2017 GCE A Level H2 FM 9649 Paper 2 Solutions

1 The terms in the sequence $u_0, u_1, u_2, \ldots$ satisfy the recurrence relation

$$u_r = 3u_{r-1} - u_{r-2}.$$

(i) Find the general solution of this recurrence relation. [2]

(ii) Find an expression for u_r in terms of r in the case that $u_0 = 2$ and $u_1 = 3$. [2]

(iii) Find the exact value of u_1 in the case that $u_0 = 2$ and u_r tends to zero as r tends to infinity. [2]

[Solution]

(i) Recurrence relation: $u_r = 3u_{r-1} - u_{r-2}$

Auxiliary equation: $m^2 - 3m + 1 = 0 \Rightarrow m = \dfrac{3 \pm \sqrt{5}}{2}$.

Hence, general solution is: $u_r = A\left(\dfrac{3+\sqrt{5}}{2}\right)^r + B\left(\dfrac{3-\sqrt{5}}{2}\right)^r$, $r = 0, 1, 2, \ldots$

(ii) Using the initial conditions,

$u_0 = 2 \Rightarrow A + B = 2$ ---- (1)

$u_1 = 3 \Rightarrow A\left(\dfrac{3+\sqrt{5}}{2}\right) + B\left(\dfrac{3-\sqrt{5}}{2}\right) = 3$ ---- (2)

$\therefore (2-B)\left(\dfrac{3+\sqrt{5}}{2}\right) + B\left(\dfrac{3-\sqrt{5}}{2}\right) = 3$

$\Rightarrow 3 + \sqrt{5} - \sqrt{5}B = 3$

$\Rightarrow B = 1, \quad A = 1.$

Hence, $u_r = \left(\dfrac{3+\sqrt{5}}{2}\right)^r + \left(\dfrac{3-\sqrt{5}}{2}\right)^r$

(iii) When $u_0 = 2$, from (1) we have $u_r = A\left(\dfrac{3+\sqrt{5}}{2}\right)^r + (2-A)\left(\dfrac{3-\sqrt{5}}{2}\right)^r$

$$= 2\left(\frac{3-\sqrt{5}}{2}\right)^r + A\left[\left(\frac{3+\sqrt{5}}{2}\right)^r - \left(\frac{3-\sqrt{5}}{2}\right)^r\right]$$

$0 < \dfrac{3-\sqrt{5}}{2} < 1$, when $r \to \infty$, $\dfrac{3-\sqrt{5}}{2} \to 0$.

So, when $r \to \infty$, $u_r \to 0 \Rightarrow A = 0$

$\therefore u_r = 2\left(\dfrac{3-\sqrt{5}}{2}\right)^r$ and $u_1 = 3 - \sqrt{5}$.

2 Let

$$C = 1 - \binom{2n}{1}\cos\theta + \binom{2n}{2}\cos 2\theta - \binom{2n}{3}\cos 3\theta + \ldots + \cos 2n\theta,$$

$$S = -\binom{2n}{1}\sin\theta + \binom{2n}{2}\sin 2\theta - \binom{2n}{3}\sin 3\theta + \ldots + \sin 2n\theta,$$

where n is a positive integer.

Show that

$$C = (-4)^n \cos n\theta \sin^{2n} \tfrac{1}{2}\theta,$$

and find the corresponding expression for S. [7]

[Solution]

$$C + iS = 1 - \binom{2n}{1}\cos\theta + \binom{2n}{2}\cos 2\theta - \binom{2n}{3}\cos 3\theta + \cdots + \cos 2n\theta$$

$$+ i\left[-\binom{2n}{1}\sin\theta + \binom{2n}{2}\sin 2\theta - \binom{2n}{3}\sin 3\theta + \cdots + \sin 2n\theta\right]$$

$$C + iS = 1 - \binom{2n}{1}(\cos\theta + i\sin\theta) + \binom{2n}{2}(\cos 2\theta + i\sin 2\theta) - \cdots + (\cos 2n\theta + i\sin 2n\theta)$$

$$C + iS = 1 - \binom{2n}{1}\mathrm{e}^{i\theta} + \binom{2n}{2}\mathrm{e}^{i2\theta} - \binom{2n}{3}\mathrm{e}^{i3\theta} + \cdots + \mathrm{e}^{i2n\theta}$$

