Question (1988 STEP I Q16)

Wondergoo is applied to all new cars. It protects them completely against rust for three years, but thereafter the probability density of the time of onset of rust is proportional to $t^2/(1+t^2)^2$ for a car of age $3+t$ years ($t \geqslant 0$). Find the probability that a car becomes rusty before it is $3+t$ years old. Every car is tested for rust annually on the anniversary of its manufacture. If a car is not rusty, it will certainly pass; if it is rusty, it will pass with probability $\frac{1}{2}$. Cars which do not pass are immediately taken off the road and destroyed. What is the probability that a randomly selected new car subsequently fails a test taken on the fifth anniversary of its manufacture? Find also the probability that a car which was destroyed immediately after its fifth anniversary test was rusty when it passed its fourth anniversary test.

Given the probability density after 3 years is proportional to $\frac{t^2}{(1+t^2)^2}$ then we must have that:

$$\begin{aligned} 1 &= A \int_0^\infty \frac{t^2}{(1+t^2)^2} \, \mathrm{d}t \\ &= A \left[-\frac{1}{2} \frac{t}{1+t^2} \right]_0^\infty + \frac{A}{2} \int_0^\infty \frac{1}{1+t^2} \mathrm{d}t \\ &= \frac{A}{2} \frac{\pi}{2} \\ \Rightarrow \qquad A &= \frac{4}{\pi} \end{aligned}$$

In order to fail a test on the fifth anniversary, there are two possibilities for when we went faulty. We could have gone faulty before 4 years, got lucky once and then failed the second test, or gone faulty in the next year and then failed the first test.

$$\begin{aligned} \mathbb{P}(\text{rusty before 4 years}) &= \frac{4}{\pi} \int_0^1 \frac{t^2}{(1+t^2)^2} \mathrm{d}t \\ &= \frac{4}{\pi} \left[-\frac{1}{2} \frac{t}{1+t^2} \right]_0^1 + \frac{2}{\pi} \int_0^1 \frac{1}{1+t^2} \mathrm{d}t \\ &= -\frac{1}{\pi} + \frac{2}{\pi} \frac{\pi}{4} \\ &= \frac{1}{2} - \frac{1}{\pi} \\ &\approx 0.181690 \cdots \end{aligned}$$

$$\begin{aligned} \mathbb{P}(\text{rusty before 5 years}) &= \frac{4}{\pi} \int_0^1 \frac{t^2}{(1+t^2)^2} \mathrm{d}t \\ &= \frac{4}{\pi} \left[-\frac{1}{2} \frac{t}{1+t^2} \right]_0^2 + \frac{2}{\pi} \int_0^2 \frac{1}{1+t^2} \mathrm{d}t \\ &= -\frac{4}{5\pi} + \frac{2}{\pi} \tan^{-1} 2 \\ &\approx 0.450184 \cdots \end{aligned}$$

Therefore:

$$\begin{aligned}\mathbb{P}(\text{fails 5th anniversary}) &= \mathbb{P}(\text{rusty before 4 years})\mathbb{P}(\text{pass one, fail other})+ \\ &\quad + \mathbb{P}(\text{rusty between 4 and 5 years})\mathbb{P}(\text{fail}) \\ &= 0.181690\cdots \cdot \frac{1}{4} + \frac{1}{2}(0.450184\cdots - 0.181690\cdots) \\ &= \frac{1}{2}0.450184\cdots - \frac{1}{4}0.181690\cdots \\ &= 0.1796688\cdots \\ &= 18.0\% \text{ (3 s.f.)}\end{aligned}$$

We also must have that:

$$\begin{aligned}\mathbb{P}(\text{rusty at 4 years}|\text{destroyed at 5}) &= \frac{\mathbb{P}(\text{rusty at 4 years and destroyed at 5})}{\mathbb{P}(\text{destroyed at 5})} \\ &= \frac{0.181690\cdots \cdot \frac{1}{4}}{\frac{1}{2}0.450184\cdots - \frac{1}{4}0.181690\cdots} \\ &= 0.252811\cdots \\ &= 25.3\% \text{ (3 s.f.)}\end{aligned}$$

Question (1989 STEP I Q14)

