Worksheet Sheet 2 - outline solutions

Questions

1. The illusionist Derren Brown famously flipped a coin on camera so that it landed heads ten times in a row; he claimed that this was because of his mind powers, in fact it was because of his patience, he simply kept trying the trick again and again until it worked. It took him nine hours. What is the probability of a coin landing heads ten times in a row? If you flip a coin ten times what is the probability of getting five heads and five tails?

   Solution: The chance of 10 heads is

   $$p(10) = 0.5^{10} = 0.0009765625 \tag{1}$$

   whereas the chance of five heads is

   $$p(5) = \left(\begin{array}{c} 10 \\ 5 \end{array} \right) 0.5^{10} = 0.24609375 \tag{2}$$

2. You are a con person with a crooked coin, it has a 0.25 chance of a head and a 0.75 chance of a harp. You make a bet with someone; if the coin shows a head you pay them two pounds, if it shows a harp they give you one. What is the mean return from a bet; what is the variance?

   Solution: So

   $$\mu = -2 \times 0.25 + 1 \times 0.75 = 0.25 \tag{3}$$

   and if X is your winnings

   $$\langle X^2 \rangle = 4 \times 0.25 + 1 \times 0.75 = 1.75 \tag{4}$$

   so $\sigma^2 = \langle X^2 \rangle - \mu^2 = 1.6875$

3. For the binomial distribution on n trials, prove $\sigma^2 = npq$; this can be done by differenciating Z twice, an easier approach is to differenciate the expression for $\mu = np$:

   $$np = \sum_{r=0}^{n} \left(\begin{array}{c} n \\ r \end{array} \right) rp^r q^{n-r} \tag{5}$$

   Solution: So

   $$np = \sum \sum_{r=0}^{n} \left(\begin{array}{c} n \\ r \end{array} \right) rp^r q^{n-r} \tag{6}$$

   and differentiating both sides and then doing the multiplying and dividing by p or q trick gives

   $$n = \frac{1}{p} \sum_{r=0}^{n} r^2 p_R(r) - \frac{1}{q} \sum_{r=0}^{n} r(n-r) p_R(r) \tag{7}$$

or

$$n = \left(\frac{1}{p} + \frac{1}{q}\right)\langle R^2\rangle - \frac{n}{q}\langle R\rangle \tag{8}$$

and after a bit of algebra and using $\langle R\rangle = np$ we get

$$npq = \langle R^2\rangle - n^2p^2 \tag{9}$$

and we note that the right hand side is $\langle R^2\rangle - \langle R\rangle^2$ which is σ^2.

4. Like the binomial distribution the geometric probability distribution is related to a series of independent trials where each trial has probability p of success and $q = 1-p$ of failure. The geometric probability $p(r)$ is the probability that the rth trial is the first success. It is

$$p(r) = q^{r-1}p \tag{10}$$

It can be shown that

$$\sum_{r=1}^{\infty} p(r) = 1 \tag{11}$$

as it must be. You can assume that here. What is the mean of the geometric probability? As a hint, this is done much as for the binomial expansion.

Solution: So we have

$$1 = \sum_{r=1}^{\infty} q^{r-1}p \tag{12}$$

If we differenciate both sides by p we get

$$0 = \sum_{r=1}^{\infty} q^{r-1} - \sum_{r=1}^{\infty}(r-1)q^{r-2}p \tag{13}$$

or

$$0 = \frac{1}{p}\sum_{r=1}^{\infty} q^{r-1}p - \sum_{r=1}^{\infty}(r-1)q^{r-2}p \tag{14}$$

In the second term set $s = r-1$ to get

$$\sum_{r=1}^{\infty}(r-1)q^{r-2}p = \sum_{s=0}^{\infty} sq^{s-1}p = \sum_{s=1}^{\infty} sq^{s-1}p \tag{15}$$

where we are able to change drop the $s = 0$ term in the sum because it gives zero. Now back to the original equation:

$$0 = \frac{1}{p}\sum_{r=1}^{\infty} q^{r-1}p - \sum_{s=1}^{\infty} sq^{s-1}p \tag{16}$$

