4 Topics: Random variables and their distributions

4.1 Prerequisites: Lecture 10

Exercise 4- 1: (Suggested for personal/peer tutorial) Poisson approximation to the Binomial: If $X \sim \text{Bin}(n,p)$ and we have $n \to \infty$ and $p \to 0$ such that $\lambda = np$ remains constant, then the p.m.f. of X converges to the p.m.f. of a Poi(λ) random variable.

Hint: Use the result that for all $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \left(1 - \frac{t}{n}\right)^n = e^{-t}.$$

Remark: The same result holds, when for $n \to \infty$ and $p \to 0$, we have that np converges to a positive constant λ. In that case, we use the result, that for a sequence (t_n) converging to t when $n \to \infty$, we have

$$\lim_{n\to\infty} \left(1 - \frac{t_n}{n}\right)^n = e^{-t}.$$

Solution: Consider the case when $\lambda = np$ is fixed when $n \to \infty$ and $p \to 0$. Let $0 \le k \le n$, then

$$\begin{aligned}
\mathrm{P}(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \frac{1}{k!} n(n-1)\cdots(n-k+1) \frac{n^k}{n^k} p^k (1-p)^{n-k} \\
&= \frac{\lambda^k}{k!} n(n-1)\cdots(n-k+1) \frac{1}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n} \left(1 - \frac{\lambda}{n}\right)^{-k}
\end{aligned}$$

Then, for fixed k:

$$\begin{aligned}
\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} &= \lim_{n\to\infty} 1 \cdot (1 - 1/n) \cdots (1 - (k-1)/n) = 1, \\
\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n &= e^{-\lambda}, \\
\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-k} &= 1.
\end{aligned}$$

Hence

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}, \text{ as } n \to \infty,$$

where the right hand side is indeed the p.m.f. of a Poisson random variable with parameter λ.

The proof of the remark goes as follows. As above, we write

$$\begin{aligned}
\mathrm{P}(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \frac{1}{k!} n(n-1)\cdots(n-k+1) \frac{n^k}{n^k} p^k (1-p)^{n-k}
\end{aligned}$$

$$= \frac{(np)^k}{k!} n(n-1)\cdots(n-k+1)\frac{1}{n^k}\left(1-\frac{np}{n}\right)^{n-k}$$
$$= \frac{(np)^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{np}{n}\right)^{n}\left(1-\frac{np}{n}\right)^{-k}$$

Note that in our set-up, p is necessarily a function of n, so we could write $p = p_n$ to stress that. We shall now define the sequence $t_n := np_n$. Recall that we assume that

$$\lim_{n\to\infty} np = \lim_{n\to\infty} np_n = \lim_{n\to\infty} t_n = \lambda < \infty.$$

This implies that $\lim_{n\to\infty} p_n \to 0$. Then, for fixed k:

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} = \lim_{n\to\infty} 1\cdot(1-1/n)\cdots(1-(k-1)/n) = 1,$$
$$\lim_{n\to\infty}\left(1-\frac{np}{n}\right)^n = \lim_{n\to\infty}\left(1-\frac{t_n}{n}\right)^n = e^{-\lambda},$$
$$\lim_{n\to\infty}(1-p_n)^{-k} = 1.$$

Hence

$$\mathrm{P}(X=k) = \binom{n}{k} p^k(1-p)^{n-k} \to \frac{\lambda^k}{k!}e^{-\lambda}, \text{ as } n\to\infty,$$

Exercise 4- 2: A company wishes to make two of a group of six employees, comprising three female and three male employees, redundant, by selecting two employees at random. Let X and Y be the random variables corresponding to the number of female and male employees made redundant, respectively.

Find the probability mass functions of X and Y.

Solution: For both variables, the range is $\{0, 1, 2\}$, and distribution is given by Hypergeometric formula with $N = 6$, $K = 3$ and $n = 2$. Hence

$$p_X(x) = p_Y(x) = \frac{\binom{3}{x}\binom{3}{2-x}}{\binom{6}{2}} \qquad x = 0, 1, 2$$

and zero otherwise.