The prevailing winds blow in a constant southerly direction from an enchanted castle. Each year, according to an ancient tradition, a princess releases 96 magic seeds from the castle, which are carried south by the wind before falling to rest. South of the castle lies one league of grassy parkland, then one league of lake, then one league of farmland, and finally the sea. If a seed falls on land it will immediately grow into a fever tree. (Fever trees do not grow in water). Seeds are blown independently of each other. The random variable L is the distance in leagues south of the castle at which a seed falls to rest (either on land or water). It is known that the probability density function f of L is given by

$$\mathrm{f}(x) = \begin{cases} \frac{1}{2} - \frac{1}{8}x & \text{for } 0 \leqslant x \leqslant 4, \\ 0 & \text{otherwise.} \end{cases}$$

What is the mean number of fever trees which begin to grow each year?

(i) The random variable Y is defined as the distance in leagues south of the castle at which a new fever tree grows from a seed carried by the wind. Sketch the probability density function of Y, and find the mean of Y.

(ii) One year messengers bring the king the news that 23 new fever trees have grown in the farmland. The wind never varies, and so the king suspects that the ancient tradition have not been followed properly. Is he justified in his suspicions?

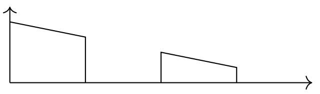
$$\mathbb{P}(\text{fever tree grows}) = \mathbb{P}(0 \leq L \leq 1) + \mathbb{P}(2 \leq L \leq 3)$$

$$= \int_0^1 \frac{1}{2} - \frac{1}{8}x\mathrm{d}x + \int_2^3 \frac{1}{2} - \frac{1}{8}x\mathrm{d}x$$
$$= \left[\frac{1}{2}x - \frac{1}{16}x^2\right]_0^1 + \left[\frac{1}{2}x - \frac{1}{16}x^2\right]_2^3$$
$$= \frac{1}{2} - \frac{1}{16} + \frac{3}{2} - \frac{9}{16} - 1 + \frac{4}{16}$$
$$= \frac{5}{8}$$

The expected number of fever trees is just $96 \cdot \frac{5}{8} = 60$.

(i) $f_Y(t)$ must match the distribution for L, but limited to the points we care about, therefore it should be:

$$f_Y(t) = \begin{cases} (\frac{4}{5} - \frac{1}{5}t) & \text{if } t \in [0,1] \cup [2,3] \\ 0 & \text{otherwise} \end{cases}$$



$$\mathbb{E}(Y) = \frac{1}{2} \cdot \frac{1}{5}(4 - \frac{1}{2}) + \frac{5}{2} \cdot (1 - \frac{1}{5}(4 - \frac{1}{2}))$$
$$= \frac{1}{2} \cdot \frac{7}{10} + \frac{5}{2} \cdot \frac{3}{10}$$
$$= \frac{22}{20}$$
$$= \frac{11}{10}$$

(ii) Given the seeds are blown independently and the wind hasn't changed, it is reasonable to model the number of fever trees as $B(96, \frac{5}{8})$, it is also acceptable to approximate this using a Normal distribution, ie $N(60, 22.5)$, 23 is $\frac{23-60}{\sqrt{22.5}}$ is a very negative number, so he should be extremely suspicious.

Question (1993 STEP I Q14)

When he sets out on a drive Mr Toad selects a speed V kilometres per minute where V is a random variable with probability density

$$\alpha v^{-2}\mathrm{e}^{-\alpha v^{-1}}$$

and α is a strictly positive constant. He then drives at constant speed, regardless of other drivers, road conditions and the Highway Code. The traffic lights at the Wild Wood cross-roads change from red to green when Mr Toad is exactly 1 kilometre away in his journey towards them. If the traffic light is green for g minutes, then red for r minutes, then green for g minutes, and so on, show that the probability that he passes them after $n(g+r)$ minutes but before $n(g+r)+g$ minutes, where n is a positive integer, is

$$\mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha(n(g+r))+g}.$$

Find the probability $\mathrm{P}(\alpha)$ that he passes the traffic lights when they are green.

Show that $\mathrm{P}(\alpha) \to 1$ as $\alpha \to \infty$ and, by noting that $(\mathrm{e}^x - 1)/x \to 1$ as $x \to 0$, or otherwise, show that

$$\mathrm{P}(\alpha) \to \frac{g}{r+g} \qquad \text{as } \alpha \to 0.$$

[NB: the traffic light show only green and red - not amber.]

