By definition

$$1 = \int_{-a}^{0} 2k\,dx + \int_{0}^{2} k\sqrt{4-x^2}\,dx$$

$$= 2ak + k\left[2\sin^{-1}\left(\tfrac{1}{2}x\right) + \tfrac{1}{2}x\sqrt{4-x^2}\right]_0^2$$

$$= 2ak + k\pi$$

$$\therefore k = \frac{1}{2a+\pi}$$

$$\text{Mean}(X) = \int_{-a}^{0} 2kx\,dx + \int_{0}^{2} kx\sqrt{4-x^2}\,dx$$

$$= \left[kx^2\right]_{-a}^{0} + \left[-\frac{k}{3}(4-x^2)^{3/2}\right]_0^2$$

$$= -ka^2 + \frac{8k}{3}$$

$$= \frac{8-3a^2}{3(2a+\pi)}$$

By above $P(X>0) = \dfrac{\pi}{2a+\pi}$ (2nd integral)

$$P(X>0) > \frac{1}{10} \Leftrightarrow \frac{\pi}{2a+\pi} < \frac{1}{10} \Leftrightarrow 9\pi < 2a$$

So $2a > 9\pi \Rightarrow d < 0$

So $P(X<d) = 2k(d+a) = \dfrac{9}{10}$

$$\therefore d = \frac{9}{20k} - a = \frac{9\pi - 2a}{20}$$

Question

A random variable X has probability density function f given by

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

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 $P(X > d) = \frac{1}{10}$. 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.

(13) contd

$$d = \sqrt{2} \Rightarrow \int_{\sqrt{2}}^{2} k\sqrt{(4-x^2)}\,dx = \frac{1}{10}$$

$$\Rightarrow k\left[2\sin^{-1}\left(\tfrac{1}{2}x\right) + \tfrac{1}{2}x\sqrt{4-x^2}\right]_{\sqrt{2}}^{2} = \frac{1}{10}$$

$$\Rightarrow k\left[\pi - \frac{\pi}{2} - 1\right] = \frac{1}{10}$$

$$\Rightarrow \quad 5\pi - 10 = 2a + \pi$$

$$\Rightarrow \quad a = 2\pi - 5$$

STEP I-2015-12

The number X of casualties arriving at a hospital each day follows a Poisson distribution with mean 8; that is,

$$P(X=n)=\frac{1}{n!}\mathrm{e}^{-8}8^n, n=0,1,2,\dots$$

Casualties require surgery with probability $\frac{1}{4}$. The number of casualties arriving on any given day is independent of the number arriving on any other day and the casualties require surgery independently of one another.

(i) What is the probability that, on a day when exactly n casualties arrive, exactly r of them require surgery?

(ii) Prove (algebraically) that the number requiring surgery each day also follows a Poisson distribution, and state its mean.

(iii) Given that in a particular randomly chosen week a total of 12 casualties require surgery on Monday and Tuesday, what is the probability that 8 casualties require surgery on Monday? You should give your answer as a fraction in its lowest terms.

(i) P (r surgeries | n arrivals) = ${}^nC_r(\frac{1}{4})^r(\frac{3}{4})^{n-r}$

(ii)

$$\mathrm{P}\,(r\text{surgeries})=\sum_{n=r}^{\infty}{}^nC_r\left(\frac{1}{4}\right)^r\left(\frac{3}{4}\right)^{n-r}\frac{8^n}{n!}\mathrm{e}^{-8}$$

$$=\sum_{n=r}^{\infty}\frac{\cancel{n!}}{(n-r)!r!}\left(\frac{1}{4}\right)^r\left(\frac{4}{3}\right)^r\left(\frac{3}{4}\right)^n\frac{8^n}{\cancel{n!}}\mathrm{e}^{-8}=\frac{\left(\frac{1}{3}\right)^r}{r!}\mathrm{e}^{-8}\sum_{n=r}^{\infty}6^n\frac{1}{(n-r)!}=\frac{\left(\frac{1}{3}\right)^r\cdot 6^r}{r!}\mathrm{e}^{-8}\sum_{t=0}^{\infty}6^t\frac{1}{t!}$$

$$=\frac{2^r}{r!}\mathrm{e}^{-8}\mathrm{e}^{6}=\frac{2^r}{r!}\mathrm{e}^{-2}$$

so r follows a Poisson distribution with mean 2.

