ST107 Outline solutions to Exercise 3

1. (a) The calculation for this part can conveniently be organised in a table as follows:

$X = x$	3	6	Total
$P(X = x)$	0.75	0.25	1
$x\,P(X = x)$	2.25	1.50	$3.75 = \mathrm{E}(X)$
$x^2 P(X = x)$	6.75	9.00	$15.75 = \mathrm{E}(X^2)$

Hence the mean is $\mathrm{E}(X) = \mu = 3.75$, while the variance is:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = 15.75 - (3.75)^2 = 1.6875.$$

Note it usually helps to have the 'Total' column and to check that the probabilities do add up to one, because it is very easy to miscalculate or to miscopy.

(b) i. The possible values of T range from 36 (all 3s) to 72 (all 6s), and every such value is a multiple of 3. So the value $T = 40$ is impossible, and hence $P(T = 40) = 0$.

ii. The only way that the total can be 45 is with nine 3s and three 6s so we want:

$$P(\text{nine 3s \& three 6s}) = P(\text{exactly nine 3s}) = P(\text{exactly three 6s})$$

which is:

$$\binom{12}{9} \times (0.75)^9 \times (0.25)^3 = 0.2581.$$

Equivalently, we can express this using the binomial distribution with $n = 12$ and $\pi = 0.75$, where a success is defined as obtaining a 3. So, if X is the number of successes, then $X \sim \mathrm{Bin}(12, 0.75)$ and:

$$P(X = 9) = \binom{12}{9} \times (0.75)^9 \times (0.25)^3 = 0.2581.$$

Note that this could also be calculated in Excel using the function:

```
=BINOM.DIST(9,12,0.75,0)
```

where the function's arguments are as follows.

1. The number of successes is $x = 9$.
2. The number of trials is $n = 12$.
3. The probability of success is $\pi = 0.75$.
4. The 0 indicates that a non-cumulative probability is returned, i.e. it returns $P(X = 9)$. (If this argument was set to 1, then the function would return the cumulative probability $P(X \leq 9)$.)

2. (a) The probability of *not* favouring the new service for every person is 8% and we are looking at a pilot sample of 17 people. Therefore, we have $\pi = 0.08$ and $n = 17$. Now, if X is the number of people *not* in favour, then the probability that at least

three people do not favour the new service is the same as the probability that $X \geq 3$. This is equivalent to 1 minus the probability for only zero, one or two persons, so:

$$\begin{aligned} P(X \geq 3) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - (0.92)^{17} - (17 \times 0.08 \times (0.92)^{16}) - (136 \times (0.08)^2 \times (0.92)^{15}) \\ &= 0.1503. \end{aligned}$$

The assumption made in order to justify this method of calculation is that the population within the M25 is large, we can assume that the probability of not wanting mobile telephone access is the same for every person asked, at 0.08, i.e. that we can treat the process of sampling several people as a sequence of independent and identical Bernoulli trials. Another way of saying the same thing is that we are assuming we can treat the sampling process as being with replacement. The same assumption applies in (b).
Note this probability can be computed in Excel using the function:

```
=1-BINOM.DIST(2,17,0.08,1)
```

which returns $1 - P(X \leq 2)$, noting that the final argument of '1' makes Excel calculate a cumulative probability.

(b) Now a suitable approximation, since the probability for the event is small ($\pi = 0.08$) and the number of events is large ($n = 110$), is the Poisson approximation to the binomial. If the random variable this time is Y, then the Poisson approximation says that the probability that $Y = y$ is given by:

$$P(Y = y) = \begin{cases} e^{-\lambda}\lambda^y/y! & \text{for } y = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

where λ is the rate parameter, that is the average number of successes – the mean of the Poisson distribution. Since $\pi = 0.08$ and $n = 110$, the mean will be $\lambda = n\pi = 8.8$, and $8.8 < 10$ (as required) while n is certainly 'large'. So, using the Poisson approximation with $\lambda = 8.8$, we get:

$$\begin{aligned} P(Y \geq 3) &= 1 - P(Y = 0) - P(Y = 1) - P(Y = 2) \\ &= 1 - \frac{e^{-8.8}(8.8)^0}{0!} - \frac{e^{-8.8}(8.8)^1}{1!} - \frac{e^{-8.8}(8.8)^2}{2!} \\ &= 1 - e^{-8.8}(1 + 8.8 + 38.72) \qquad \text{(noting the common factor, } e^{-8.8}\text{)} \\ &= 0.9927. \end{aligned}$$

The extra assumption made here is that the Poisson approximation can be applied, and the detailed checks necessary were carried out above.
Note this probability can be computed in Excel using the function:

```
=1-POISSON.DIST(2,8.8,1)
```

which returns $1 - P(X \leq 2)$, where the first argument is the value of x, the second argument is the rate parameter value, and the third argument returns a cumulative probability.

3. (a) The calculations for this part can conveniently be organised in a table as follows.

