For comments, corrections, etc…Please contact Ahnaf Abbas: ahnaf@mathyards.com

International Institute for
Technology and Management

Tutoring Sheet # 2b

Unit 04b : Statistics 2

1. i. From your tables find the probability that there are at least 6 successes for the binomial distribution with 17 trials and probability of success 0.35.

ii. From your tables find the probability that $X > 7$ when X has a Poisson distribution with mean 4.5.

For this, one needs familiarity with the statistical tables used in the examination room. It is best to use them when revising for the exam. Typical calculators in use by candidates can't give a correct answer as easily as the tables.

i. For the binomial probability one needs the complementary probability to that of no more than 5 successes:

$$P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.4197 = 0.5803$$

ii.For the Poisson probability one should look for the complementary probability to that for no more than 6:

$$P(X > 7) = 1 - P(X \leq 6) = 1 - 0.8311 = 0.1869$$

Remember that the table show cumulative probabilities.

2. Consider two random variables X and Y. X can take the values 0, 1 and 2 and Y can take the values 1 and 2. The joint probabilities for each pair are given by the following table.

	$X=0$	$X=1$	$X=2$
$Y=1$	0.1	0.2	0.3
$Y=2$	0.1	0.1	0.2

Let $Z = \max(X,Y)$ be the larger of the two variables. Find $E(Y)$, $E(Y|X=1)$, $E(Z)$ and $E(Z|X=1)$.

$$\mu_Y = E(Y) = \sum_y y p_Y(y) = 1 \times (0.1 + 0.2 + 0.3) + 2 \times (0.1 + 0.1 + 0.2) = 1.4$$

$$P(X = x) = p_X(x) = \sum_y p_{XY}(x,y) \Rightarrow P(X = 1) = 0.2 + 0.1 = 0.3$$

$$P_{Y|X}(y \mid x) = \frac{p_{XY}(x,y)}{p_X(x)} \Rightarrow P(Y = 1 \mid X = 1) = 0.2 / 0.3 = 2/3$$

Similarly , P(Y = 2 | X = 1) = 0.1/0.3 = 1/3

$E(Y|X) = \sum_y y p_{Y|X}(x,y) \Rightarrow$ E(Y | X = 1) = 1 × 2/3 + 2 × 1/3 = 4/3

E (Z) = 1× (0.1 + 0.2) + 2×(0.3 + 0.1 + 0.1 + 0.2) = 1.7

Now P(Z = 1 | X = 1) = P (Y =1 | X = 1) = 2/3
P(Z = 2 | X = 1) = P (Y = 2 | X = 1) = 1/3

E(Z) = 1 × 2/3 + 2 × 1/3 = 4/3

3. The random variable X has a density function given by

$$f(x) = \frac{3x^2 + 2x}{2}$$

defined over the region $0 \le x \le 1$. Find $\Pr(X > 0.8 | X > 0.6)$, $E(X)$, $\mathrm{Var}(X)$ and $\mathrm{Cov}\left(X, \frac{1}{X}\right)$.

$$P(X > 0.8 \mid X > 0.6) = \frac{P(X > 0.8)}{P(X > 0.6)} = \frac{1 - \int_0^{0.8} \frac{3x^2 + 2x}{2} dx}{1 - \int_0^{0.6} \frac{3x^2 + 2x}{2} dx} = \ldots\ldots.$$

$$E(X) = \int_0^1 x \frac{3x^2 + 2x}{2} dx = 17/21$$

$$E(X^2) = \int_0^1 x^2 \frac{3x^2 + 2x}{2} dx = \ldots\ldots.$$

$$Var(X) = E(X^2) - E^2(X) = \ldots\ldots.$$

$$Cov(X, \frac{1}{X}) = E(X.\frac{1}{X}) - E(X)E(\frac{1}{X}) = 1 - (17/21)\int_0^1 \frac{1}{x}\frac{3x^2 + 2x}{2} dx = \ldots\ldots.$$

4. The random variable X is normally distributed with mean 0 and variance 9.

Find $\Pr(X > 3.6 \mid X > 1.8)$ and $\Pr(|X| > 3.6 \mid |X| > 1.8)$

The standard deviation is 3. So

$$\Pr(X > 3.6 | X > 1.8) = \frac{\Pr(X > 3.6)}{\Pr(X > 1.8)} = \frac{1 - \Phi\left(\frac{3.6}{3}\right)}{1 - \Phi\left(\frac{1.8}{3}\right)} = \frac{1 - \Phi(1.2)}{1 - \Phi(0.6)} = \frac{1 - 0.8849}{1 - 0.7257} = 0.4196. \text{ Also}$$

$$\Pr(|X| > 3.6 \,||\, X| > 1.8) = \frac{\Pr(|X| > 3.6)}{\Pr(|X| > 1.8)} = \frac{2\left(1 - \Phi\left(\frac{3.6}{3}\right)\right)}{2\left(1 - \Phi\left(\frac{1.8}{3}\right)\right)} = \frac{1 - \Phi(1.2)}{1 - \Phi(0.6)} = \frac{1 - 0.8849}{1 - 0.7257} = 0.4196.$$

5. Consider two random variables X and Y. X can take the values -1, 0 and 1 and Y can take the values 0, 1 and 2. The joint probabilities for each pair are given by the following table.