(iii)

Let m = number of surgeries on Monday, t = number of surgeries on Tuesday: m is Po(2), t is Po(2), m+t is Po(4)

$$\mathrm{P}(m=8,t=4|m+t=12)=\frac{\mathrm{P}(m=8)\cdot\mathrm{P}(t=4)}{\mathrm{P}(m+t=12)}=\frac{\frac{2^8}{8!}\mathrm{e}^{-2}\cdot\frac{2^4}{4!}\mathrm{e}^{-2}}{\frac{4^{12}}{12!}\mathrm{e}^{-4}}=\frac{\frac{1}{8!}\cdot\frac{1}{4!}}{\frac{2^{12}}{12!}}=\frac{12\times 11\times 10\times 9}{4!\times 2^{12}}$$

$$=\frac{495}{2^{12}}$$

Note on (ii): The question asks for an algebraic proof. However, the result can be proved without algebra. The casualties show up along a continuous timeline as occurrences of a random event of constant expected frequency per unit time, with each occurrence independent of the others. To get that continuous timeline to show the surgery cases, we toss a tetrahedral dice at each casualty case, and light up the case if the dice lands on one particular face, fade it out if the dice lands on any other face. The new timeline still shows occurrences of a random event of constant expected frequency per unit time, with each occurrence independent of the others, so it's still a Poisson process, only now the expected frequency per unit time is $\frac{1}{4} \times 8 = 2$.

Question 12

In this question, you may use without proof the results:

$$\sum_{r=1}^{n} r = \tfrac{1}{2}n(n+1) \qquad \textit{and} \qquad \sum_{r=1}^{n} r^2 = \tfrac{1}{6}n(n+1)(2n+1).$$

The independent random variables X_1 *and* X_2 *each take values* $1, 2, \ldots, N$, *each value being equally likely. The random variable* X *is defined by*

$$X = \begin{cases} X_1 & \text{if } X_1 \geqslant X_2 \\ X_2 & \text{if } X_2 \geqslant X_1. \end{cases}$$

(i) *Show that* $\mathrm{P}(X = r) = \dfrac{2r-1}{N^2}$ *for* $r = 1, 2, \ldots, N$.

We have $X = r$ when either $X_1 = r$ and $X_2 < r$, or $X_2 = r$ and $X_1 < r$ or $X_1 = X_2 = r$. Therefore

$$\begin{aligned} \mathrm{P}(X = r) &= \mathrm{P}(X_1 = r \cap X_2 < r) + \mathrm{P}(X_2 = r \cap X_1 < r) + \mathrm{P}(X_1 = X_2 = r) \\ &= \frac{1}{N}.\frac{r-1}{N} + \frac{1}{N}.\frac{r-1}{N} + \frac{1}{N}.\frac{1}{N} \\ &= \frac{2r-1}{N^2}\,. \end{aligned}$$

Alternatively, one can argue as follows:

$$\begin{aligned} \mathrm{P}(X = r) &= \mathrm{P}(X_1 = r \cap X_2 \leqslant r) + \mathrm{P}(X_2 = r \cap X_1 \leqslant r) - \mathrm{P}(X_1 = X_2 = r) \\ &= \frac{1}{N}.\frac{r}{N} + \frac{1}{N}.\frac{r}{N} - \frac{1}{N}.\frac{1}{N} \\ &= \frac{2r-1}{N^2}\,. \end{aligned}$$

(ii) *Find an expression for the expectation,* μ, *of* X *and show that* $\mu = 67.165$ *in the case* $N = 100$.

