PROBLEM1

Tyagi Mahir

2022-11-16

Suppose X denote the number of goals scored by home team in premier league. We can assume X is a random variable. Then we have to build the probability distribution to model the probability of number of goals. Since X takes value in $\mathbb{N} = \{0, 1, 2, \cdots\}$, we can consider the geometric progression sequence as possible candidate model, i.e.,

$$S = \{a, ar, ar^2, ar^3, \cdots\}.$$

QUESTION1
Here, we have:

$$\sum\nolimits_{x=0}^{\infty} ar^x = 1$$

or:

$$a/r - 1 = 1$$

hence a = r-1. The necessary conditions are $a > 0$ and $r < 1$.
QUESTION 2 & 3 The mean is:

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} xP(X = x)$$

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} x(r-1)r^x$$

$$E[X] = (1-r)(0 + r + 2r^2 ...)..(i)$$

$$rE[X] = (1-r)(0 + r^2 + 2r^3 ...)..(ii)$$

Subtracting (ii) from (i):

$$(1-r)E[X] = (1-r)(0 + r^2 + r^3..)$$

i.e:

$$E[X] = \frac{r}{1-r}$$

Also for variance, we note that the MGF of this distribution will be $a\sum(e^t r)^x = \frac{1-r}{1-e^t r}$. Differentiating twice we get:

$$E[X^2] = [\frac{(r^2(e^t) + r(e^t))(1-r)^2}{(1-e^t r)^4}]at \;\; t = 0$$

$$V[X] = E[X^2] - E[X]^2$$

$$= \frac{r^2 + r}{(1-r)^2} - (\frac{r}{1-r})^2$$

$$= \frac{r}{(1-r)^2}$$

Hence both mean and variance exist. We knew this from the fact that this is a geometric distribution.
QUESTION 4
a) $P(X >= 1) = 1 - P(X = 0) = 1 - (1 - r) = 0.6$.

b) $P(0 < X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= (1 - r)(r + r^2 + r^3) = (1 - r)[(1 - r^4)/(1 - r) - 1] = 0.348$$

QUESTION 5

Since we're taking an "off the shelf" model, we'll take $\lambda = 1.5$
a)$P(X >= 1) = 1 - P(X = 0) = 1 - exp(-\lambda) = 1 - 0.223 = 0.777.$
b)$P(0 < X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= exp(-\lambda)(\lambda + \lambda^2/2 + \lambda^3/6) = 0.710$$

QUESTION 6

If our model's mean is taken to be 1.5, $r = 1.5(1 - r) \implies r = 0.6 \implies var(x) = 3.75$, which is far off. Stats like median approximation of poisson=1 whereas median approximation of geometric is 2 (sample median is 1) convince us that poisson is a better model.

QUESTION 7

For Geometric distribution we have:-

$$\mathcal{L}(r|x_1, x_2, x_3, ..., x_n) = \prod_{i=1}^{n}(1 - r)r^{x_i}$$
$$\mathcal{L}(r|x_1, x_2, x_3, ..., x_n) = (1 - r)[r^{x_1}.r^{x_2}.r^{x_3}...r^{x_n}]$$
$$\mathcal{L}(r|x_1, x_2, x_3, ..., x_n) = (1 - r)r^{\sum_{i=1}^{n} x_i}$$

For Poisson distribution we have:-

$$\mathcal{L}(\lambda|x_1, x_2, x_3, ..., x_n) = \prod_{i=1}^{n}\left(\frac{e^{-\lambda}\lambda^{k_i}}{k_i!}\right)$$
$$\mathcal{L}(\lambda|x_1, x_2, x_3, ..., x_n) = \frac{e^{-\lambda}\lambda^{k_1}}{k_1!} \cdot \frac{e^{-\lambda}\lambda^{k_2}}{k_2} \cdot \frac{e^{-\lambda}\lambda^{k_3}}{k_3!} ... \frac{e^{-\lambda}\lambda^{k_n}}{k_n!}$$
$$\mathcal{L}(\lambda|x_1, x_2, x_3, ..., x_n) = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} k_i}}{\prod_{i=1}^{n} k_i!}$$