PBSR_QUESTION_NO: 1

AKASH

2022-11-16

Question 1

```
library(tinytex)
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 4.2.2
```

```
## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output
## %in% : 'length(x) = 2 > 1' in coercion to 'logical(1)'
```

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 N = {0, 1, 2, · · · }, we can consider the geometric progression sequence as possible candidate model, i.e

$$S = \left\{a, ar^2, ar^3, \cdot\cdot\cdot\cdot\cdot\right\}$$

But we have to be careful and put proper conditions in place and modify S in such a way so that it becomes proper probability distributions

1. Figure out the necesary conditions and define the probability distribution model using S.

```
x=c(0,1,2,3)
Probability=c("a","ar","ar^2","ar^3")
Distribution_table=data.frame(x,Probability)
```

```
kable(Distribution_table,"pipe",align=c("r","r"),table.attr = "class=\"striped\"")
```

x	Probability
0	a
1	ar
2	ar^2
3	ar^3

under the consideration $|r| <= 1$

$$\sum_{k=0}(ar^k) = 1$$
$$\frac{a}{1-r} = 1$$
$$a = 1-r$$
$$a+r = 1$$

2)Check if mean and variance exists for the probability model

To check the existance of $E(X)$:(By Usiing Dalemberts Ratio test)

$$E(x) = \sum_{k=0}^{inf} k(ar^k)$$
$$\lim_{n->inf} \frac{(n+1))ar^(n+1)}{(n)ar^n}$$
$$= r * \lim_{n->inf}(1+\frac{1}{n})$$
$$= r$$

We already know that $|r| <= 1$ .Thus through D alemberts ratio test for infinite series we can say that $E(X)$ exists

To check the existance of $E(X^2)$:(By Usiing Dalemberts Ratio test)

$$E(x^2) = \sum_{k=0}^{inf} k^2(ar^k)$$
$$\lim_{n->inf} \frac{(n+1)^2ar^(n+1)}{(n)^2ar^n}$$
$$= r * \lim_{n->inf}(1+\frac{1}{n^2})$$
$$= r$$

We already know that $|r| <= 1$ .Thus through D alemberts ratio test for infinite series we can say that $E(X^2)$ exists

Since Both $E(X^2)$ and $E(X)$ exists $V(X) = E(X^2) - (E(X))^2$ also exists.(Algebra of Convergent Series)

3)Can you find the analytically expression of mean and variance

$E(X)$:

$$E(x) = \sum_{k=0}^{inf} k(ar^k)$$
$$= ar + 2a(r^2) + 3a(r^3) + \cdots$$
$$= ar(1-r)^-2$$

$E(X^2)$:

$$
\begin{aligned}
E(x^2) &= \sum_{k=0}^{inf} k^2(ar^k) \\
&= ar + 4(ar^2) + 9(ar^3) + \cdots \\
&= ar(1+r)(1-r)^-3
\end{aligned}
$$

$V(X) = E(X^2) - (E(X))^2$

$$
V(X) = ar(1+r)^-3 - a^2r^2(1-r)^-4
$$

4)From historical data we found the following summary statistics Using the summary statistics and your newly defined probability distribution model find the following: a. What is the probability that home team will score at least one goal?

b. What is the probability that home team will score at least one goal but less than four goal?

We have three equations

$$
\begin{aligned}
V(X) = 9/4 \\
E(X) = 3/2 \\
a + r = 1
\end{aligned}
$$

From the three equations we compute the value of a=2/5 and r=3/5

Probability that the team will score atleast one goal

```
geom=function(a,r,n){
  ans=0
  for (i in 0:n){
    ans=ans+a*(r)^i
  }
  return (ans)
}
```

```
1-geom(2/5,3/5,0)
```

```
## [1] 0.6
```

Probability that the team will score atleast one goal but less than 4 goals

```
geom(2/5,3/5,3)-geom(2/5,3/5,0)
```

```
## [1] 0.4704
```

5. Suppose on another thought you want to model it with off-the shelf Poisson probability models. Under the assumption that underlying distribution is Poisson probability find the above probabilities, i.e.,

a. What is the probability that home team will score at least one goal?
b. What is the probability that home team will score at least one goal but less than four goal?

```
pois=function(lambda,n){

  ans=0
  for (i in 0:n){
    ans=ans+((exp(-lambda))*(lambda)^i)/factorial(i)
  }
  return (ans)
}
```

```
1-pois(3/2,0)
```

```
## [1] 0.7768698
```

```
pois(3/2,3)-pois(3/2,0)
```

```
## [1] 0.7112274
```

6. Which probability model you would prefer over another?

I would prefer the geometric model over the poisson model because n ie the no of goals scored cannot be infinitely large and as calculated the r is not indefinitely small.

7.Write down the likelihood functions of your newly defined probability models and Poisson models. Clearly mention all the assumptions that you are making.

Let, $x_1, x_2, x_3, ..., x_n$ are "n" random samples from our newly defined model. Now the likelihood function (L1) of this model is given below :-

$$
\begin{aligned}
L1(x) &= \prod_{i=1}^{n} f(x_i) \\
&= \prod_{i=1}^{n} ar^{x_i} \\
&= a^n r^{\sum_{i=i}^{n} x_i}
\end{aligned}
$$

Assumptions :

1> For this model we assume that $|r| <= 1$.

2> $x_1, x_2, x_3, ..., x_n$ are independent.

3> $x_1, x_2, x_3, ..., x_n$ are identical.

Now for the Poisson model,

Let, $x_1, x_2, x_3, ..., x_n$ are "n" random samples from the Poisson model. Now the likelihood function (L2) of this model is given below :-

$$
\begin{aligned}
L2(x) &= \prod_{i=1}^{n} f(x_i) \\
&= \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \\
&= e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}
\end{aligned}
$$

Assumptions :

1> $x_1, x_2, x_3, \ldots.., x_n$ are independent.

2> $x_1, x_2, x_3, \ldots.., x_n$ are identical.