Exam-style practice AS Level

1 a Probabilities sum to 1, so:

$\sum P(X = x) = 0.1 + a + 0.15 + 0.2 + b = 1$

$\Rightarrow 04.5 + a + b = 1$

$\Rightarrow a + b = 0.55$ **(1)**

$Y = 2X + 3 \Rightarrow X = \frac{Y-3}{2}$

$E(X) = E\left(\frac{Y-3}{2}\right) = E(0.5Y - 1.5)$

$= 0.5E(Y) - 1.5$

$= 0.5 \times 4.48 - 1.5$ (using $E(Y) = 4.48$)

$= 0.74$

So $E(X) = \sum x P(X = x) = 0.74$ gives

$\Rightarrow -2 \times 0.1 + (-1) \times a + 0 \times 0.15 + 1 \times 0.2 + 2 \times b = 0.74$

$\Rightarrow -0.2 - a + 0.2 + 2b = 0.74$

$\Rightarrow -a + 2b = 0.74$ **(2)**

Add equation (1) and (2)

$3b = 1.29 \Rightarrow b = 0.43$

Substitute value of b in (1)

$a = 0.55 - 0.43 = 0.12$

Solution: $a = 0.12 \quad b = 0.43$

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

$= 4 \times 0.1 + 1 \times 0.12 + 1 \times 0.2 + 4 \times 0.43 = 2.44$

$Var(X) = E(X^2) - (E(X))^2$

$= 2.44 - 0.74^2 = 2.44 - 0.5476$

$= 1.8924$

c $Y - 2 > X \Rightarrow 2X + 3 - 2 > X \Rightarrow X > -1$

$P(Y - 2 > X) = P(X > -1) = P(X \geqslant 0)$

$= 0.15 + 0.2 + 0.43 = 0.78$

2 a Let the random variable X denote the number of calls about insurance in a 10-minute interval, so $X \sim \text{Po}(3.2)$; and the random variable Y denote the number of calls about utility bills in a 10-minute interval, so $Y \sim \text{Po}(4.1)$

As the calls are independent of each other

$$\text{P}(X=3 \cap Y=3) = \text{P}(X=3) \times \text{P}(Y=3)$$
$$= \frac{e^{-3.2}\,3.2^3}{3!} \times \frac{e^{-4.1}\,4.1^3}{3!} = 0.22262 \times 0.19037 = 0.0424 \text{ (4 d.p.)}$$

b $X+Y \sim \text{Po}(3.2+4.1)$, i.e. $X+Y \sim \text{Po}(7.3)$

By calculator

$$\text{P}((X+Y) \geqslant 7) = 1 - \text{P}((X+Y) \leqslant 6)$$
$$= 1 - 0.4060 = 0.5940 \text{ (4 d.p.)}$$

c Let the random variable T denote the number of calls received in a one-hour period, so $T \sim \text{Po}(6 \times 7.3)$, i.e. $T \sim \text{Po}(43.8)$

By calculator

$\text{P}(T < 45) = \text{P}(T \leqslant 44) = 0.5520$ (4 d.p.)

3 Find the respective totals for each sport and gender in the sample.

	Hockey	Cricket	Squash	Totals
Male	61	45	32	138
Female	66	23	23	112
Totals	127	68	55	250

a H_0: There is no association between sport and gender.

H_1: There is an association between sport and gender.

3 b Calculate the expected values as follows:

$$\text{P(male and hockey)} = \frac{138}{250} \times \frac{127}{250}$$

$$\text{Expected frequency of male and hockey} = \frac{138}{250} \times \frac{127}{250} \times 25 = \frac{138 \times 127}{250} = 70.104$$

	Hockey	**Cricket**	**Squash**	**Totals**
Male	$\frac{138 \times 127}{250} = 70.104$	$\frac{138 \times 68}{250} = 37.536$	$\frac{138 \times 55}{250} = 30.36$	138
Female	$\frac{112 \times 127}{250} = 56.896$	$\frac{112 \times 68}{250} = 30.464$	$\frac{112 \times 55}{250} = 24.64$	112
Totals	127	68	55	250

All expected frequencies > 5

O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
61	70.104	1.1822
66	56.896	1.4567
45	37.536	1.4842
23	30.464	1.8288
32	30.36	0.0886
23	24.64	0.1092

$$\text{Test statistic} = \sum \frac{(O_i - E_i)^2}{E_i} = 6.1497 \text{ (4 d.p.)}$$

c This is a 2×3 contingency table. The number of degrees of freedom $= (2-1)(3-1) = 2$

d At the 2.5% level of significance for $\nu = 2$, the critical value is $\chi_2^2 = 7.378$
As $6.1497 < 7.378$, this is not significant and H_0 should not be rejected.

e At the 5% level of significance for $\nu = 2$, the critical value for $\chi_2^2 = 5.991$
As $6.1497 > 5.991$, this is significant and therefore reject H_0

4 a Let the random variable X be the number of defects found in a the sample of 350 bowls,
$X \sim B(750, 0.005)$

$\text{Mean} = E(X) = np = 750 \times 0.005 = 3.75$

$\text{Variance} = \text{Var}(X) = np(1-p) = 3.75 \times 0.995 = 3.73125$

4 b Using the approximation $X \approx\sim \text{Po}(3.75)$

By calculator

$\text{P}(X > 3) = 1 - \text{P}(X \leqslant 3) = 1 - 0.4838 = 0.5162$ (4 d.p.)

c The mean n is large and p is small, so the mean is approximately equal to the variance. Hence, the Poisson distribution is a good approximation for the binomial distribution.

5 a Let probability of any single coin landing on heads = p, then an estimate of p is:

$$p = \frac{\sum(x \times f_x)}{\text{number of trials } \times \text{number of observations}}$$

$$= \frac{0\times6+1\times18+2\times35+3\times26+4\times15}{4\times(6+18+35+26+15)}$$

$$= \frac{226}{400} = 0.565$$

b H_0: A B(4, 0.565) distribution is a suitable model for the results.

H_1: A B(4, 0.565) distribution is not a suitable model for the results.

From part **a**, the null hypothesis is that B(4, 0.565) is a suitable model.

x	P(x)	E_i
0	$1\times0.565^0\times0.435^4 = 0.035806$	3.5806
1	$4\times0.565^1\times0.435^3 = 0.186027$	18.6027
2	$6\times0.565^2\times0.435^2 = 0.362432$	36.2432
3	$4\times0.565^3\times0.435^1 = 0.313830$	31.3830
4	$1\times0.565^4\times0.435^0 = 0.101905$	10.1905

Since 3.5806 < 5 combine cells.

x	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
0 or 1	24	22.1833	0.14878
2	35	36.2432	0.04264
3	26	31.3830	0.92332
4	15	10.1905	2.26989

Number of degrees of freedom = number of cells – 2 = 4 – 2 =2 (p is estimated by calculation)

From the tables: χ^2_2 (10%) is 4.605

$$\sum\frac{(O_i - E_i)^2}{E_i} = 3.3846 \text{ (4 d.p.)}$$

As 3.3846 < 4.605, do not reject H_0. At the 10 % level of significance it appears that B(4, 0.565) is a suitable model for the data.