By the definition of expectation, we have

$$\begin{aligned}
\mu = \mathrm{E}(X) &= \sum_{r=1}^{N} r.\mathrm{P}(X=r) \\
&= \sum_{r=1}^{N} \frac{r(2r-1)}{N^2} \\
&= \frac{1}{N^2}\sum_{r=1}^{N}(2r^2-r) \\
&= \frac{1}{N^2}\big(\tfrac{2}{6}N(N+1)(2N+1) - \tfrac{1}{2}N(N+1)\big) \quad \text{using the given results} \\
&= \frac{\tfrac{1}{6}N(N+1)(4N+2-3)}{N^2} \\
&= \frac{(N+1)(4N-1)}{6N}\,.
\end{aligned}$$

In the case $N=100$, this is $101 \times 399/600 = 13\,433/200 = 67.165$ as required.

> ***(iii)*** *The median, m, of X is defined to be the integer such that $\mathrm{P}(X \geqslant m) \geqslant \frac{1}{2}$ and $\mathrm{P}(X \leqslant m) \geqslant \frac{1}{2}$. Find an expression for m in terms of N and give an explicit value for m in the case $N = 100$.*

We have

$$\begin{aligned}
\mathrm{P}(X \leqslant k) &= \sum_{r=1}^{k} \frac{2r-1}{N^2} \\
&= \frac{1}{N^2}\big(2.\tfrac{1}{2}k(k+1) - k\big) \\
&= \frac{k^2}{N^2}
\end{aligned}$$

so that

$$\mathrm{P}(X \geqslant k) = 1 - \mathrm{P}(X \leqslant k-1) = 1 - \frac{(k-1)^2}{N^2}\,.$$

We are looking for the value of m which makes $\mathrm{P}(X \geqslant m) \geqslant \frac{1}{2}$ and $\mathrm{P}(X \leqslant m) \geqslant \frac{1}{2}$. The first condition gives

$$\begin{aligned}
& 1 - \frac{(m-1)^2}{N^2} \geqslant \tfrac{1}{2} \\
\iff \quad & \frac{(m-1)^2}{N^2} \leqslant \tfrac{1}{2} \\
\iff \quad & (m-1)^2 \leqslant \tfrac{1}{2}N^2 \\
\iff \quad & m-1 \leqslant \tfrac{N}{\sqrt{2}} \\
\iff \quad & m \leqslant \tfrac{N}{\sqrt{2}} + 1.
\end{aligned}$$

The second condition, $\mathrm{P}(X \leqslant m) \geqslant \frac{1}{2}$, yields

$$\begin{aligned} & \frac{m^2}{N^2} \geqslant \tfrac{1}{2} \\ \iff \quad & m^2 \geqslant \tfrac{1}{2}N^2 \\ \iff \quad & m \geqslant \tfrac{N}{\sqrt{2}}\,. \end{aligned}$$

So we have $N/\sqrt{2} \leqslant m \leqslant (N/\sqrt{2}) + 1$, thus m is the smallest integer greater than $N/\sqrt{2}$ (which is not itself an integer as $\sqrt{2}$ is irrational). The smallest integer greater than or equal to a number x is called the *ceiling* of x, and is denoted by $\lceil x \rceil$, so we can state our result as $m = \lceil N/\sqrt{2} \rceil$.

In the case $N = 100$, $m = \lceil 100/\sqrt{2} \rceil = \lceil 70.7\ldots \rceil = 71$.

(iv) *Show that when N is very large,*

$$\frac{\mu}{m} \approx \frac{2\sqrt{2}}{3}\,.$$

We have formulæ for μ from part (ii) and for m from part (iii). Therefore we have, for large N,

$$\begin{aligned} \frac{\mu}{m} &= \frac{(N+1)(4N-1)}{6N} \Big/ \lceil N/\sqrt{2} \rceil \\ &\approx \frac{(N+1)(4N-1)}{6N(N/\sqrt{2})} \\ &= \frac{4N^2+3N-1}{6N^2/\sqrt{2}} \\ &= \frac{4+\frac{3}{N}-\frac{1}{N^2}}{6/\sqrt{2}} \\ &\approx \frac{4}{6/\sqrt{2}} \\ &= \frac{2\sqrt{2}}{3}\,. \end{aligned}$$