	$X=-1$	$X=0$	$X=1$
$Y=0$	0.1	0.2	0.1
$Y=1$	0.1	0.05	0.1
$Y=2$	0.1	0.05	0.2

a) Calculate the marginal distributions and expected values of X and Y.

(9 marks)

P(X = x) = $p_X(x) = \sum_y p_{XY}(x,y) \Rightarrow$ The marginal distributions of ***X*** are obtained by computing <u>column sums</u>:

P(X= -1) = 0.1 + 0.1 + 0.1 = 0.3
P(X = 0) = 0.2 + 0.05 + 0.05 = 0.3
P(X = 1) = 0.1 + 0.1 + 0.2 = 0.4

μ_X = E(X) = $\sum_x xp_X(x)$ = (-1)(0.3) + (0)(0.3) + (1)(0.4) = 0.1

P(Y = y) = $p_Y(y) = \sum_x p_{XY}(x,y) \Rightarrow$ The marginal distributions of ***X*** are obtained by computing <u>row sums</u>:

P(Y= 0) = 0.1 + 0.2 + 0.1 = 0.4
P(Y = 1) = 0.1 + 0.05 + 0.1 = 0.25
P(Y = 2) = 0.1 + 0.05 + 0.2 = 0.35

μ_Y = E(Y) = $\sum_y yp_Y(y)$ = (0)(0.4) + (1)(0.25) + (2)(0.35) = 0.95

b) Calculate the covariance of the random variables U and V, where $U = X + Y$ and $V = X - Y$.

(7 marks)

Cov(U,V) = E[(X+Y)(X-Y)] – E(X+Y)E(X-Y)
= E(X^2 – Y^2) – [E(X) + E(Y)][E(X)-E(Y)]
= E(X^2) – E(Y^2) – (0.1+0.95)(0.1-0.95)
= E(X^2) – E(Y^2) – 0.8925

E(X^2) = $\sum_x x^2 p_X(x)$ = $(-1)^2$(0.3) + $(0)^2$(0.3) + $(1)^2$(0.4) = 0.7

E(Y^2) = $\sum_y y^2 p_Y(y)$ = $(0)^2$(0.4) + $(1)^2$(0.25) + $(2)^2$(0.35) = 1.65

Cov(U,V) = E(X^2) – E(Y^2) – 0.8925 = 0.7 – 1.65 – 0.8925 = -1.8425

c) Calculate $E(V|U=1)$

(7 marks)

U = 1 can be obtained by the pairs (X,Y) = { (-1,2) , (0,1) , (1,0) }
The corresponding values of V are : - 3 , - 1 , 1

	U= x+y = 1
V= x-y = -3	0.1
V= x-y = -1	0.05
V = x-y = 1	0.1

P(U =1) = $p_U(u) = \sum_V p_{UV}(u,v)$ = 0.1 + 0.05 + 0.1 = 0.25 (sum of column)

$P_{V|U}$ (V | U) = $\dfrac{p_{UV}(u,v)}{p_U(u)} \Rightarrow$ P (V= -3 | U =1) = 0.1/ 0.25 = 2/5

P (V= -1 | U =1) = 0.05/ 0.25 = 1/5

P (V= 1 | U =1) = 0.1/ 0.25 = 2/5

$$E(V|U) = \sum_V v p_{V|U}(u,v) = \sum_{x-y} (x-y) p_{X-Y|X+Y}(x+y, x-y)$$

Hence E(V | U = 1) = (-3)(2/5) + (-1)(1/5) + (1)(2/5) = - 1

d) The random variable W has the same distribution as X and the random variable Z has the same distribution as Y. The random variables W and Z are independent. Write down the table for the joint probabilities of W and Z and calculate their covariance.

(7 marks)

From part (a) :
P(X= -1) = 0.1 + 0.1 + 0.1 = 0.3 , P(Y= 0) = 0.1 + 0.2 + 0.1 = 0.4
P(X = 0) = 0.2 + 0.05 + 0.05 = 0.3 , P(Y = 1) = 0.1 + 0.05 + 0.1 = 0.25
P(X = 1) = 0.1 + 0.1 + 0.2 = 0.4 , P(Y = 2) = 0.1 + 0.05 + 0.2 = 0.35

	W = -1	W = 0	W = 1
Z = 0	0.3×0.4	0.3×0.4	0.4×0.4
Z = 1	0.3×0.25	0.3×0.25	0.4×0.25
Z = 2	0.3×0.35	0.3×0.35	0.4×0.35

The covariance is of course **0** since they are independent.