Compare with: $(1-x)^n = 1 - \binom{n}{1}x + \binom{n}{2}x^2 - \binom{n}{3}x^3 + \cdots + \binom{n}{n}x^n$ for $n \in \mathbb{Z}^+$

$$\Rightarrow \quad C + iS = \left(1 - e^{i\theta}\right)^{2n}$$

$$= \left[e^{i\frac{\theta}{2}}\left(e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}\right)\right]^{2n}$$

$$= e^{in\theta}\left[\left(\cos\frac{\theta}{2} - i\sin\frac{\theta}{2}\right) - \left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right)\right]^{2n}$$

$$= e^{in\theta}\left(-2i\sin\frac{\theta}{2}\right)^{2n}$$

$$= e^{in\theta}\left[(-2)^2\right]^n \left(i^2\right)^n \sin^{2n}\frac{\theta}{2}$$

$$= (\cos n\theta + i\sin n\theta)(-4)^n \sin^{2n}\frac{\theta}{2}$$

Comparing real parts: $C = (-4)^n \cos n\theta \sin^{2n}\dfrac{\theta}{2}$

Comparing imaginal parts: $S = (-4)^n \sin^{2n}\dfrac{\theta}{2}\sin n\theta$

3 An arc of a parabola has parametric equations $x = at^2$, $y = 2at$ for $0 \leqslant t \leqslant p$.

(i) Find, in terms of a and p, the area of the surface of revolution formed when this arc is rotated about the x-axis through 2π radians. [5]

(ii) Find an approximation to this surface area when p is sufficiently small for powers above p^2 to be ignored. Interpret this approximate area geometrically. [3]

[Solution]

$x = at^2, \quad y = 2at, \quad 0 \le t \le p$

(i) Curve surface area

$$= \int_0^p 2\pi y \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

$$= \int_0^p 2\pi(2at)\sqrt{(2at)^2 + (2a)^2}\, dt$$

$$= 4\pi a^2 \int_0^p 2t\sqrt{t^2+1}\, dt$$

$$= 4\pi a^2 \left[\frac{(t^2+1)^{3/2}}{3/2}\right]_0^p$$

$$= \frac{8}{3}\pi a^2\left[(p^2+1)^{3/2} - 1\right]$$

(ii) When p is small,

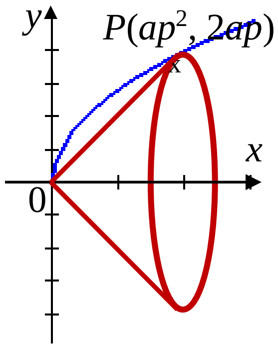
Curve surface area

$= \pi rl$

$= \frac{8}{3}\pi a^2\left[(1+p^2)^{3/2} - 1\right]$

$\approx \frac{8}{3}\pi a^2\left[1 + \frac{3}{2}p^2 - 1\right]$

$= 4\pi a^2 p^2$



When p is small, the arc of the parabola is almost a line segment joining (0, 0) to $(ap^2, 2ap)$, so the approximate surface area $4\pi a^2 p^2$ is the surface area of the cone generated when this line segment is rotated about the x-axis

$= \pi(2ap)\sqrt{(ap^2)^2 + (2ap)^2}$

$= 2\pi a^2 p^2 (p^2+4)^{\frac{1}{2}} = 4\pi a^2 p^2\left(1 + \frac{p^2}{4}\right)^{\frac{1}{2}} \approx 4\pi a^2 p^2$

4 The matrix **M** is defined as follows.

$$\mathbf{M} = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 3 & 1 & 2 \\ 4 & 3 & 1 \end{pmatrix}.$$

(i) Find the number of independent rows of **M** and deduce the number of independent columns of **M**. [4]

(ii) The matrix **M** defines a transformation from $\mathbb{R}^3$ to $\mathbb{R}^4$. Find a basis for

(a) the null space of **M**,

(b) the range space of **M**.

[4]

[Solution]

(i) $$\mathbf{M} = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 3 & 1 & 2 \\ 4 & 3 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

rank of **M** = 2

$\Rightarrow$ number of independent rows of **M** = 2

$\Rightarrow$ number of independent columns of **M** = 2

(ii) (a) $$\begin{cases} 2x + y + z = 0 \\ y - z = 0 \end{cases}$$

$$\Rightarrow \quad x = -z, \quad y = z$$

$\therefore$ A basis for the null space of **M** = $\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

(b) Since the column space = range space, and the first two columns are linearly independent,

$\therefore$ a basis for the range space of **M** = $\left\{ \begin{pmatrix} 2 \\ 3 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \end{pmatrix} \right\}$.

5 Solve the differential equation

$$\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 2e^{-2x}$$

given that when $x = 0$, $y = 0$ and $\frac{dy}{dx} = 5$.