$X = x$	$P(X = x)$	$x\,P(X = x)$	$x^2 P(X = x)$	$(X - 2)^2$
0	$\binom{6}{0} \times (0.3)^0 \times (0.7)^6 = 0.117649$	0.000000	0.000000	4
1	$\binom{6}{1} \times (0.3)^1 \times (0.7)^5 = 0.302526$	0.302526	0.302526	1
2	$\binom{6}{2} \times (0.3)^2 \times (0.7)^4 = 0.324135$	0.648270	1.296540	0
3	$\binom{6}{3} \times (0.3)^3 \times (0.7)^3 = 0.185220$	0.555660	1.666980	1
4	$\binom{6}{4} \times (0.3)^4 \times (0.7)^2 = 0.059535$	0.238140	0.952560	4
5	$\binom{6}{5} \times (0.3)^5 \times (0.7)^1 = 0.010206$	0.051030	0.255150	9
6	$\binom{6}{6} \times (0.3)^6 \times (0.7)^0 = 0.000729$	0.004374	0.026244	16
Total	1	1.8 $= \mathrm{E}(X)$	4.5 $= \mathrm{E}(X^2)$	

The final column shows that different values of X give identical values for $(X - 2)^2$. So, to find its probability distribution, we need to identify the *new* sample space, and then to calculate the probability for each item in it. This can conveniently be done as in the following table, where we define $Y = (X - 2)^2$.

$Y = y$	$P(Y = y)$	$y\,P(Y = y)$
0	0.324135	0.000000
1	$0.302526 + 0.185220 = 0.487746$	0.487746
4	$0.117649 + 0.059535 = 0.177184$	0.708736
9	0.010206	0.091854
16	0.000729	0.011664
Total	1	1.3 $= \mathrm{E}(Y)$

(b) i. From the tables above we see that $\mathrm{E}(X) = 1.8$, $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = 4.5 - (1.8)^2 = 1.26$ and $\mathrm{E}((X - 2)^2) = \mathrm{E}(Y) = 1.3$.

ii. Since $X \sim \mathrm{Bin}(6, 0.3)$, $\mathrm{E}(X) = 6 \times 0.3 = 1.8$ and $\mathrm{Var}(X) = 6 \times 0.3 \times 0.7 = 1.26$. Also note that $\mathrm{E}(X^2) = \mathrm{Var}(X) + (\mathrm{E}(X))^2 = 1.26 + (1.8)^2 = 4.5$.

For any random variables X and Y, and for any constants α and β (with both non-zero), we have a new random variable $Z = \alpha X + \beta Y$ such that:

$$\mathrm{E}(Z) = \mathrm{E}(\alpha X + \beta Y) = \alpha\,\mathrm{E}(X) + \beta\,\mathrm{E}(Y).$$

Hence:

$$\begin{aligned} \mathrm{E}(Y) &= \mathrm{E}((X - 2)^2) \\ &= \mathrm{E}(X^2 - 4X + 4) \\ &= \mathrm{E}(X^2) - 4 \times \mathrm{E}(X) + \mathrm{E}(4) \\ &= 4.5 - 4 \times 1.8 + 4 \\ &= 1.3. \end{aligned}$$

4. (a) A process has a Poisson distribution if it is a purely random point process – that is, 'events' have no length or duration, and they can be counted – and the occurrences of any pair of distinct 'events' are independent. The probability of two or more occurrences at the *same* time is negligibly small.

The average density of points per unit of medium is usually written as λ. The relevant random variable X is the number of points which will occur in a fixed interval of unit size. It takes values $0, 1, 2, \ldots, n, \ldots$ (so it has no upper limit, unlike the binomial distribution).

(b) The rate is given as 78 per hour, but it is convenient to work in numbers of minutes, so note that this is the same as $\lambda = 1.3$ arrivals per minute.

i. For two minutes, use $\lambda = 1.3 \times 2 = 2.6$. Hence:

$$P(X = 6) = \frac{\mathrm{e}^{-2.6}(2.6)^6}{6!} = 0.0319.$$

ii. For 90 seconds, $\lambda = 1.3 \times 1.5 = 1.95$. Hence:

$$\begin{aligned} P(X > 2) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - \frac{\mathrm{e}^{-1.95}(1.95)^0}{0!} - \frac{\mathrm{e}^{-1.95}(1.95)^1}{1!} - \frac{\mathrm{e}^{-1.95}(1.95)^2}{2!} \\ &= 1 - \mathrm{e}^{-1.95}(1 + 1.95 + 1.90125) \\ &= 0.3098. \end{aligned}$$

iii. The probability that the time to arrival of the next customer is less than one minute is $1 - P(\text{no arrivals in one minute}) = 1 - P(X = 0)$. For one minute we use $\lambda = 1.3$, hence:

$$1 - P(X = 0) = 1 - \frac{\mathrm{e}^{-1.3}(1.3)^0}{0!} = 1 - \mathrm{e}^{-1.3} = 1 - 0.2725 = 0.7275.$$

(c) The time to the next customer is more than d if there are no arrivals in the interval from 0 to d, which means that we need to use $\lambda = d \times 1.3$. Now the conditional probability formula yields:

$$P(D > 1.9 \,|\, D > 1) = \frac{P(\{D > 1.9\} \cap \{D > 1\})}{P(D > 1)}$$

and, as in other instances, the two events $\{D > 1.9\}$ and $\{D > 1\}$ collapse to a single event, $\{D > 1.9\}$. Hence:

$$P(D > 1.9 \,|\, D > 1) = \frac{P(D > 1.9)}{P(D > 1)} = \frac{P(D > 1.9)}{0.2725}.$$

To calculate the numerator, use $\lambda = 1.9 \times 1.3 = 2.47$, hence (by the same method as in (iii.):

$$P(D > 1.9) = \frac{\mathrm{e}^{-2.47}(2.47)^0}{0!} = \mathrm{e}^{-2.47} = 0.0846.$$

Hence:

$$P(D > 1.9 \,|\, D > 1) = \frac{P(D > 1.9)}{P(D > 1)} = \frac{0.0846}{0.2725} = 0.3105.$$