Discrete Distributions

John S Butler (TU Dublin)

Introduction

Binomial Distribution

A Binomial Distribution is used to describe the probability distribution of n experiments that have two possible outcomes "success" or "failure". The probability of "success" is p, the probability of failure is $q = 1-p$. This gives the general deifntion of the distribution as:

$$\Pr(k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k}, \quad k = 0, 1, 2, ...n,$$

where k is the number of "successes". If the data can be described using a Binomial Distribution then the expected outcome, $E[k]$ of the distribution is given by:

$$E[k] = np,$$

and the variance of the distribution $Var[k]$, is

$$Var[k] = npq.$$

Example

The example we shall use to illustrate the discrete distibutions is the New Zealand vs Ireland Wolrd Cup Rugby Quarter Final. Let's assume the probability of "success", Ireland beating New Zealand is

$$p = 0.15,$$

the probability of failure is

$$q = 1 - p = 1 - 0.15 = 0.85.$$

Now imagine that Ireland have to play New Zealand 10 times, this gives the Binomial Distribution

$$\Pr(k) = \left(\begin{array}{c} 10 \\ k \end{array} \right) (0.15)^k (0.85)^{10-k}, \quad k = 0, 1, 2, ...10,$$

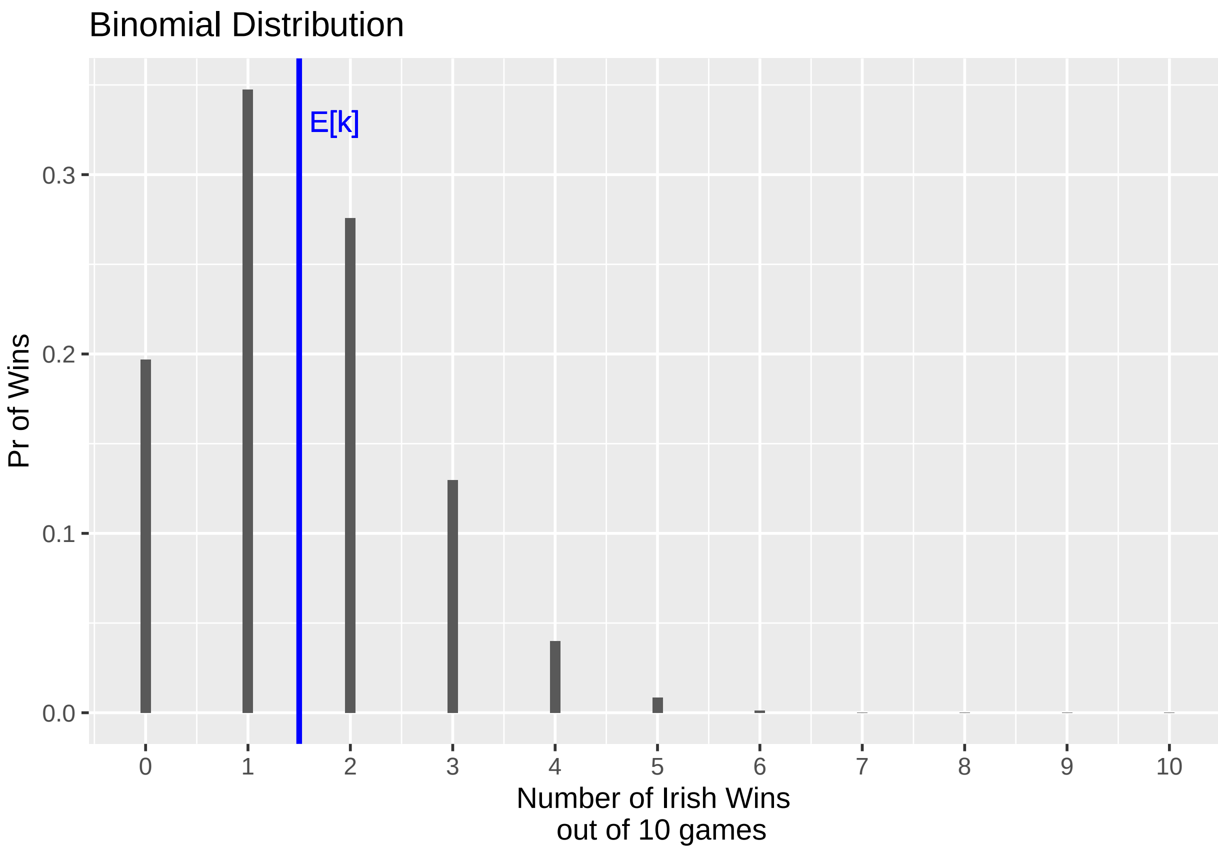
where k is the number of times Ireland beats New Zealand. As it is a binomial Distribution we can state that the expected number of times Ireland will win is

$$E[k] = 10 \times 0.15 = 1.5,$$

ie Ireland is only expected to win once, the variance of the distribution is

$$Var[k] = 10 \times 0.15 \times 0.85 = 1.35.$$

```
## Warning: Ignoring unknown parameters: text
```