Sketch the solution curve for $x \geqslant 0$. Confirm, by considering $\frac{dy}{dx}$, that the curve has only one turning point. [12]

[Solution]

$$\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 2e^{-2x}$$

Auxiliary equation: $m^2 + 5m + 6 = 0 \Rightarrow m = -2$ or -3

Complimentary Function: $y = Ae^{-2x} + Be^{-3x}$ where A and B are arbitrary constants.

Let particular integral be $y = pxe^{-2x}$

$$\frac{dy}{dx} = pe^{-2x} - 2pxe^{-2x}$$

$$\frac{d^2y}{dx^2} = -4pe^{-2x} + 4pxe^{-2x}$$

Substitute into DE: $-4pe^{-2x} + 4pxe^{-2x} + 5pe^{-2x} - 10pxe^{-2x} + 6pxe^{-2x} = 2e^{-2x}$

$$pe^{-2x} = 2e^{-2x}$$

$$\therefore p = 2$$

General Solution is $y = Ae^{-2x} + Be^{-3x} + 2xe^{-2x}$

$$\frac{dy}{dx} = -2Ae^{-2x} - 3Be^{-3x} + 2e^{-2x} - 4xe^{-2x}$$

When $x = 0, y = 0$, $A + B = 0$ ---- (1)

When $x = 0, \frac{dy}{dx} = 5$, $-2A - 3B + 2 = 5$

$2A + 3B = -3$ ----- (2)

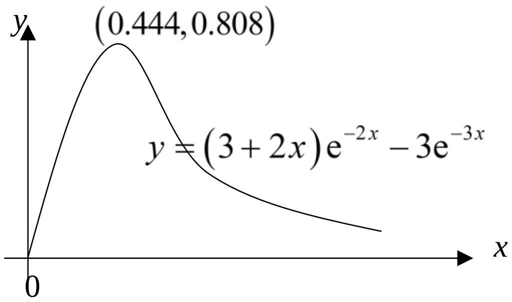
Solving (1) and (2): $A = 3, B = -3$

Particular solution is $y = (3+2x)e^{-2x} - 3e^{-3x}$

$$\frac{dy}{dx} = 9e^{-3x} - 4e^{-2x} - 4xe^{-2x}$$

When $\frac{dy}{dx} = 0$, $9e^{-3x} - 4e^{-2x} - 4xe^{-2x} = 0$

$$e^{-x} = \frac{4}{9}(x+1)$$



Since $y = e^{-x}$ is a strictly decreasing function and $y = \frac{4}{9}(x+1)$ is a strictly increasing function, therefore the two graphs $y = e^{-x}$ and $y = \frac{4}{9}(x+1)$ intersect only once. Hence $\frac{dy}{dx} = 0$ only has one real solution. Thus, $(0.444, 0.808)$ is the only turning point on the graph of $y = Ae^{-2x} + Be^{-3x} + 2xe^{-2x}$.

6 **(i)** Show that the matrix **A**, where

$$\mathbf{A} = \begin{pmatrix} 0.7 & 0.6 & 0.4 \\ 0.2 & 0 & 0.2 \\ 0.1 & 0.4 & 0.4 \end{pmatrix},$$

has an eigenvalue equal to 1. Find a corresponding eigenvector. [3]

(ii) Show that $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix}$ are also eigenvectors of **A**, and determine corresponding eigenvalues.

Hence write down a matrix **Q** and a diagonal matrix **D** such that $\mathbf{A} = \mathbf{QDQ}^{-1}$.

Explain briefly how this form for **A** may be used to find powers of **A**. [6]

[Solution]

(i) Consider $\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 0.7-\lambda & 0.6 & 0.4 \\ 0.2 & -\lambda & 0.2 \\ 0.1 & 0.4 & 0.4-\lambda \end{pmatrix}$

When $\lambda = 1$,

$$\det(A-\lambda I) = -0.3\left[-(0.4-1)-0.08\right]-0.6\left[(0.2)(0.4-1)-0.02\right]+0.4\left[0.08-(-0.1)\right]$$
$$= 0$$

$\therefore$ **A** has an eigenvalue equal to 1.

And $A - \lambda I = \begin{pmatrix} -0.3 & 0.6 & 0.4 \\ 0.2 & -1 & 0.2 \\ 0.1 & 0.4 & -0.6 \end{pmatrix}$

Let $\mathbf{x} = \begin{pmatrix} 0.2 \\ -1 \\ 0.2 \end{pmatrix} \times \begin{pmatrix} 0.1 \\ 0.4 \\ -0.6 \end{pmatrix} = \begin{pmatrix} 0.52 \\ 0.14 \\ 0.18 \end{pmatrix}$, we have $\begin{pmatrix} -0.3 & 0.6 & 0.4 \\ 0.2 & -1 & 0.2 \\ 0.1 & 0.4 & -0.6 \end{pmatrix}\begin{pmatrix} 0.52 \\ 0.14 \\ 0.18 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

Therefore, an eigenvector of **A** corresponding to eigenvalue 1 is $\begin{pmatrix} 0.52 \\ 0.14 \\ 0.18 \end{pmatrix}$.