Now note that the first sum gives one and the second gives μ so

$$\mu = \frac{1}{p} \tag{17}$$

We weren't asked to prove

$$\sum_{r=1}^{\infty} p(r) = 1 \tag{18}$$

but in case you are interested it is discussed here

$$\sum_r q^{r-1} p = 1 \tag{19}$$

For this we use the expansion

$$\frac{1}{1-q} = 1 + q + q^2 + \dots \tag{20}$$

you can think of this as coming from the binomial expansion of $(1-q)^{-1}$ using the generalized binomial formula discovered by Newton

$$(1+x)^n = \sum_r \frac{n(n-1)\dots(n-r+1)}{r!} x^r \tag{21}$$

with $n = -1$; it can also be derived as the Taylor expansion. Either way we have

$$\frac{1}{1-q} = \sum_{r=0}^{\infty} q^r = \sum_{r=1}^{\infty} q^{r-1} \tag{22}$$

and since $1 - q = p$ this gives the result.

Extra questions

1. Oranmore, the village I grew up in had more people with the surname Furey than any other village or town in the world. Since I left the village has expanded ten-fold and has gone from being a small village to a commuter town for Galway. However, when I was young one in ten people in the village had the surname Furey. Imagine there are 35 children in a class at school, what is the probability that five of them were Fureys?

   Solution: So this is just

   $$p(r=5) = \begin{pmatrix} 35 \\ 5 \end{pmatrix} \left(\frac{1}{10}\right)^5 \left(\frac{9}{10}\right)^{30} \tag{23}$$

   which is about 0.14.

2. The **Fano factor** is sometimes used to describe distributions, it is

   $$F = \frac{\sigma^2}{\mu} \tag{24}$$

   What is the Fano factor for the Poisson distribution?

   Solution: Well the mean is np and the variance is npq so $F = q$.

3. What is the variance of the geometric distribution?

   Solution: Well we have
   $$\frac{1}{p} = \sum_{r=0}^{\infty} r q^{r-1} p \tag{25}$$
   so differentiating both sides gives
   $$-\frac{1}{p^2} = -\sum_{r=0}^{\infty} r(r-1) q^{r-2} p + \sum_{r=0}^{\infty} r q^{r-1} \tag{26}$$
   and so with a bit of algebra we get
   $$-\frac{1}{p^2} = -\frac{1}{q}\langle R^2\rangle + \left(\frac{1}{q} + \frac{1}{p}\right)\langle R\rangle \tag{27}$$
   and substituting for $\langle R\rangle = 1/p$ and doing still more algebra we get
   $$\sigma^2 = \frac{q}{p^2} \tag{28}$$

4. The negative binomial distribution models a binomial process where, instead of n trials, the experiment continues until there are r failures. The probability of k successes is
   $$p_K(k) = \begin{pmatrix} k+r-1 \\ k \end{pmatrix} (1-p)^r p^k \tag{29}$$
   Obviously if
   $$Z = \sum_{k=0}^{\infty} p_K(k) \tag{30}$$
   then $Z = 1$; this relies on the negative binomial expansion and you don't need to do this calculation. By differenciating calculate the mean of the negative binomial distribution.

   Solution: It's just the usual craic, we differentiate Z, for the same of space lets write
   $$C_k = \begin{pmatrix} k+r-1 \\ k \end{pmatrix} \tag{31}$$
   and we get
   $$0 = \sum_{k=0}^{\infty} C_k k (1-p)^r p^{k-1} - \sum_{k=0}^{\infty} C_k r (1-p)^{r-1} p^k \tag{32}$$
   and so
   $$\frac{1}{p}\sum_{k=0}^{\infty} C_k k q^r p^k = \frac{1}{q}\sum_{k=0}^{\infty} C_k r q^r p^k \tag{33}$$
   or
   $$\langle K\rangle = \frac{pr}{q} \tag{34}$$