Geometric Distribution

A Geometric distribution
is used to describe the probability distribution if you do an experiment until you succeed, the experiment has e two possible outcomes "success" or "failure". The probability of "success" is p, the probability of failure is $q = 1 - p$. This gives the general deifintion of the distribution as:

$$\Pr(k) = q^{(k-1)}p, \quad k = 1, 2, ...$$

with the expected outcome of,

$$E[k] = \frac{1}{p},$$

and variacne of

$$Var[k] = \frac{q}{p^2}.$$

Example

The example we shall use to illustrate the Geometric distibution is the New Zealand vs Ireland Wolrd Cup Rugby Quarter Final. Let's assume the probability of "success", Ireland beating New Zealand is

$$p = 0.15,$$

the probability of failure is

$$q = 1 - p = 1 - 0.15 = 0.85.$$

Now imagine that Ireland have to play New Zealand until Ireland wins this gives the Geometric Distribution

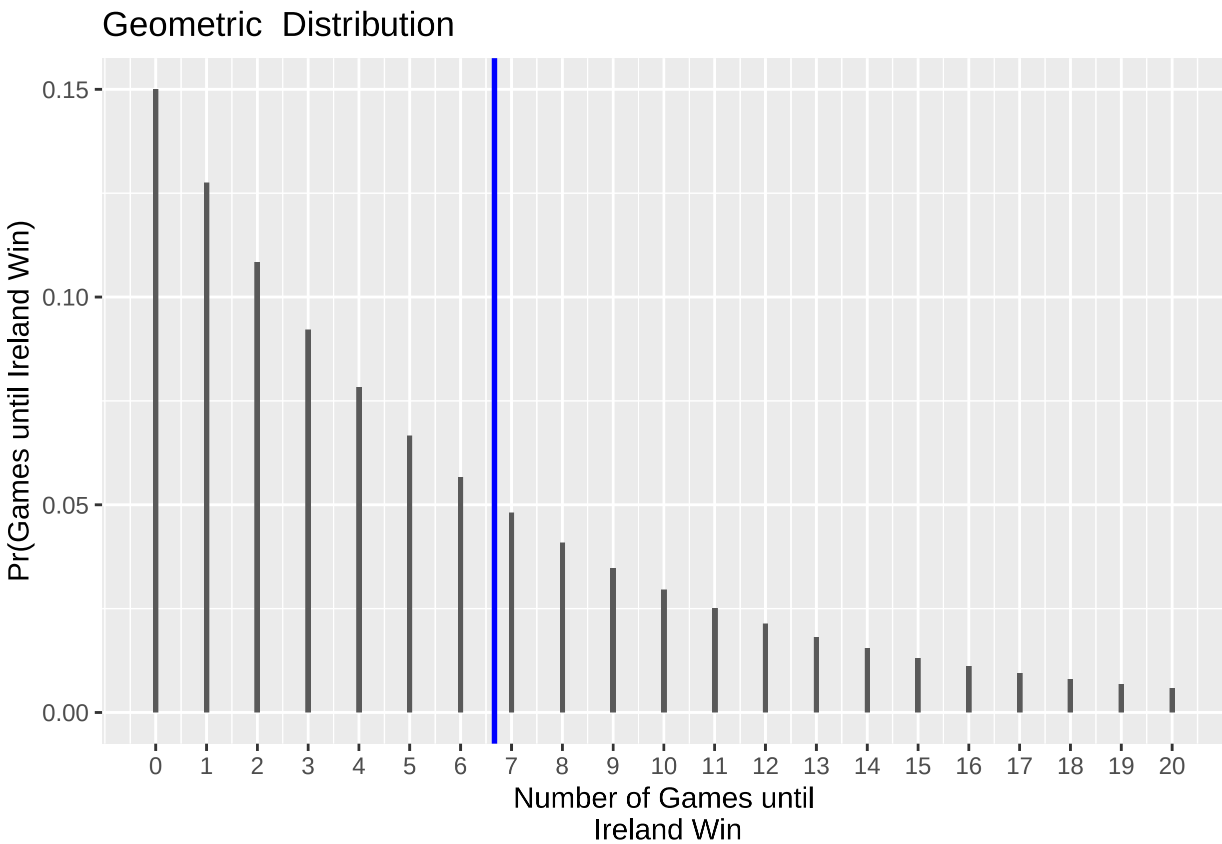
$$\Pr(k) = (0.85)^{k-1}0.15, \quad k = 0, 1, 2, \cdots$$