(ii) $(\mathbf{A}-\lambda\mathbf{I})\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.7-\lambda & 0.6 & 0.4 \\ 0.2 & -\lambda & 0.2 \\ 0.1 & 0.4 & 0.4-\lambda \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.3-\lambda \\ 0 \\ -0.3+\lambda \end{pmatrix}$

$\therefore \lambda = 0.3 \iff (\mathbf{A}-\lambda\mathbf{I})\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

Therefore, an eigenvector of **A** corresponding to eigenvalue 0.3 is $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

Similarly,

$$(\mathbf{A}-\lambda\mathbf{I})\begin{pmatrix}2\\-5\\3\end{pmatrix}=\begin{pmatrix}0.7-\lambda & 0.6 & 0.4\\ 0.2 & -\lambda & 0.2\\ 0.1 & 0.4 & 0.4-\lambda\end{pmatrix}\begin{pmatrix}2\\-5\\3\end{pmatrix}=\begin{pmatrix}-2\lambda-0.4\\ 5\lambda+1\\ -0.6-3\lambda\end{pmatrix}$$

$$\therefore\ \lambda=-0.2 \iff (\mathbf{A}-\lambda\mathbf{I})\begin{pmatrix}2\\-5\\3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Therefore, an eigenvector of **A** corresponding to eigenvalue -0.2 is $\begin{pmatrix}2\\-5\\3\end{pmatrix}$.

$\mathbf{A}=\mathbf{QDQ}^{-1}$

$$\mathbf{A}=\begin{pmatrix}0.7 & 0.6 & 0.4\\ 0.2 & 0 & 0.2\\ 0.1 & 0.4 & 0.4\end{pmatrix},\ \text{let}\ \ \mathbf{D}=\begin{pmatrix}-0.2 & 0 & 0\\ 0 & 0.3 & 0\\ 0 & 0 & 1\end{pmatrix},\ \text{and}\ \ \mathbf{Q}=\begin{pmatrix}2 & 1 & 0.52\\ -5 & 0 & 0.14\\ 3 & -1 & 0.18\end{pmatrix}$$

$$\mathbf{AQ}=\begin{pmatrix}0.7 & 0.6 & 0.4\\ 0.2 & 0 & 0.2\\ 0.1 & 0.4 & 0.4\end{pmatrix}\begin{pmatrix}2 & 1 & 0.52\\ -5 & 0 & 0.14\\ 3 & -1 & 0.18\end{pmatrix}=\begin{pmatrix}-0.4 & 0.3 & 0.52\\ 1 & 0 & 0.14\\ -0.6 & -0.3 & 0.18\end{pmatrix}$$

$$\mathbf{QD}=\begin{pmatrix}2 & 1 & 0.52\\ -5 & 0 & 0.14\\ 3 & -1 & 0.18\end{pmatrix}\begin{pmatrix}-0.2 & 0 & 0\\ 0 & 0.3 & 0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}-0.4 & 0.3 & 0.52\\ 1 & 0 & 0.14\\ -0.6 & -0.3 & 0.18\end{pmatrix}$$

Thus $\mathbf{AQ}=\mathbf{QD} \ \Rightarrow\ \mathbf{A}=\mathbf{QDQ}^{-1}$

If $\mathbf{A}=\mathbf{QDQ}^{-1}$,

$\mathbf{A}^2=(\mathbf{QDQ}^{-1})(\mathbf{QDQ}^{-1})=\mathbf{QD}(\mathbf{Q}^{-1}\mathbf{Q})\mathbf{DQ}^{-1}=\mathbf{QD}(\mathbf{I})\mathbf{DQ}^{-1}=\mathbf{QD}^2\mathbf{Q}^{-1}$

$\mathbf{A}^3=(\mathbf{QD}^2\mathbf{Q}^{-1})(\mathbf{QDQ}^{-1})=\mathbf{QD}^3\mathbf{Q}^{-1}$