Exercise 4- 3: Five balls numbered 1,2,3,4 and 5 are placed in a bag. Two balls are selected without replacement. Find the probability mass function of the following random variables:

(a) $X =$ the largest of the two selected numbers,

(b) $Y =$ the sum of the two selected numbers

Solution:

(a) Range $\mathrm{Im}X = \{2, 3, 4, 5\}$. Now $p_X(x) = \mathrm{P}(X = x) = \mathrm{card}(E)/\mathrm{card}(\Omega)$, say, and

$\mathrm{card}(E)$= "number of ways of choosing two from five with largest equal to x" $= x - 1$, $\mathrm{card}(\Omega)$ ="number of ways of choosing two from five"=$\binom{5}{2} = 10$. So $p_X(x) = \mathrm{P}(X = x) = (x-1)/10$.

(b) Range $\text{Im}Y = \{3,4,5,6,7,8,9\}$. As above, define $p_Y(y) = \text{P}(Y = y) = \text{card}(E)/\text{card}(\Omega)$, say, and again $\text{card}(\Omega) = \binom{5}{2} = 10$. Enumeration of $\text{card}(E)$ achieved by considering distinguishable partitions of y into the sum of two integers in the range $\{1,2,3,4,5\}$. Hence if $y = 3,4,8,9$, $\text{card}(E) = 1$, but if $y = 5,6,7$, $\text{card}(E) = 2$, so

$$p_Y(y) = \begin{cases} 1/10 & y = 3,4,8,9 \\ 2/10 & y = 5,6,7 \end{cases}$$

Exercise 4- 4: A surgical procedure is successful with probability θ. The surgery is carried out on five patients, with the success or failure of each operation independent of all other operations. Let X be the discrete random variable corresponding to the number of successful operations.

Find the probability mass function of X, and evaluate the probability that

(a) all five operations are successful, if $\theta = 0.8$,

(b) exactly four operations are successful, if $\theta = 0.6$,

(c) fewer than two are successful, if $\theta = 0.3$.

Solution: $X \sim \text{Bin}(n, \theta)$, so

(a) $\theta = 0.8, \text{P}(X = 5) = 0.3227$

(b) $\theta = 0.6, \text{P}(X = 4) = 0.2592$

(c) $\theta = 0.3, \text{P}(X < 2) = \text{P}(X = 0) + \text{P}(X = 1) = 0.5282$

Exercise 4- 5: If X has a Geometric distribution with parameter θ, so that

$$p_X(x) = (1-\theta)^{x-1}\theta, \qquad x = 1,2,3,...$$

and zero otherwise, show that, for $n, k \geq 1$,

$$\text{P}(X = n + k | X > n) = \text{P}(X = k).$$

This result is known as the *Lack of Memory* property (for a discrete random variable).

Solution: If $X \sim \text{Geo}(\theta)$, then

$$p_X(x) = (1-\theta)^{x-1}\theta,$$

and

$$\text{P}(X \leq x) = 1 - (1-\theta)^x, \text{ for } x \in \{1,2,3...\}.$$

Thus $\text{P}(X > n) = (1-\theta)^n$, and hence

$$\begin{aligned}\text{P}(X = n + k | X > n) &= \frac{\text{P}(X = n + k, X > n)}{\text{P}(X > n)} = \frac{\text{P}(X = n + k)}{\text{P}(X > n)} = \frac{(1-\theta)^{n+k-1}\theta}{(1-\theta)^n} \\ &= (1-\theta)^{k-1}\theta = \text{P}(X = k).\end{aligned}$$

4.2 Prerequisites: Lecture 11

Exercise 4- 6: Suppose $X \sim \mathrm{DUnif}(\{1, \ldots, n\})$. Find the c.d.f. of X.

Solution: We have $\mathrm{P}(X = x) = 1/n$ for $x \in \{1, \ldots, n\}$ and zero otherwise. Hence

$$F_X(x) = \mathrm{P}(X \leq x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{\lfloor x \rfloor}{n}, & \text{if } 0 \leq x < n, \\ 1, & \text{if } x \geq n, \end{cases}$$