6. The random variable X has density function given by $f(x) = \dfrac{12x^2(x+1)}{7}$

defined over the region $0 < x < 1$.
(a) Calculate Pr ($X > 0.5 \mid X > 0.25$) and $E(X)$.

a)

$$\Pr(X > 0.5 | X > 0.25) = \frac{\Pr(X>0.5)}{\Pr(X>0.25)} = \frac{1-\int_0^{0.5} \frac{12x^3+12x^2}{7} \mathrm{dx}}{1-\int_0^{0.25} \frac{12x^3+12x^2}{7} \mathrm{dx}}$$

$= 0.911.$

$$E(X) = \int_0^1 x \frac{12x^3+12x^2}{7} \mathrm{dx} = \frac{12}{7}\left(\frac{1}{5}+\frac{1}{4}\right) = \frac{108}{140} = \frac{27}{35} = 0.771.$$

(b) Calculate Cov($\frac{1}{1+X}, \frac{1}{X^2}$)

$E\left(\frac{1}{X+1}\right) = \int_0^1 \frac{12x^2}{7} \mathrm{dx} = \frac{4}{7}$

$E\left(\frac{1}{X^2}\right) = \int_0^1 \frac{12(x+1)}{7} \mathrm{dx} = \frac{18}{7}$

$E\left(\frac{1}{(X+1)X^2}\right) = \int_0^1 \frac{12}{7} \mathrm{dx} = \frac{12}{7}$

and so $\mathrm{Cov}\left(\frac{1}{X+1}, \frac{1}{X^2}\right) = \frac{12}{7} - \frac{4}{7} \times \frac{18}{7} = \frac{12}{49} = 0.245.$

7. Consider two random variables *X* and *Y*. They both take the values 0, 1 and 2. The joint probabilities for each pair are given by the following table.

	X = 0	**X = 1**	**X = 2**
Y = 0	**0.10**	**0.06**	**0.14**
Y = 1	**0.08**	**0.06**	**0.16**
Y = 2	**0.20**	**0.08**	**0.12**

(a) Calculate the marginal distributions, and the expected values of *X* and *Y*.

$\Pr(X=0) = 0.10 + 0.08 + 0.20 = 0.38$,
$\Pr(X=1) = 0.06 + 0.06 + 0.08 = 0.20$, $\Pr(X=2) = 0.14 + 0.16 + 0.12 = 0.42$,
$\Pr(Y=0) = 0.10 + 0.06 + 0.14 = 0.30$, $\Pr(Y=1) = 0.08 + 0.06 + 0.16 = 0.30$ and
$\Pr(Y=2) = 0.20 + 0.08 + 0.12 = 0.40$. So

$$E(X) = 0 \times 0.38 + 1 \times 0.20 + 2 \times 0.42 = 1.04$$

$$E(Y) = 0 \times 0.30 + 1 \times 0.30 + 2 \times 0.4 = 1.1.$$

(b) Calculate $E(X \mid Y=1)$ and $E(X \mid X+Y=3)$.

We have that:

$$P(X=0|Y=1) = \frac{0.08}{0.3} = \frac{8}{30} = \frac{4}{15},$$
$$P(X=1|Y=1) = \frac{0.06}{0.3} = \frac{6}{30} = \frac{1}{5}$$

and

$$P(X=2|Y=1) = \frac{0.16}{0.3} = \frac{16}{30} = \frac{8}{15}$$

and therefore

$$E(X|Y=1) = 0 \times \frac{4}{15} + 1 \times \frac{1}{5} + 2 \times \frac{8}{15} = \frac{19}{15} = 1.267.$$

We also have

$$P(X+Y=3) = 0.16 + 0.08 = 0.24$$

and so

$$P(X=1|X+Y=3) = \frac{0.08}{0.24} = \frac{1}{3}$$

and

$$P(X=2|X+Y=3) = \frac{0.16}{0.24} = \frac{2}{3}.$$

Hence

$$E(X|X+Y=3) = 1 \times \frac{1}{3} + 2 \times \frac{2}{3} = \frac{5}{3} = 1.667.$$

(c) Define $U = |X-1|$ and $V = Y$. Calculate the covariance of U and V.

