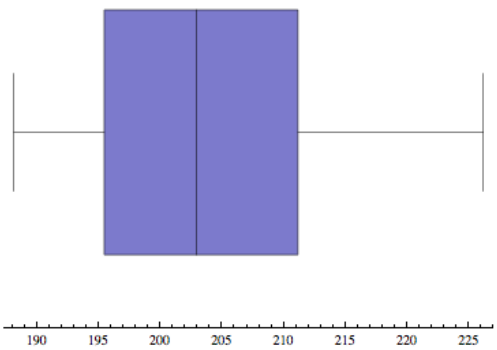
```
categories = Table[Min[Data] - 0.005 + i CatLength, {i, 0, n}]
```

```
{188.115, 195.725, 203.335, 210.945, 218.555, 226.165}
```

```
Histogram[Data, {Min[Data] - 0.005, Min[Data] - 0.005 + n * CatLength,
  CatLength}, AxesOrigin → Min[Data] - 3, Ticks → {categories, {1, 2, 3, 4}}]
```



ii)



iii) The histogram does not have the characteristic shape of a normal distribution, but also shows no significant skewing. The boxplot confirms that the values are distributed symmetrically. The lack of "Gauss-curve" shape may be due to the small sample size. While there is no evidence for a normal distribution, there is also non against it.

iv) We calculate the sample mean and standard deviation: $\overline{x} = 206.045$, $s = 11.572$. Using the value $t_{0.025,15} = 2.131$ we obtain the confidence interval for the mean

$$\overline{x} \pm 2.131 s/\sqrt{16} = 206.045 \pm 5.968.$$

Similarly, using $\chi^2_{0.025,15} = 27.5$, $\chi^2_{0.975,15} = 6.26$ we have

$$\left[15s^2/\chi^2_{0.025,15},\ 15s^2/\chi^2_{0.975,15}\right] = [73.03, 320.82]$$

for the variance.

Exercise 15.
The manufacturer of a power supply is interested in the variability of output voltage. He has tested 16 units, chosen at random, with the following results:

5.34	5.00	5.07	5.25	5.65	5.55	5.35	5.35
4.96	5.54	5.54	4.61	5.28	5.93	5.38	5.47

i) It is thought that the output voltage is normally distributed. Create a stem-and-leaf diagram for the data. Does it support this assumption?

ii) Create a histogram for the data.

iii) Create a boxplot for the data.

((2 + 1) + 2 + 2 Marks)

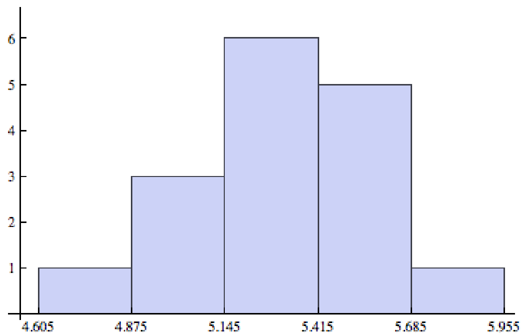
Solution.

i)

Stem	Leaves
46	1
49	6
50	07
52	58
53	4558
54	7
55	445
56	5
59	3

The stem-and-leaf diagram has a shape that might be expected if the data followed a normal distribution. There is no evidence that the data is not normally distributed.

ii)



iii)