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

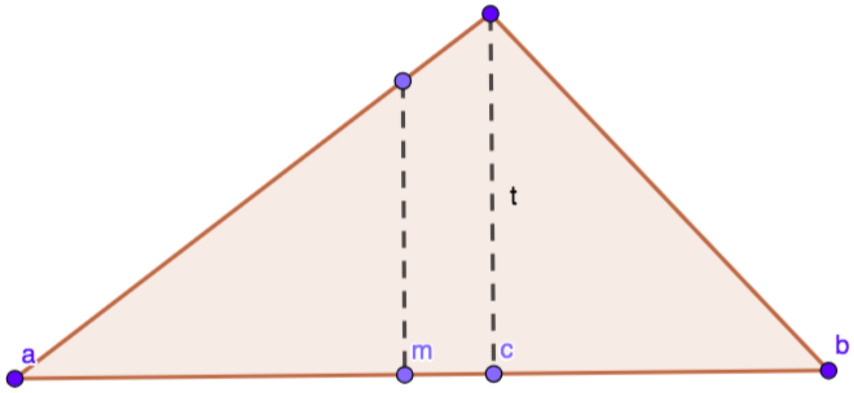
$$f(x) = \begin{cases} g(x) & \text{for } a < x \leqslant c \\ h(x) & \text{for } c \leqslant x < b \\ 0 & \text{otherwise,} \end{cases}$$

1/ 2014/13 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.



$$g(x) = \frac{t(x-a)}{c-a}, \qquad h(x) = \frac{t(b-x)}{b-c}$$

and whole area of triangle $= \frac{1}{2}t(b-a) = 1$ so $t = \frac{2}{b-a}$: $\quad g(x) = \frac{2(x-a)}{(b-a)(c-a)}, \quad h(x) = \frac{2(b-x)}{(b-a)(b-c)}$

The problem of finding the mean of the distribution is exactly the same as the problem of finding the x-coordinate of the centre of mass of the triangle considered as a uniform lamina, so the mean = the x-coordinate of the centroid = $\frac{1}{3}$(sum of x-coordinates of vertices) = $\frac{1}{3}(a+b+c)$

If the left-hand triangle is larger (as shown), so the median m is to the left of c, then area of triangle to left of median = half area of whole triangle, and so:

$$2(m-a)\cdot t\cdot\frac{m-a}{c-a} = t(b-a); \qquad m = a + \sqrt{\frac{1}{2}(c-a)(b-a)}$$

If the right-hand triangle is larger, similarly

$$2(b-m)\cdot t\cdot\frac{b-m}{b-c} = t(b-a); \qquad m = b - \sqrt{\frac{1}{2}(b-c)(b-a)}$$

An internet tester sends n e-mails simultaneously at time $t = 0$. Their arrival times at their destinations are independent random variables each having probability density function $\lambda e^{-\lambda t}$ $(0 \leqslant t < \infty,\ \lambda > 0)$.

(i) The random variable T is the time of arrival of the e-mail that arrives first at its destination. Show that the probability density function of T is

$$n\lambda e^{-n\lambda t},$$

I /2016/13

and find the expected value of T.

(ii) Write down the probability that the second e-mail to arrive at its destination arrives later than time t and hence derive the density function for the time of arrival of the second e-mail. Show that the expected time of arrival of the second e-mail is

$$\frac{1}{\lambda}\left(\frac{1}{n-1} + \frac{1}{n}\right).$$

The probability that any one email has not yet arrived at time s:

$$= 1 - \int_0^s \lambda e^{-\lambda t}\, dt = 1 - \left[-\frac{e^{-\lambda t}}{\lambda}\right]_0^s = e^{-\lambda s}$$

So probability that none of the emails have yet arrived at time s

$= e^{-\lambda s} \cdot e^{-\lambda s} \cdot e^{-\lambda s} \ldots e^{-\lambda s}$ (n terms) $= e^{-n\lambda s}$

So cdf for T = $1 - e^{-n\lambda s}$. By differentiating, pdf for T = $n\lambda e^{-n\lambda s}$

