IYGB - FS1 PAPER O - QUESTION 1

$$P(X=x) = \begin{cases} kx^2 & x = 3,4,5 \\ 0 & \text{OTHERWISE} \end{cases}$$

a) WRITE THE FORMULA AS A TABLE

x	3	4	5
$P(X=x)$	$9k$	$16k$	$25k$

$$9k + 16k + 25k = 1$$
$$50k = 1$$
$$k = \frac{1}{50}$$

b)I) $E(X) = \sum x P(X=x)$

$$\Rightarrow E(X) = (3 \times 9k) + (4 \times 16k) + (5 \times 25k)$$
$$\Rightarrow E(X) = 27k + 64k + 125k$$
$$\Rightarrow E(X) = 216k$$
$$\Rightarrow E(X) = 4.32$$

II) $E(X^2) = \sum x^2 P(X=x)$

$$\Rightarrow E(X^2) = (3^2 \times 9k) + (4^2 \times 16k) + (5^2 \times 25k)$$
$$\Rightarrow E(X^2) = 81k + 256k + 625k$$
$$\Rightarrow E(X^2) = 962k$$
$$\Rightarrow E(X^2) = 19.24$$

IYGB – FS1 PAPER O – QUESTION 1

$$\underline{\mathrm{Var}(X) = E(X^2) - [E(X)]^2}$$

$$\Longrightarrow \mathrm{Var}(X) = 19.24 - 4.32^2$$

$$\Longrightarrow \mathrm{Var}(X) = 0.5776$$

c) I) $E(5X-4) = 5E(X) - 4$

$= 5 \times 4.32 - 4$

$= 17.6$

II) $\mathrm{Var}(5X-4) = 5^2\,\mathrm{Var}(X)$

$= 25 \times 0.5776$

$= 14.44$

IYGB - FP PAPER O - QUESTION 2

a) START BY REWRITING THE TABLE IN RANKS

LAP TIME RANK	4	6	3	1	7	5	8	2
FINISH ORDER RANK	5	6	1	3	7	4	8	2
d^2	1	0	4	4	0	1	0	0

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6 \times 10}{8 \times 63} = 1 - \frac{5}{42} = \frac{37}{42} \simeq 0.8810$$

b) $\boxed{\begin{array}{ll} H_0: \rho_s = 0 & \text{(NO ASSOCIATION, "POSITIVE" OR "NEGATIVE")} \\ H_1: \rho_s \neq 0 & \text{(ASSOCIATION EXISTS)} \end{array}}$

THE CRITICAL VALUE FOR $n=8$, AT 5%, TWO TAILED, IS ± 0.7381

AS $0.8810 > 0.7381$, THERE IS SIGNIFICANT EVIDENCE OF (POSITIVE) ASSOCIATION BETWEEN THE FASTEST QUALIFYING LAP TIME AND THE RACE FINISHING POSITION. — REJECT H_0

IYGB – FS1 PAPER O – QUESTION 3

SETTING SUITABLE HYPOTHESES

- H_0 : THERE IS NO ASSOCIATION BETWEEN THE EDUCATION LEVEL AND THE EVENTUAL AVERAGE EARNINGS (INDEPENDENCE)
- H_0 : THERE IS ASSOCIATION BETWEEN THE EDUCATION LEVEL AND THE EVENTUAL AVERAGE EARNINGS (NOT INDEPENDENT)

EARNINGS	EDUCATION LEVEL: NON GRADUATES	GRADUATES	POST GRADUATES	TOTAL
UP TO £10000	17 16.56 0.0117	~~3~~6 6.44 0.0301	~~3~~	23
£10001 TO £25000	97 83.52 2.1757	~~16~~19 32.48 5.5945	~~3~~	116
£25001 TO £40000	42 51.12 1.6270	~~21~~29 19.6 4.5081	~~8~~	71
OVER £40000	24 28.8 0.8000	~~10~~16 11.2 2.0571	~~6~~	40
TOTAL	180	~~50~~70	~~20~~	250