Here is the table of probabilities:

	$U=0$	$U=1$
$V=0$	0.06	0.24
$V=1$	0.06	0.24
$V=2$	0.08	0.32

We then have $\Pr(U=0) = 0.06 + 0.06 + 0.08 = 0.2$, $\Pr(U=1) = 0.24 + 0.24 + 0.32 = 0.8$. Of course $\Pr(V=0) = 0.3$, $\Pr(V=1) = 0.3$ and $\Pr(V=2) = 0.4$. So

$$E(U) = 0 \times 0.2 + 1 \times 0.8 = 0.8,$$
$$E(V) = E(Y) = 1.1$$

and

$$E(UV) = 1 \times 0.24 + 2 \times 0.32 = 0.88.$$

Hence $\mathrm{Cov}(U,V) = 0.88 - 0.8 \times 1.1 = 0$. **(6 marks)**.

(d) Are *U* and *V* are independent variables? Explain your answer.

The fact that their covariance is 0 is not enough to show that they are independent. However the table of probabilities can be rewritten as

	$U = 0$	$U = 1$
$V = 0$	0.2×0.3	0.8×0.3
$V = 1$	0.2×0.3	0.8×0.3
$V = 2$	0.2×0.4	0.8×0.4

We observe that $\Pr(U = i, V = j) = \Pr(U = i)\Pr(V = j)$ for all possible pairs (i, j) and so they are independent.

Note that all pairs have to be checked. (**5 marks**).

8. Consider random variables *X* and *Y* with joint density function

$$f_{XY}(x, y) = \begin{cases} k(3x-2) & 0 < y < x < 5 \\ 0 & otherwise \end{cases}$$

(a) Find *k*.

To find k we choose it to ensure that the total probability is 1. See the same being checked in Example 4.2.10. For $0 < x < 2$,

$$f_X(x) = \int_0^x k(3x-2)dy = k(3x-2)y|_0^x = k(3x-2)x = k(3x^2 - 2x),$$

and for x outside this domain, $f_X(x) = 0$. Integrating over x (we need only the range for which the density is positive)

$$1 = \int f_X(x) = \int_0^2 k(3x^2 - 2x)dx = k(x^3 - x^2)|_0^2 = k(8-4) = 4k.$$

So k=1/4.

(b) Find $f_X(x)$. Hence evaluate $E(X)$.

As above we have

$$f_X(x) = \begin{cases} 0 & x \le 0 \\ \frac{3x^2-2x}{4} & 0 < x < 2 \\ 0 & x \ge 2. \end{cases}$$

Then,

$$E(X) = \int_0^2 x(3x^2 - 2x)/4dx = \int_0^2 (3x^3 - 2x^2)/4dx = (3x^4/4 - 2x^3/3)/4|_0^2 = (12 - 16/3)/4 = 5/3.$$

(c) Write down an expression for $f_{Y|X}(y|x)$. Find $E(Y \mid X)$ and hence evaluate $E(Y)$.

The conditional density is defined only for $0 < x < 2$. For $0 < y < x$,

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{(3x-2)/4}{(3x^2-2x)/4} = 1/x.$$

This is a uniform density over $(0, x)$. One should write the conditional density in full (for $0 < x < 2$) as

$$f_{Y|X}(y|x) = \begin{cases} 0 & y \le 0 \\ 1/x & 0 < y \le x \\ 0 & x < y. \end{cases}$$

For $0 < x < 2$,

$$E(Y|X=x) = \int_0^x y f_{Y|X}(y|x)dy = \int_0^x y/x dy = y^2/(2x)\Big|_0^x = x^2/(2x) = x/2.$$

This is obvious anyway because the conditional distribution is a uniform density over $(0, x)$. Then

$$E(Y) = E[E(Y|X)] = E(X/2) = E(X)/2 = 5/6.$$

(e) Evaluate $P(2Y > X)$.

The region $2Y > X$ is the region where $Y > X/2$, so the probability required is obtained by integrating the joint density over the region $0 < x < 2$, $x/2 < y < x$. It is obvious from the uniform conditional distribution of $Y|X$ that this probability is 1/2, but we will carry it through, first integrating over y and then x. The calculations amount to finding the probability for the second half of a uniform distribution.

$$\begin{aligned}
P(2Y > X) &= \int_0^2 \int_{x/2}^x f_{X,Y}(x,y)dydx \\
&= \int_0^2 \int_{x/2}^x f_{Y|X}(y|x) f_X(x)dydx \\
&= \int_0^2 \int_{x/2}^x \frac{1}{x} f_X(x)dydx = \int_0^2 \left[\int_{x/2}^x \frac{1}{x}dy\right] f_X(x)dx \\
&= \int_0^2 [y/x|_{x/2}^x] f_X(x)dx = \int_0^2 [(x - x/2)/x] f_X(x)dx = \int_0^2 \frac{1}{2} f_X(x)dx \\
&= \frac{1}{2}\int_0^2 f_X(x)dx = \frac{1}{2}.
\end{aligned}$$