Question (1994 STEP I Q14)

Each of my n students has to hand in an essay to me. Let T_i be the time at which the ith essay is handed in and suppose that $T_1, T_2, \ldots, T_n$ are independent, each with probability density function $\lambda \mathrm{e}^{-\lambda t}$ $(t \geqslant 0)$. Let T be the time I receive the first essay to be handed in and let U be the time I receive the last one.

(i) Find the mean and variance of T_i.

(ii) Show that $\mathrm{P}(U \leqslant u) = (1 - \mathrm{e}^{-\lambda u})^n$ for $u \geqslant 0$, and hence find the probability density function of U.

(iii) Obtain $\mathrm{P}(T > t)$, and hence find the probability density function of T.

(iv) Write down the mean and variance of T.

Question (1999 STEP II Q13)

A stick is broken at a point, chosen at random, along its length. Find the probability that the ratio, R, of the length of the shorter piece to the length of the longer piece is less than r. Find the probability density function for R, and calculate the mean and variance of R.

Question (2001 STEP II Q13)

The life times of a large batch of electric light bulbs are independently and identically distributed. The probability that the life time, T hours, of a given light bulb is greater than t hours is given by

$$\mathbb{P}(T > t) \;=\; \frac{1}{(1+kt)^\alpha}\,,$$

where α and k are constants, and $\alpha > 1$. Find the median M and the mean m of T in terms of α and k. Nine randomly selected bulbs are switched on simultaneously and are left until all have failed. The fifth failure occurs at 1000 hours and the mean life time of all the bulbs is found to be 2400 hours. Show that $\alpha \approx 2$ and find the approximate value of k. Hence estimate the probability that, if a randomly selected bulb is found to last M hours, it will last a further $m - M$ hours.

Question (2005 STEP I Q14)

The random variable X can take the value $X = -1$, and also any value in the range $0 \le X < \infty$. The distribution of X is given by

$$\mathbb{P}(X = -1) = m\,, \qquad \mathbb{P}(0 \le X \le x) = k(1 - \mathrm{e}^{-x})\,,$$

for any non-negative number x, where k and m are constants, and $m < \frac{1}{2}$.

(i) Find k in terms of m.

(ii) Show that $\mathbb{E}(X) = 1 - 2m$.

(iii) Find, in terms of m, $\mathrm{Var}(X)$ and the median value of X.

(iv) Given that

$$\int_0^\infty y^2 \mathrm{e}^{-y^2}\,\mathrm{d}y = \tfrac{1}{4}\sqrt{\pi}\,,$$

find $\mathbb{E}\big(|X|^{\frac{1}{2}}\big)$ in terms of m.

Question (2006 STEP II Q14)

Sketch the graph of

$$y = \frac{1}{x \ln x} \text{ for } x > 0,\ x \neq 1.$$

You may assume that $x \ln x \to 0$ as $x \to 0$. The continuous random variable X has probability density function

$$f(x) = \begin{cases} \dfrac{\lambda}{x \ln x} & \text{for } a \leq x \leq b\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

where a, b and λ are suitably chosen constants.

(i) In the case $a = 1/4$ and $b = 1/2$, find λ.

(ii) In the case $\lambda = 1$ and $a > 1$, show that $b = a^{\mathrm{e}}$.

(iii) In the case $\lambda = 1$ and $a = \mathrm{e}$, show that $\mathbb{P}(\mathrm{e}^{3/2} \leq X \leq \mathrm{e}^2) \approx \frac{31}{108}$.

(iv) In the case $\lambda = 1$ and $a = \mathrm{e}^{1/2}$, find $\mathbb{P}(\mathrm{e}^{3/2} \leq X \leq \mathrm{e}^2)$.