FIRSTLY LOOKING AT SOME OF THE FREQUENCIES, WE SUSPECT THAT THE EXPECTED FREQUENCIES MIGHT FALL BELOW 5

- CHECK "POST GRADS UNDER £10000" : $\frac{23 \times 20}{250} = 1.84 < 5$
- COMBINE THE LAST TWO COLUMNS OF THE TABLE

COMPUTE EXPECTED FREQUENCIES & CONTRIBUTIONS

- ● (black) : OBSERVED FREQUENCIES (ACTUAL DATA), O_i
- ● (red) : EXPECTED FREQUENCIES FOR INDEPENDENCE, E_i
- ● (blue) : CONTRIBUTIONS, $\frac{(O_i - E_i)^2}{E_i}$

IYGB - FS1 PAPER O - QUESTION 3

SUMMARIZING ALL RESULTS

$$\nu = 3$$

$$\chi^2_3(1\%) = 11.345$$

$$\sum_{i=1}^{8} \frac{(O_i - E_i)^2}{E_i} = 18.861$$

AS $18.861 > 11.345$ THERE IS SIGNIFICANT EVIDENCE OF ASSOCIATION (DEPENDENCE) BETWEEN THE LEVEL OF EDUCATION AND THE EXPECTED AVERAGE ANNUAL EARNINGS

THERE IS SUFFICIENT EVIDENCE TO REJECT H_0

IYGB – FS1 PAPER O – QUESTION 4

- AS THE TEST IS TWO TAILED, THE SIGNIFICANCE MUST BE 2.5% IN EACH TAIL

- LOOKING AT THE POISSON TABLES (ON PAPER OR CALCULATOR)

  - "BOTTOM END"

    $W = 0, 1$ THIS OCCURS FOR INTEGER VALUES OF λ IF $k = 6, 7$

  - "TOP END"

    $W = 12, 13, 14, \ldots$ THIS OCCURS FOR INTEGER VALUES OF λ IF $k = 6$ (ONLY)

- HENCE $k = 6$ AND BY USING TABLES STILL

$$P(W \leq 1) = \ldots \text{tables} = 0.0174 = 1.74\%$$

$$P(W \geq 12) = \ldots \text{tables} = 1 - P(W \leq 11) = 1 - 0.9799$$

$$= 0.0201 = 2.01\%$$

- ACTUAL SIGNIFICANCE IS $1.74\% + 2.01\% = 3.75\%$

IYGB - FS1 PAPER O - QUESTION 5

a) START BY DEFINING DISTRIBUTIONS

I) X = no of tube fails per weekday
$X \sim Po(1)$

II) Y = no of tube fails per weekend
$Y \sim Po(0.5)$

$$P(X=4) = \frac{e^{-1} \times 1^4}{4!} = \underline{\underline{0.0153}}$$

$$P(Y>2) = P(Y \geq 3)$$
$$= 1 - P(Y \leq 2)$$
... tables ...
$$= 1 - 0.9856$$
$$= \underline{\underline{0.0144}}$$

III)
$X+Y \sim Po(5 \times 1 + 0.5)$
$X+Y \sim Po(5.5)$

$$P(X+Y<4) = P(X+Y \leq 3) = \text{... tables ...} = \underline{\underline{0.2017}}$$

b) USING $X+Y \sim Po(5.5)$

$$\Rightarrow P(X+Y>n) < 1\%$$
$$\Rightarrow P(X+Y \geq n+1) < 0.01$$
$$\Rightarrow 1 - P(X+Y \leq n) < 0.01$$
$$\Rightarrow -P(X+Y \leq n) < -0.99$$
$$\Rightarrow P(X+Y \leq n) > 0.99$$

LOOKING AT THE TABLES OF $Po(5.5)$

$$\Rightarrow \underline{\underline{n = 12}}$$

IYGB - FS1 PAPER O - QUESTION 6

$\sum d = 385.5 \quad \sum v = 22.5 \quad \sum dv = 650.25$

$\sum d^2 = 11543.25 \quad \sum v^2 = 38.25 \quad n = 15.$

a) DEPTH IS THE EXPLANATORY VARIABLE, I.E INDEPENDENT VARIABLE AND FREQUENCY IS THE RESPONSE VARIABLE (DEPENDENT VARIABLE)