E(T) = $\int\limits_0^\infty sn\lambda e^{-n\lambda s} ds = n\lambda \int\limits_0^\infty s e^{-n\lambda s} = n\lambda\left(\left[\frac{-s}{n\lambda} e^{-n\lambda s}\right]_0^\infty + \int\limits_0^\infty \frac{1}{n\lambda} e^{-n\lambda s} ds\right) = n\lambda\left[\frac{1}{(n\lambda)^2} e^{-n\lambda s}\right]_0^\infty = \frac{1}{n\lambda}$

Probability that second email arrives later than t?

The distribution is "memoryless", i.e. if one email has not arrived by time s, then probability that it will take at least time t longer and will not not arrived by time t+s = $\frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$, independent of s, and similar for any selection of m emails.

Thus probability that second email has not arrived by time t = $e^{-(n-1)\lambda(t-t_1)}$ if t_1 is time of arrival of first email.

Because the distribution is "memoryless", the expected waiting time for the second email after the first has arrived = $\frac{1}{(n-1)\lambda}$

So expected time of arrival of second email = $\frac{1}{\lambda}\left(\frac{1}{n-1} + \frac{1}{n}\right)$

(i) The random variable X has a binomial distribution with parameters n and p, where $n = 16$ and $p = \frac{1}{2}$. Show, using an approximation in terms of the standard normal density function $\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}x^2}$, that

$$\mathrm{P}(X = 8) \approx \frac{1}{2\sqrt{2\pi}}\,.$$

II /2016/13 (ii) By considering a binomial distribution with parameters $2n$ and $\frac{1}{2}$, show that

$$(2n)! \approx \frac{2^{2n}(n!)^2}{\sqrt{n\pi}}\,.$$

(iii) By considering a Poisson distribution with parameter n, show that

$$n! \approx \sqrt{2\pi n}\,\mathrm{e}^{-n}\,n^n\,.$$

The normal approximation for $\mathrm{X} \sim \mathrm{B}(16, \frac{1}{2})$ is $\mathrm{Y} \sim \mathrm{N}(8, 4)$

P(X=8) is approximated by $\mathrm{P}(7.5 < \mathrm{Y} < 8.5)$, i.e. probability in standard normal distribution that $\mathrm{Z} - \mu$ is less than $\frac{1}{4}$ standard deviations. Approximate the pdf of the standard normal distribution in that interval by $\frac{1}{\sqrt{2\pi}}$:

$\mathrm{P}(\mathrm{X} = 8) \approx \frac{1}{2\sqrt{2\pi}}$

The normal approximation for $\mathrm{X} \sim \mathrm{B}(2n, \frac{1}{2})$ is $\mathrm{Y} \sim \mathrm{N}(n, 0.5n)$

$\mathrm{P}(\mathrm{X} = n) = \frac{(2n)!}{2^{2n}(n!)^2}$

Normal approximation for this probability is the probability in a standard normal distribution that $\mathrm{Z} - \mu$ is less than $\sqrt{\frac{1}{2n}}$ standard deviations, so

$\frac{(2n)!}{2^{2n}(n!)^2} \approx \frac{2\sqrt{\frac{1}{2n}}}{\sqrt{2\pi}}$ $\qquad (2n)! \approx \frac{2^{2n}(n!)^2}{\sqrt{n\pi}}$

The normal approximation for $\mathrm{X} \sim \mathrm{Po}(n)$ is $\mathrm{Y} \sim \mathrm{N}(n, n)$

$\mathrm{P}(\mathrm{X} = n) = \mathrm{e}^{-n}\frac{n^n}{n!}$

Normal approximation for this probability is the probability in a standard normal distribution that $\mathrm{Z} - \mu$ is less than $\frac{1}{2\sqrt{n}}$ standard deviations, so

$\mathrm{e}^{-n}\frac{n^n}{n!} \approx \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{n}}$ $\qquad n! \approx \mathrm{e}^{-n}n^n\sqrt{2n\pi}$