(i)

$$\begin{aligned} 1 &= \int_{1/4}^{1/2} \frac{\lambda}{x \ln x}\, dx \\ &= \lambda \left[\ln|\ln x|\right]_{1/4}^{1/2} \\ &= \lambda\left(\ln|-\ln 2| - \ln|-2\ln 2|\right) \\ &= \lambda(-\ln 2) \end{aligned}$$

So $\lambda = -\frac{1}{\ln 2} = \frac{1}{\ln \frac{1}{2}}$

(ii)

$$\begin{aligned} 1 &= \int_a^b \frac{1}{x \ln x}\, dx \\ &= \left[\ln|\ln x|\right]_a^b \\ &= (\ln\ln b - \ln\ln a) \\ &= \ln\left(\frac{\ln b}{\ln a}\right) \end{aligned}$$

So $b = e^a$

(iii) If $\lambda = 1, a = e, b = e^e$, then

$$\mathbb{P}(\mathrm{e}^{3/2} \leq X \leq \mathrm{e}^2) = \int_{e^{3/2}}^{e^2} \frac{1}{x \ln x}\, dx$$

$$
\begin{aligned}
&= [\ln \ln x]_{e^{3/2}}^{e^2} \\
&= \ln 2 - \ln \frac{3}{2} \\
&= \ln \frac{4}{3} \\
&= \ln \left(1 + \frac{1}{3}\right) \\
&\approx \frac{1}{3} - \frac{1}{2 \cdot 3^2} + \frac{1}{3 \cdot 3^3} - \frac{1}{4 \cdot 3^4} \\
&= \frac{31}{108}
\end{aligned}
$$

(iv) Note that $2 > e^{\frac{1}{2}} > 1$ so $a = e^{\frac{1}{2}}, b = e^{\frac{e}{2}}$. Since $3 > e \Rightarrow e^{3/2} > e^{\frac{e}{2}}$ this probability is out of range, therefore $\mathbb{P}(\mathrm{e}^{3/2} \leq X \leq \mathrm{e}^2) = 0$

Question (2011 STEP I Q13)

In this question, you may use without proof the following result:

$$\int \sqrt{4 - x^2}\, \mathrm{d}x = 2 \arcsin(\tfrac{1}{2}x) + \tfrac{1}{2}x\sqrt{4 - x^2} + c\,.$$

A random variable X has probability density function f given by

$$f(x) = \begin{cases} 2k & -a \leq x < 0 \\ k\sqrt{4 - x^2} & 0 \leq x \leq 2 \\ 0 & \text{otherwise,} \end{cases}$$

where k and a are positive constants.

(i) Find, in terms of a, the mean of X.

(ii) Let d be the value of X such that $\mathbb{P}(X > d) = \frac{1}{10}$. Show that $d < 0$ if $2a > 9\pi$ and find an expression for d in terms of a in this case.

(iii) Given that $d = \sqrt{2}$, find a.

First notice that

$$
\begin{aligned}
1 &= \int_{-a}^{2} f(x) \mathrm{d}x \\
&= 2ka + k\pi \\
\Rightarrow \quad k &= (\pi + 2a)^{-1}
\end{aligned}
$$

(i)

$$\mathbb{E}[X] = \int_{-a}^{2} x f(x) \mathrm{d}x$$

$$
\begin{aligned}
&= \int_{-a}^{0} 2kx\mathrm{d}x + k\int_{0}^{2} x\sqrt{4-x^2}\mathrm{d}x \\
&= \left[kx^2\right]_{-a}^{0} + k\left[-\frac{1}{3}(4-x^2)^{\frac{3}{2}}\right]_0^2 \\
&= -ka^2 + \frac{8}{3}k \\
&= \frac{\frac{8}{3} - a^2}{\pi + 2a}
\end{aligned}
$$

(ii) Consider $\mathbb{P}(X<0)$ then $d < 0 \Leftrightarrow \mathbb{P}(X<0) > \frac{9}{10}$

$$
\begin{aligned}
\frac{9}{10} &< \mathbb{P}(X<0) \\
&= \int_{-a}^{0} 2k\mathrm{d}x \\
&= \frac{2a}{\pi+2a} \\
\Leftrightarrow \quad 9\pi &< 2a
\end{aligned}
$$

$$
\begin{aligned}
\frac{9}{10} &= \int_{-a}^{d} 2k\mathrm{d}x \\
&= \frac{2(d+a)}{\pi+2a} \\
\Rightarrow \quad 9\pi &= 2a + 20d \\
\Rightarrow \quad d &= \frac{2a - 9\pi}{20}
\end{aligned}
$$

