Assignment 2: Question 1

Name: Abhishek Sharma Roll Number: MDS202204

PART-1:

The necessary conditions are:

1. $P(X = i) = a * r^i, i = 1, 2, 3 \ldots$
2. $a \leq 1$
3. $0 \leq r \leq 1$
4. $a = 1 - r$

PART-2:

The idea is to use ratio test on E$[X]$ and Var(X) :

$$E[X] = \sum_{n=0}^{\infty} n * a * r^n = \sum_{n=0}^{\infty} a_n$$

now,

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{(n+1) * a * r^{n+1}}{(n) * a * r^n} = r < 1$$

So, by ratio test, E$[X]$ exists!

$$\text{Var}[X] = \sum_{n=0}^{\infty} (n - E[X]) * a * r^n = \sum_{n=0}^{\infty} a_n$$

now,

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{(n+1-E[X]) * a * r^{n+1}}{(n-E[X]) * a * r^n} = r < 1$$

So, by ratio test, Var$[X]$ exists!

PART-3:

$$E[X] = \sum_{n=0}^{\infty} n * a * r^n = a * r \left[\frac{d}{dr} \sum_{n=0}^{\infty} r^n \right] = \frac{r}{1-r}$$

$$E[X^2] = \sum_{n=0}^{\infty} n^2 * a * r^n = a * r * \frac{d}{dr}\left[r \left[\frac{d}{dr} \sum_{n=0}^{\infty} r^n \right]\right] = \frac{r(1+r)}{(1-r)^2}$$

$\implies \text{Var}(X) = \dfrac{r}{(1-r)^2}$

PART-4:

From the data, we can see that sample mean is 1.5 and the number of data points is 380. Hence, for method of moments estimate of r, we have

$$\frac{\hat{r}}{(1-\hat{r})} = 1.5 \implies \hat{r} = 0.6$$

$$\implies P(X \geq 1) = 1 - P(X = 0) = \hat{r} = 0.6$$

and

$$P(1 \leq X < 4) = P(X = 1) + P(X = 2) + P(X = 3) = (1-\hat{r})(\hat{r})(1 + \hat{r} + \hat{r}^2) = 0.4704$$

PART-5:

Now, we assume that $X \sim Poisson(\lambda)$

From the data, we can see that sample mean is 1.5 and the number of data points is 380. Hence, the method of moments estimate for λ is

$$\hat{\lambda} = 1.5$$

$$\implies P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\hat{\lambda}} = 0.7768$$

and

$$P(1 \leq X < 4) = P(X = 1) + P(X = 2) + P(X = 3) = 0.7112$$

PART-6:

We have to find a model for number of goals scored in a fixed amount of time, which can be thought of as number of arrivals in a fixed amount of time. For such a scenario, i believe, it is best to use a possion probability model.

PART-7:

Let $f(r)$ be the probability model defined using the geometric sequence $S = \{a, ar, ar^2, \cdots\}$

$$\implies \text{ if } X_1, X_2, \cdots X_n \overset{i.i.d}{\sim} f(r) \text{ then, likelihood of r is } L(r) = \prod_{i=1}^{n} ar^{X_i}$$

and

$$\implies \text{ if } X_1, X_2, \cdots X_n \overset{i.i.d}{\sim} Poisson(\lambda) \text{ then, likelihood of } \lambda \text{ is } L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{X_i}}{X_i!}$$