THIS IS BECAUSE IT IS THE DEPTH WHICH AFFECTS FREQUENCY AND <u>NOT</u> THE OTHER WAY ROUND

b) CALCULATE S_{dd} S_{vv} & S_{dv}

$$S_{dd} = \sum d^2 - \frac{\sum d \sum d}{n} = 11543.25 - \frac{385.5 \times 385.5}{15} = 1635.9$$

$$S_{vv} = \sum v^2 - \frac{\sum v \sum v}{n} = 38.25 - \frac{22.5 \times 22.5}{15} = 4.5$$

$$S_{dv} = \sum dv - \frac{\sum d \sum v}{n} = 650.25 - \frac{385.5 \times 22.5}{15} = 72$$

c) FIND THE PMCC.

$$r = \frac{S_{dv}}{\sqrt{S_{dd} S_{vv}}} = \frac{72}{\sqrt{1635.9 \times 4.5}} \simeq 0.839$$

d) POSITIVE CORRELATION, I.E THE GREATER THE DEPTH, THE HIGHER THE FREQUENCY AND VISE VERSA

e) THE P.M.C.C IS REASONABLY HIGH TO SUGGEST A GOOD LINEAR MODEL MIGHT BE APPROPRIATE

f)
$$b = \frac{S_{dv}}{S_{dd}}$$

$$b = \frac{72}{1635.9}$$

$$d = \frac{240}{5453} \approx 0.0440$$

3 s.f.

IYGB - FS1 PAPER O - QUESTION 6

"$a = \bar{y} - b\bar{x}$", HERE $a = \bar{v} - b\bar{d}$ ← $\frac{\sum d}{n} = \frac{385.5}{15} = \underline{\underline{25.7}}$

$\frac{\sum v}{n} = \frac{22.5}{15} = \underline{\underline{1.5}}$

$a = 1.5 - 0.0440... \times 25.7$

$a = 0.369$ 3 sf

g) $a =$ "y INTERCEPT"

THIS REPRESENTS THE "FREQUENCY OF DOLPHIN" WHEN IT IS AT THE SURFACE (ZERO DEPTH)

$b =$ "gradient"

INCREASE IN THE FREQUENCY PER METRE DEPTH, FOR EVERY METRE FOR WHICH THE DOLPHIN DIVES, THE FREQUENCY INCREASES BY 0.044 KHz

IYGB - FS1 PAPER O - QUESTION 7

a) $E(X) = \int_a^b x f(x)\,dx$

$E(X) = \int_2^4 x\left(\frac{1}{60}x^3\right)dx = \int_2^4 \frac{1}{60}x^4 = \left[\frac{1}{300}x^5\right]_2^4$

$= \frac{1}{300}\left[1024 - 32\right] = \frac{248}{75} \simeq 3.31$

b) START BY FINDING $E(X^2) = \int_a^b x^2 f(x)\,dx$

$E(X^2) = \int_2^4 x^2\left(\frac{1}{60}x^3\right)dx = \int_2^4 \frac{1}{60}x^5\,dx = \frac{1}{360}\left[x^6\right]_2^4$

$= \frac{1}{360}\left[4096 - 64\right] = \frac{56}{5}$

NOW USING $Var(X) = E(X^2) - \left[E(X)\right]^2$

$Var(X) = \frac{56}{5} - \left(\frac{248}{75}\right)^2$

$Var(X) = 0.2659555\ldots$

$\therefore$ STANDARD DEVIATION $= \sqrt{0.2659555\ldots}$

$\simeq 0.516$ // 3 d.p

c) $F(x) = \int_a^x f(x)\,dx$

$F(x) = \int_2^x \frac{1}{60}x^3\,dx = \left[\frac{1}{240}x^4\right]_2^x = \frac{1}{240}\left(x^4 - 16\right)$

$$\therefore\ F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{240}(x^4 - 16) & 2 \leqslant x \leqslant 4 \\ 1 & x > 4 \end{cases}$$

IYGB - FS1 PAPER O - QUESTION 7

d) $P(X > 3.5) = 1 - P(X < 3.5)$

$= 1 - F(3.5)$

$= 1 - \frac{1}{240}(3.5^4 - 16)$

$= \frac{113}{256}$

e) SOLVING $F(x) = \frac{1}{2}$

$\frac{1}{240}(x^4 - 16) = \frac{1}{2}$

$x^4 - 16 = 120$

$x^4 = 136$

$x = +\sqrt[4]{136}$

$x \simeq 3.41$

2 d.p.