(iii) Suppose $d = \sqrt{2}$ then

$$
\begin{aligned}
\frac{1}{10} &= \int_{\sqrt{2}}^{2} f(x)\mathrm{d}x \\
&= \int_{\sqrt{2}}^{2} k\sqrt{4-x^2}\mathrm{d}x \\
&= k\left[2\sin^{-1}\tfrac{1}{2}x + \tfrac{1}{2}x\sqrt{4-x^2}\right]_{\sqrt{2}}^{2} \\
&= k\left(\pi - \frac{\pi}{2} - 1\right) \\
\Rightarrow \quad \pi + 2a &= 5\pi - 10 \\
\Rightarrow \quad a &= 2\pi - 5
\end{aligned}
$$

Question (2012 STEP I Q12)

Fire extinguishers may become faulty at any time after manufacture and are tested annually on the anniversary of manufacture. The time T years after manufacture until a fire extinguisher becomes faulty is modelled by the continuous probability density function

$$f(t) = \begin{cases} \frac{2t}{(1+t^2)^2} & \text{for } t \geq 0\,, \\ \\ 0 & \text{otherwise.} \end{cases}$$

A faulty fire extinguisher will fail an annual test with probability p, in which case it is destroyed immediately. A non-faulty fire extinguisher will always pass the test. All of the annual tests are independent. Show that the probability that a randomly chosen fire extinguisher will be destroyed exactly three years after its manufacture is $p(5p^2 - 13p + 9)/10$. Find the probability that a randomly chosen fire extinguisher that was destroyed exactly three years after its manufacture was faulty 18 months after its manufacture.

The probability it becomes faulty in each year is:

$$\begin{aligned} \mathbb{P}(\text{faulty in Y1}) &= \int_0^1 \frac{2t}{(1+t^2)^2}\,dt \\ &= \left[-\frac{1}{(1+t^2)}\right]_0^1 \\ &= 1 - \frac{1}{2} = \frac{1}{2} \\ \mathbb{P}(\text{faulty in Y2}) &= \frac{1}{2} - \frac{1}{5} = \frac{3}{10} \\ \mathbb{P}(\text{faulty in Y3}) &= \frac{1}{5} - \frac{1}{10} = \frac{1}{10} \end{aligned}$$

The probability of failing for the first time after exactly 3 years is:

$$\begin{aligned} \mathbb{P}(\text{faulty in Y1, } PPF) + \mathbb{P}(\text{faulty in Y2, } PF) + +\mathbb{P}(\text{faulty in Y3, } F) \\ &= \frac{1}{2}(1-p)^2 p + \frac{3}{10}(1-p)p + \frac{1}{10}p \\ &= \frac{p}{10}\left(5(1-p)^2 + 3(1-p) + 1\right) \\ &= \frac{p}{10}\left(5 - 10p + 5p^2 + 3 - 3p + 1\right) \\ &= \frac{p}{10}\left(9 - 13p + 5p^2\right) \end{aligned}$$

as required.

The probability that a randomly chosen fire extinguisher that was destroyed exactly three years after its manufacture was faulty 18 months after its manufacture is:

$$\mathbb{P}(\text{faulty 18 months after}|\text{fails after 3 tries}) = \frac{\mathbb{P}(\text{faulty 18 months after and fails after 3 tries})}{\mathbb{P}(\text{fails after exactly 3 tries})}$$

We can compute $\mathbb{P}(\text{faulty 18 months after and fails after 3 tries})$ by looking at 2 cases, fails between 12 months and 18 years, and between 0 years and 1 year.

$$\begin{aligned}\mathbb{P}(\text{faulty between 1y and 18m}) &= \int_1^{\frac{3}{2}} \frac{2t}{(1+t^2)^2}\,dt \\ &= \left[-\frac{1}{(1+t^2)}\right]_1^{\frac{3}{2}} \\ &= \frac{1}{2} - \frac{4}{13} = \frac{5}{26}\end{aligned}$$