$\vdots$

$\mathbf{A}^n=\mathbf{QD}^n\mathbf{Q}^{-1}$ for $n=0, 1, 2, \dots$

When $n=-1$, $\mathbf{A}^{-1}=(\mathbf{QDQ}^{-1})^{-1}=(\mathbf{Q}^{-1})^{-1}\mathbf{D}^{-1}\mathbf{Q}^{-1}=\mathbf{QD}^{-1}\mathbf{Q}^{-1}$

Thus $\mathbf{A}^n=\mathbf{QD}^n\mathbf{Q}^{-1}$ for $n=-1, 0, 1, 2, \dots$

$$=\mathbf{Q}\begin{pmatrix}1 & 0 & 0\\ 0 & 0.3^n & 0\\ 0 & 0 & (-0.2)^n\end{pmatrix}\mathbf{Q}^{-1}$$

Section B: Probability and Statistics [50 marks]

7 A medical trial is carried out to test the effectiveness of a new tablet intended to be used for headache relief. A random sample of 30 patients are given paracetamol tablets and a random sample of 26 patients are given the new tablets. Each patient is asked to record the time from taking the tablet to gaining pain relief. The data are as follows.

	Tablet	Sample size	Sum of times (minutes)	Sum of squares of times (minutes^2)
X	Paracetamol	30	$\Sigma x = 1609$	$\Sigma x^2 = 92\,314$
Y	New	26	$\Sigma y = 1242$	$\Sigma y^2 = 67\,173$

Carry out an appropriate test to determine whether there is evidence, at the 5% level of significance, that the new tablet acts more quickly than paracetamol.

State the assumptions you have made in carrying out the test. [8]

[Solution]

Let X and Y be the respective time from taking the paracetamol tablets and the new tablet till gaining pain relief.

Let μ_x and μ_y be the population mean time of X and Y respectively.

Assumptions: X and Y follow two independent normal distribution with a common variance.

$$n_1 = 30,\ \bar{x} = \frac{1609}{30} = 53.633,\ s_x^2 = \frac{1}{30-1}\left[92314 - \frac{1609^2}{30}\right] = 207.5161$$

$$n_2 = 26,\ \bar{y} = \frac{1242}{26} = 47.76923,\ s_y^2 = \frac{1}{26-1}\left[67173 - \frac{1242^2}{26}\right] = 313.7446$$

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2} = 16.0217^2$$

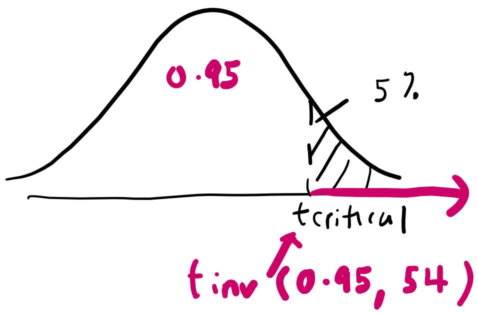
H$_0$: $\mu_y - \mu_x = 0$

H$_1$: $\mu_x - \mu_y > 0$

Level of significance: 5%

Under H$_0$,

Test statistic: $\dfrac{\bar{X} - \bar{Y} - 0}{S_p\sqrt{\dfrac{1}{30} + \dfrac{1}{26}}} \sim t_{30+26-2} = t_{54}$



From GC, $t_{cal} = 1.366$

p-value = 0.0888 > 0.05

Since p-value > level of significance, we do not reject H$_0$.

There is insufficient evidence at 5% level of significance to conclude that the new tablet acts more quickly than paracetamol.

8 It is thought that a set of data, denoted by x, might have arisen from a probability distribution for which

$$P(X > x) = (1 + x)e^{-x}, \text{ where } x \geqslant 0.$$

The data, 100 values in total, are as follows.

Range	$0 \leqslant x < 1$	$1 \leqslant x < 2$	$2 \leqslant x < 3$	$3 \leqslant x < 4$	$4 \leqslant x$
Frequency	21	25	17	16	21

Carry out an appropriate test, using a 1% significance level, to assess whether or not it is reasonable to suppose that the data are drawn from the specified distribution. [8]

[Solution]

H_0: The data are ~~drawn~~ consistent with the specified distribution with

$$\mathrm{P}(X > x) = (1+x)e^{-x}, \quad x \geq 0.$$

H_1: The data are not ~~drawn~~ consistent with the above specified distribution.
Level of significance: 1%

Based on H_0, the expected frequencies are shown below:

$E_1 = 100\mathrm{P}(0 \leq X < 1) = 100[\mathrm{P}(X \geq 0) - \mathrm{P}(X \geq 1)] = 100(e^{-0} - 2e^{-1}) = 26.424$