where k is the number of games played until Ireland beats New Zealand. As it is a Geometic Distribution we can state that the expected number of games played until Ireland will win is

$$E[k] = \frac{1}{0.15} = 6.6667,$$

ie it will take this many games on average for Ireland to win once, the variance of the distribution is

$$Var[k] = \frac{0.85}{0.15^2}.$$



Poisson Distribution

A Poisson distribution
is used to describe the probability distribution of the average number of events in a specific time period or distance. The distribution is described by the average, λ,

$$\Pr(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, ...$$

where k is the number of times the event happens. The expeceted value of the Poisson Distribution is

$$E[k] = \lambda,$$

with the variance

$$Var[k] = \lambda.$$

Example

The example we shall use to illustrate the Poisson distibution is the New Zealand vs Ireland Wolrd Cup Rugby Quarter Final. Let's assume that over a 10 year period Ireland win on average $\lambda = 2.2$ games. This gives the Poisson distribution

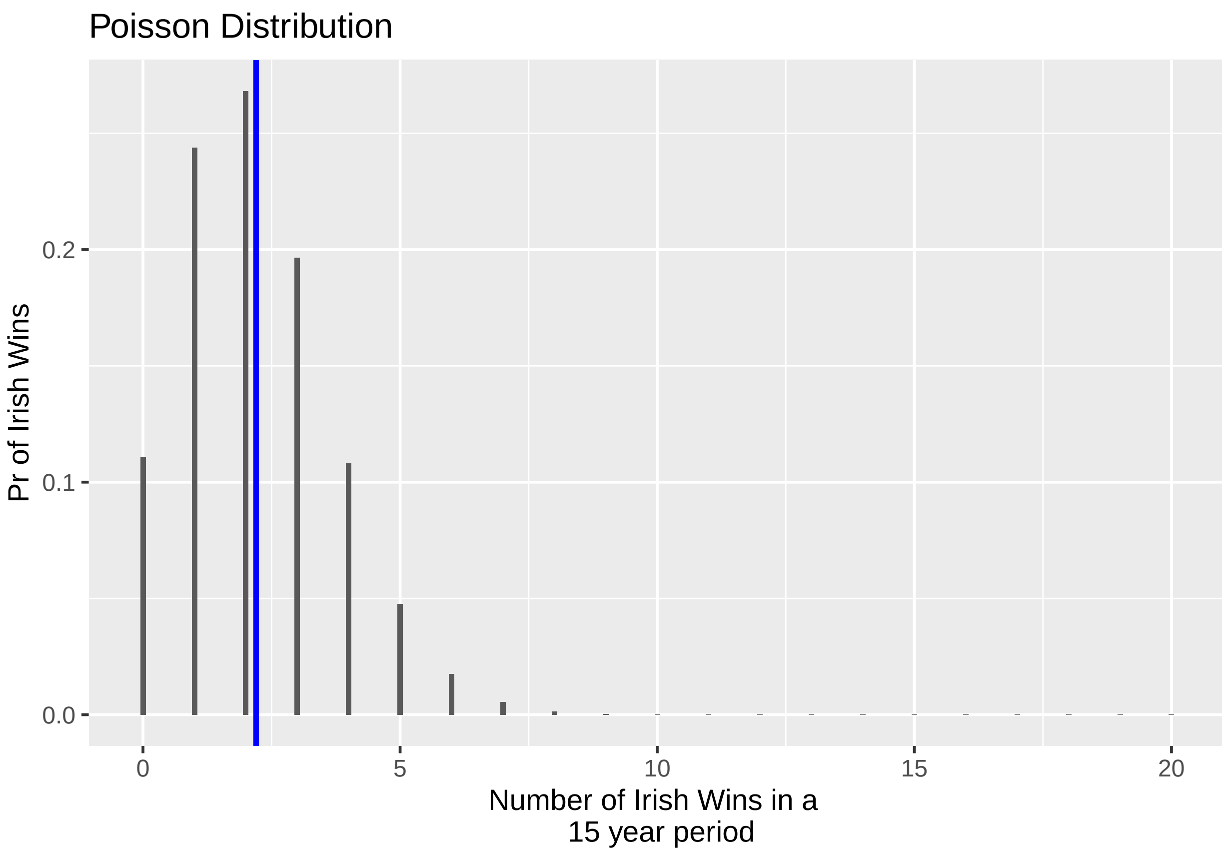
$$\Pr(k) = \frac{2.2^k e^{-2.2}}{k!}, \quad k = 0, 1, 2, ...$$

where k is the number of games Ireland beats New Zealand in a 10 year period that Ireland will win is

$$E[k] = 2.2,$$

the variance of the distribution is

$$Var[k] = 2.2.$$



Normal Distribution

Definition:

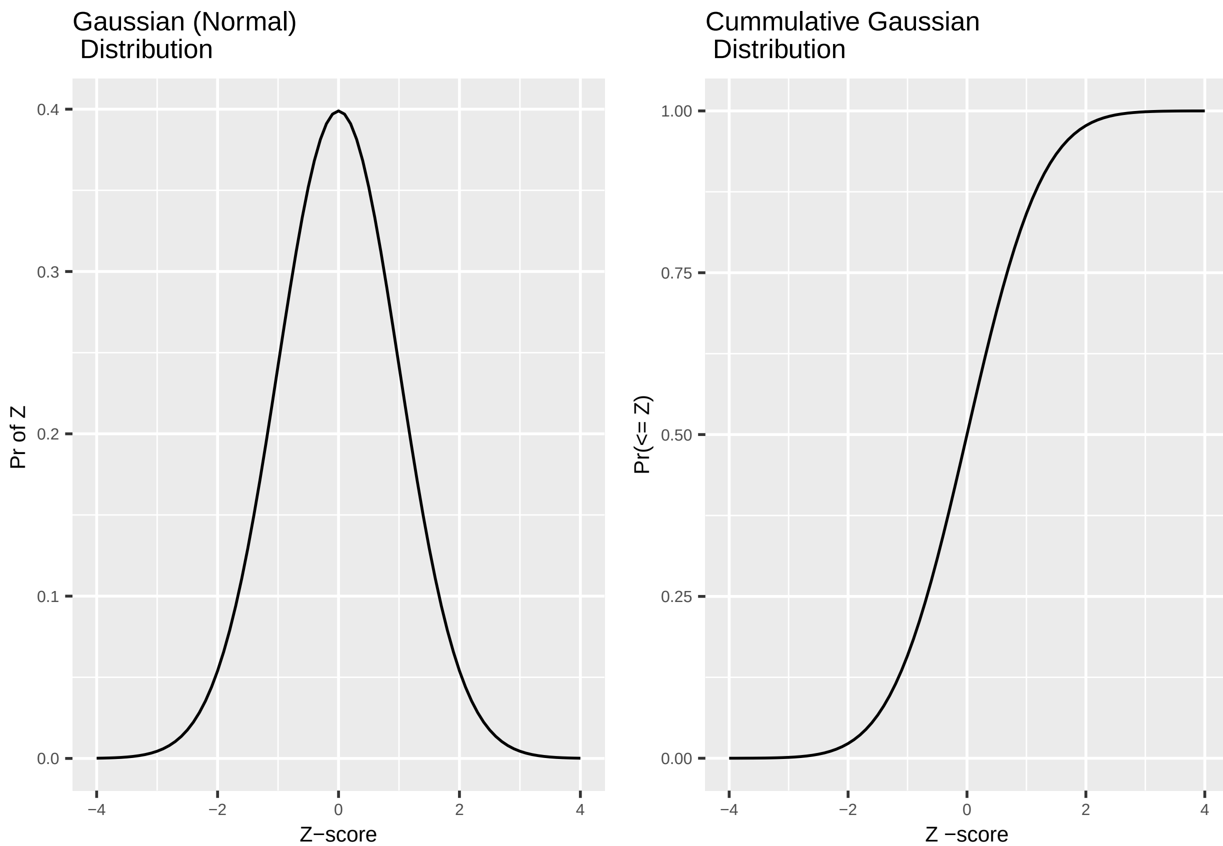
$$\Pr(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, ...$$

$$E[k] = \lambda, \quad Var[k] = \lambda.$$

In R code:

```
## Warning: Ignoring unknown parameters: width
```

```
## Warning: Ignoring unknown parameters: width
```


Gaussian (Normal)
Distribution
Pr of Z
0.4
0.3
0.2
0.1
0.0
−4
−2
0
2
4
Z−score
Cummulative Gaussian
Distribution
Pr(<= Z)
1.00
0.75
0.50
0.25
0.00
−4
−2
0
2
4
Z −score

Confidence Intervals