So the probability is:

$$\begin{aligned}\mathbb{P} &= \frac{\frac{5}{26}(1-p)p + \frac{1}{2}(1-p)^2p}{\frac{p}{10}\left(9-13p+5p^2\right)} \\ &= \frac{\frac{25}{13}(1-p) + 13(1-p)^2p}{9-13p+5p^2} \\ &= \frac{5}{13}\frac{(1-p)\left(5+13(1-p)\right)}{9-13p+5p^2} \\ &= \frac{5}{13}\frac{(1-p)\left(18-13p\right)}{9-13p+5p^2}\end{aligned}$$

Question (2014 STEP I Q13)

A continuous random variable X has a *triangular* distribution, which means that it has a probability density function of the form

$$f(x) = \begin{cases} g(x) & \text{for } a < x \leq c \\ h(x) & \text{for } c \leq x < b \\ 0 & \text{otherwise,} \end{cases}$$

where $g(x)$ is an increasing linear function with $g(a) = 0$, $h(x)$ is a decreasing linear function with $h(b) = 0$, and $g(c) = h(c)$. Show that $g(x) = \dfrac{2(x-a)}{(b-a)(c-a)}$ and find a similar expression for $h(x)$.

(i) Show that the mean of the distribution is $\frac{1}{3}(a+b+c)$.

(ii) Find the median of the distribution in the different cases that arise.

Since $\int f(x)\,dx = 1$, and $f(x)$ is a triangle with base $b-a$, it must have height $\frac{2}{b-a}$ in order to have the desired area.

Since $g(a) = 0, g(c) = \frac{2}{b-a}$, $g(x) = A(x-a)$ and $\frac{2}{b-a} = A(c-a) \Rightarrow g(x) = \frac{2(x-a)}{(b-a)(c-a)}$ as required.

Similarly, $h(x) = B(x-b)$ and $\frac{2}{b-a} = B(c-b) \Rightarrow h(x) = \frac{2(b-x)}{(b-a)(b-c)}$

The mean of the distribution will be:

$$\int_a^b xf(x)\,dx = \int_a^c xg(x)\,dx + \int_c^b xh(x)\,dx$$

$$
\begin{aligned}
&= \frac{2}{(b-a)(c-a)}\int_a^c x(x-a)dx + \frac{2}{(b-a)(b-c)}\int_c^b x(b-x)\,dx \\
&= \frac{2}{(b-a)}\left(\frac{1}{c-a}\left[\frac{x^3}{3} - a\frac{x^2}{2}\right]_a^c + \frac{1}{b-c}\left[b\frac{x^2}{2} - \frac{x^3}{3}\right]_c^b\right) \\
&= \frac{2}{(b-a)}\left(\frac{1}{c-a}\left(\frac{c^3}{3} - a\frac{c^2}{2} - \frac{a^3}{3} + \frac{a^3}{2}\right) + \frac{1}{b-c}\left(\frac{b^3}{2} - \frac{b^3}{3} - \frac{bc^2}{2} + \frac{c^3}{3}\right)\right) \\
&= \frac{2}{(b-a)}\left(\left(\frac{c^2+ac+a^2}{3} - \frac{a(a+c)}{2}\right) + \left(\frac{b(b+c)}{2} - \frac{b^2+bc+c^2}{3}\right)\right) \\
&= \frac{2}{(b-a)}\left(\frac{2c^2+2ac+2a^2}{6} - \frac{3a^2+3ac}{6} + \frac{3b^2+3bc}{6} - \frac{2b^2+2bc+2c^2}{6}\right) \\
&= \frac{2}{(b-a)}\left(\frac{-a^2+b^2-ac+bc}{6}\right) \\
&= \frac{a+b+c}{3}
\end{aligned}
$$

The median M satisfies:

$$\int_a^M f(x)\,dx = \frac{1}{2}$$

The left hand triangle will have area: $\frac{c-a}{b-a}$ which will be $\geq \frac{1}{2}$ if $c \geq \frac{a+b}{2}$. In this case we need

$$
\begin{aligned}
\frac{(M-a)^2}{(b-a)(c-a)} &= \frac{1}{2} \\
\Rightarrow \qquad M &= a + \sqrt{\frac{1}{2}(b-a)(c-a)}
\end{aligned}
$$

Otherwise, we need:

$$
\begin{aligned}
\frac{(b-M)^2}{(b-a)(b-c)} &= \frac{1}{2} \\
\Rightarrow \qquad M &= b - \sqrt{\frac{1}{2}(b-a)(b-c)}
\end{aligned}
$$

These are consistent, if $c = \frac{b+a}{2}$

Question (2019 STEP II Q12)

The random variable X has the probability density function on the interval $[0,1]$:

$$f(x) = \begin{cases} nx^{n-1} & 0 \le x \le 1, \\ 0 & \text{elsewhere,} \end{cases}$$

where n is an integer greater than 1.