$E_2 = 100[\mathrm{P}(X \geq 1) - \mathrm{P}(X \geq 2)] = 100(2e^{-1} - 3e^{-2}) = 32.975$

$E_3 = 100[\mathrm{P}(X \geq 2) - \mathrm{P}(X \geq 3)] = 100(3e^{-2} - 4e^{-3}) = 20.686$

$E_4 = 100[\mathrm{P}(X \geq 3) - \mathrm{P}(X \geq 4)] = 100(4e^{-3} - 5e^{-4}) = 10.757$

$E_5 = 100 - 26.424 - 32.975 - 20.686 - 10.757 = 9.158$

The only constraint is: $\sum O_i = \sum E_i = 100$,

Degrees of freedom $= 5 - 1 = 4$

Test statistic: $\sum \frac{(O_i - E_i)^2}{E_i} \sim \chi_4^2$

From GC, we have $\chi^2_{cal} = 21.567$. very strong evidence

p- value $= 0.000244 < 0.01$

We reject H_0, there is sufficient evidence at the 1% level of significance to claim that the data are not consistent with the above specified distribution.

9 A laser aimed at a target misses by a distance X which is modelled as a random variable with the probability density function $\mathrm{f}(x)$ defined as follows.

$$\mathrm{f}(x) = \begin{cases} \dfrac{k}{1+x^2} & \text{for } 0 \leqslant x \leqslant a, \\ 0 & \text{otherwise,} \end{cases}$$

where a is a parameter which determines the accuracy with which the laser is aimed and k is a constant.

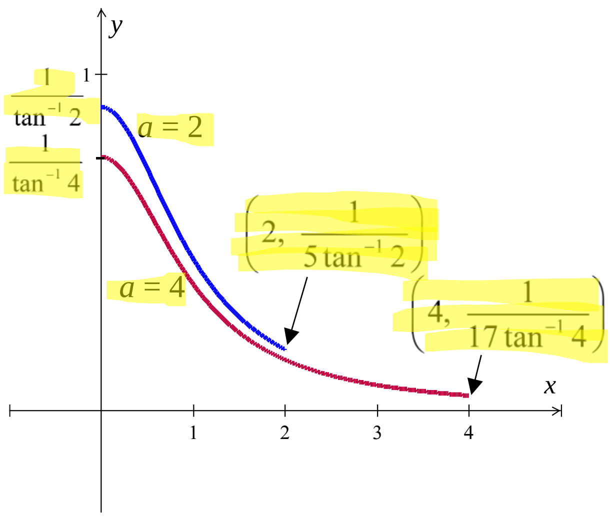
(i) Obtain an expression for k in terms of a. Sketch, on the same axes, the probability density function for $a = 2$ and for $a = 4$. [5]

(ii) Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ in terms of a. [5]

(iii) Find, in terms of a, the probability that the laser misses the target by a distance less than $\frac{1}{2}a$. Give the approximate value of this probability when a is small. State what happens to this probability when a is very large. [3]

[Solution]

(i) $$\int_0^a \frac{k}{1+x^2}\,\mathrm{d}x = 1 \Rightarrow k\left[\tan^{-1} x\right]_0^a = 1 \Rightarrow k\tan^{-1} a = 1 \Rightarrow k = \frac{1}{\tan^{-1} a}$$



(ii) $$\mathrm{E}(X) = \int_0^a x\mathrm{f}(x)\,\mathrm{d}x = \frac{1}{2\tan^{-1} a}\int_0^a \left(\frac{2x}{1+x^2}\right)\mathrm{d}x = \frac{1}{2\tan^{-1} a}\left[\ln\left(1+x^2\right)\right]_0^a = \frac{\ln\left(1+a^2\right)}{2\tan^{-1} a}$$

$$\mathrm{E}\left(X^2\right) = \int_0^a x^2\mathrm{f}(x)\,\mathrm{d}x = \frac{1}{\tan^{-1} a}\int_0^a \left(1 - \frac{1}{1+x^2}\right)\mathrm{d}x = \frac{1}{\tan^{-1} a}\left[x - \tan^{-1} x\right]_0^a = \frac{a}{\tan^{-1} a} - 1$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \left[\mathrm{E}(X)\right]^2 = \frac{a}{\tan^{-1} a} - 1 - \left[\frac{\ln\left(1+a^2\right)}{2\tan^{-1} a}\right]^2$$

(iii) $$\mathrm{P}\left(X < \frac{1}{2}a\right) = \int_0^{\frac{1}{2}a} \frac{1}{\left(1+x^2\right)\tan^{-1} a}\,\mathrm{d}x = \frac{1}{\tan^{-1} a}\left[\tan^{-1} x\right]_0^{\frac{1}{2}a} = \frac{\tan^{-1}\frac{1}{2}a}{\tan^{-1} a}$$