(i) Let $\mu = E(X)$. Find an expression for μ in terms of n, and show that the variance, σ^2, of X is given by

$$\sigma^2 = \frac{n}{(n+1)^2(n+2)}.$$

(ii) In the case $n = 2$, show without using decimal approximations that the interquartile range is less than 2σ.

(iii) Write down the first three terms and the $(k+1)$th term (where $0 \le k \le n$) of the binomial expansion of $(1+x)^n$ in ascending powers of x. By setting $x = \frac{1}{n}$, show that μ is less than the median and greater than the lower quartile. Note: You may assume that

$$1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots < 4.$$

(i)

$$\begin{aligned} \mu &= \mathbb{E}[X] \\ &= \int_0^1 xf(x)\mathrm{d}x \\ &= \int_0^1 nx^n \mathrm{d}x \\ &= \frac{n}{n+1} \end{aligned}$$

$$\begin{aligned} \text{Var}[X] &= \sigma^2 \\ &= \mathbb{E}[X^2] - \mu^2 \\ &= \int_0^1 x^2 f(x)\mathrm{d}x - \mu^2 \\ &= \int_0^1 nx^{n+1}\mathrm{d}x - \mu^2 \\ &= \frac{n}{n+2} - \frac{n^2}{(n+1)^2} \\ &= \frac{n(n+1)^2 - n^2(n+2)}{(n+1)^2(n+2)} \\ &= \frac{n}{(n+1)^2(n+2)} \end{aligned}$$

(ii)

$$\begin{aligned}
\frac{1}{4} &= \int_0^{Q_1} 2x\mathrm{d}x \\
&= Q_1^2 \\
\Rightarrow \qquad Q_1 &= \frac{1}{2} \\
\frac{3}{4} &= \int_0^{Q_3} 2x\mathrm{d}x \\
&= Q_3^2 \\
\Rightarrow \qquad Q_3 &= \frac{\sqrt{3}}{2}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \qquad IQR &= Q_3 - Q_1 = \frac{\sqrt{3}-1}{2} \\
2\sigma &= 2\sqrt{\frac{2}{3^2 \cdot 4}} \\
&= \frac{\sqrt{2}}{3}
\end{aligned}$$

$$\begin{aligned}
2\sigma - IRQ &= \frac{\sqrt{2}}{3} - \frac{\sqrt{3}-1}{2} \\
&= \frac{2\sqrt{2} - 3\sqrt{3} + 3}{6} \\
(3+2\sqrt{2})^2 &= 17 + 12\sqrt{2} > 29 \\
(3\sqrt{3})^2 &= 27
\end{aligned}$$

Therefore $2\sigma > IQR$

(iii)

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \cdots + \binom{n}{k}x^k + \cdots$$

$$\begin{aligned}
Q_1^{-n} &= 4 \\
Q_2^{-n} &= 2 \\
\mu &= \frac{n}{n+1} \\
\Rightarrow \quad \mu^{-n} &= \left(1 + \frac{1}{n}\right)^n \\
&\geq 1 + n\frac{1}{n} + \cdots > 2 \\
\Rightarrow \qquad \mu &< Q_2
\end{aligned}$$

$$
\begin{aligned}
\mu^{-n} &= \left(1+\frac{1}{n}\right)^n \\
&= 1 + n\frac{1}{n} + \frac{n(n-1)}{2!}\frac{1}{n^2} + \cdots + \frac{n(n-1)\cdots(n-k+1)}{k!}\frac{1}{n^k} + \cdots \\
&= 1 + 1 + \left(1-\frac{1}{n}\right)\frac{1}{2!} + \cdots + \left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\frac{1}{k!} + \cdots \\
&< 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{k!} \\
&< 4
\end{aligned}
$$

$$\Rightarrow \qquad \mu > Q_1$$