$\tan\alpha \approx \alpha$

$\alpha \approx \tan^{-1}\alpha$

$\approx \frac{1}{2}a$

When a is small, $$\mathrm{P}\left(X < \frac{1}{2}a\right) = \lim_{a\to 0}\left(\frac{\tan^{-1}\frac{1}{2}a}{\tan^{-1}a}\right) = \lim_{a\to 0}\frac{\frac{\mathrm{d}}{\mathrm{d}a}\left(\tan^{-1}\frac{1}{2}a\right)}{\frac{\mathrm{d}}{\mathrm{d}a}\left(\tan^{-1}a\right)} = \lim_{a\to 0}\frac{\frac{\frac{1}{2}}{1+\left(\frac{1}{2}a\right)^2}}{\frac{1}{1+a^2}} = \frac{1}{2}$$

$\approx a$

L'Hopital Rule

When a is very large, $\mathrm{P}\left(X < \frac{1}{2}a\right) \approx$ Total probability = 1.

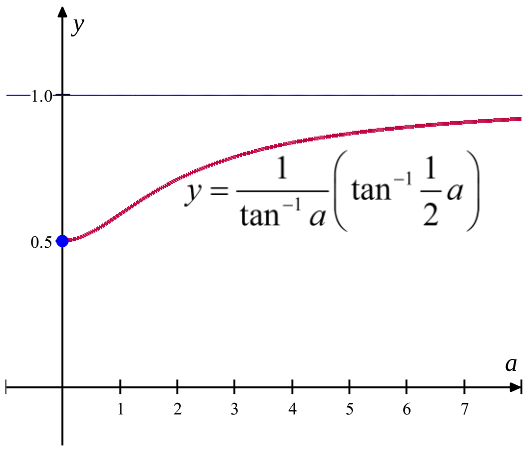
<u>Alternative Method</u>

From the graph of $y = \frac{1}{\tan^{-1}a}\left(\tan^{-1}\frac{1}{2}a\right)$, we see that as $a \to 0$, $y \to \frac{1}{2}$.

$\therefore \mathrm{P}\left(X < \frac{1}{2}a\right) \approx \frac{1}{2}$

When $a \to \infty$, $y \to 1$.

$\therefore \mathrm{P}\left(X < \frac{1}{2}a\right) \approx 1$



10 A company supplying home broadband services estimates that about 8% of its customers will not be happy with their download speed. The company calls randomly chosen customers in sequence and asks them if they are happy with their download speed.

(i) Assuming the company's estimate is correct, state the distribution of the number of customers the company will have to contact in order to find one who is not happy with their download speed. State the mean and standard deviation of this distribution, and find the probability that the number of customers contacted exceeds 25. [3]

(ii) The company wishes to check how accurate the estimate of 8% is. A worker is instructed to call customers at random until *either* the first unhappy customer is found *or* 25 customers have been contacted.

Assuming that the proportion of unhappy customers is p, find the probability that exactly 25 customers are contacted. [2]

(iii) The procedure in part **(ii)** is carried out repeatedly. Records of workers' phone calls show that, in the long run, fewer than 25 customers are contacted on 89.5% of occasions. Use this information to estimate the value of p. [2]

(iv) The median download speed provided by the company is claimed to be 75 Mb/s, but a consumer group believes the true figure to be lower than this. In order to check, a random sample of 50 customers are contacted and asked to measure and report their download speeds. A sign test is then carried out, using a 1% level of significance. Suppose that k customers report download speeds of less than 75 Mb/s.

State the null and alternative hypotheses for this test. Determine the possible values of k if the null hypothesis is rejected. [5]

[Solution]

(i) Let X be the number of customers the company will have to contact until one who is not happy with their download speed.

$X \sim \text{Geo}(0.08)$

$$\mathrm{E}(X) = \frac{1}{0.08} = 12.5$$

$$\mathrm{Var}(X) = \frac{1-0.08}{0.08^2} = 143.75$$

Standard deviation of $X = \sqrt{143.75} = 11.9896 \approx 12.0$

$\mathrm{P}(X > 25) = \mathrm{P}(\text{the first 25 customers contacted are all happy with the download speed})$

$= (1-0.08)^{25}$ or $P(X>25) = 1-P(X\le 24)$ use GC

$= 0.12436 \approx 0.124$

(ii) p = P(a random chosen customer is unhappy)

P(exactly 25 customers are contacted)

= P(the first 24 customers are all happy with the download speed)

$= (1-p)^{24}$ $(1-p)^{25}$ $(1-p)^{24}p$

(iii) P(fewer than 25 customers are contacted) = 0.895

$\Rightarrow 1 - \text{P(exactly 25 customers are contacted)} = 0.895$

$\Rightarrow 1-(1-p)^{24} = 0.895$

$\Rightarrow p = 1 - 0.105^{\frac{1}{24}} \approx 0.0896$

(iv) Let M Mb/s be the population median download speed. Not sample median!

$H_0: \ M = 75$

$H_1: \ M < 75$

Level of significance: 1%

Let S_+ be the number of customers report download speeds > 75 Mb/s

S_- be the number of customers report download speeds < 75 Mb/s

Test Statistic: $S = S_+ \sim B\left(50, \frac{1}{2}\right)$ if H_0 is true.

Computation: $S_- = k$, $S_+ = 50 - k$.

s	$P(S \le s)$
15	0.00330 < 0.01
16	0.00767 < 0.01
17	0.01642 > 0.01

Since $P(S \le 16) = 0.00767 < 0.01$

and $P(S \le 17) = 0.0164 > 0.01$

At 1% level of significance,

critical region = { S: $S \le 16$ }

observed value of $S = 50 - k$

For H_0 to be rejected, $50 - k \le 16$

$\Rightarrow k \ge 34$

So possible values of k are $34 \le k \le 50$, $k \in \mathbb{Z}^+$.

11 **(i)** In order to reduce levels of atmospheric pollution, a country's Ministry of Transport proposes setting a new limit for pollutants emitted by car exhausts. The Ministry sets up a checkpoint and stops cars at random for testing. In one week, 435 cars are tested and it is found that 133 of them generate more than the proposed new limit for pollutants.

Construct a 95% confidence interval for the proportion of cars that generate more than the proposed new limit for pollutants. [2]

(ii) The Ministry gives considerable publicity to the data over the next few weeks, and it recommends that car owners have their cars serviced to reduce pollutants. The Ministry then carries out another survey, in which n cars are tested and k of them are found to be generating more than the proposed new limit for pollutants. The 95% confidence interval constructed from these data is

(0.227, 0.305).

Find the values of k and n. [5]

(iii) The Ministry claims that the publicity drive has been successful. Discuss briefly whether or not the figures support this view. [2]

[Solution]

(i) Let p be the population proportion of cars that generate more than the proposed new limit for pollutants

$n = 435, \quad p_s = \frac{133}{435} = 0.30574$

$p_s = \frac{k}{n}$

A 95% confidence interval for p is given by

$$p_s - 1.9600\sqrt{\frac{p_s(1-p_s)}{n}} < p < p_s + 1.9600\sqrt{\frac{p_s(1-p_s)}{n}}$$

$\Rightarrow 0.26245 < p < 0.34904$

$\Rightarrow 0.262 < p < 0.349$ ← use GC to check ans!

(ii) After the publicity drive, $p_s = \dfrac{k}{n}$.

Let p_1 be the new population proportion after the publicity.
A 95% confidence interval for p_1 is given by

$$p_s - 1.95996\sqrt{\frac{p_s(1-p_s)}{n}} < p_1 < p_s + 1.95996\sqrt{\frac{p_s(1-p_s)}{n}}$$

Since the 95% confidence interval for p_1 is (0.227, 0.305),

$$p_s = \frac{0.227+0.305}{2} = 0.266$$

$$\therefore \quad \frac{k}{n} = 0.266 \quad \Rightarrow \quad k = 0.266n$$

Also, $0.266 + 1.95996\sqrt{\dfrac{0.266(1-0.266)}{n}} = 0.305$

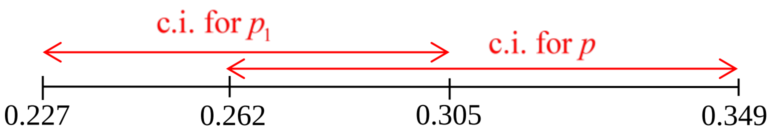
$$\Rightarrow \quad n = 493.11 \approx 493$$

$$k = 0.226n$$

$$\approx 131$$

(iii) Before publicity, $0.262 < p < 0.349$

After publicity, $0.227 < p_1 < 0.305$



The two confidence intervals overlap, so it is possible that $p_1 = 0.29$ (say) and $p = 0.27$ (say) so that $p_1 > p$, this suggests that the proportion of cars that generate more than the new limit for pollutants after the publicity is greater that the proportion before the publicity. Hence the figures may not necessary support the claim that the publicity drive has been successful.