THEORY OF STATISTICS

EDWARD O'CALLAGHAN

Contents

1. Bernoulli Distribution.

$$\boxed{X \sim Bernoulli(p) : 0 \leq p \leq 1} \tag{1}$$

Take some random variable X with a Bernoulli parametric model with success parameter p defined by:

$$X \triangleq \begin{cases} 1 & \text{if success } p, \\ 0 & \text{otherwise.} \end{cases}$$

Then, $X \sim Bernoulli(p)$ read as; "X has distribution Bernoulli with parameter p". The probability mass function is then defined by:

$$\boxed{f_X(x;p) \triangleq p^x(1-p)^{1-x} : x \in \{0,1\}.} \tag{2}$$

2. Binomial Distribution.

$$\boxed{X \sim Bin(n,p) : 0 \leq p \leq 1} \tag{3}$$

This is a finite generalisation of a series of Brnoulli random variables with replacement.

Take some random variable X with each observation $x_i \in X : 0 \leq i \leq n$ where n is the number of trials and that each $x_i \sim Bernoulli(p)$. Then, provided all the observations x_i are independent and since the parameter p is the same (a special case of the Poisson binomial distribution) for all. Then, we claim that the each x_i are "independent and identically distributed" **i.i.d**. Assuming now, we select each x_i with replacement, then there are $\binom{n}{x}$ ways to select x observations from a trial, or sample size of n.

That is, $X \sim Bin(n,p) : 0 \leq p \leq 1$ and hence has probability mass function defined by:

$$\boxed{f_X(x;n,p) \triangleq \binom{n}{x} p^x(1-p)^{n-x} : 0 \leq x \leq n.} \tag{4}$$

Clear, this is just an extension of a Bernoulli trial sequentially repeated n times.

3. Geometric Distribution.

$$\boxed{X \sim Geom(p) : 0 \leq p \leq 1} \tag{5}$$

This is a special case result of a Binomial random variable, for when $x = 1$.

Take some random variable X such that $X \sim Bin(n,p)$, given that p is the probability of success. Take $x = 1$ with $n = k$ so that, the probability of first success on the k^{th} trial is given by,

$$f_X(1;k,p) = \binom{k}{1} p^1(1-p)^{k-1}.$$

Then we have the following special case result as our probability mass function:

$$\boxed{f_X(k;p) \triangleq p(1-p)^{k-1} : k \in \{1,2,\dots\}.} \tag{6}$$

4. Poisson Distribution.

$$\boxed{X \sim Poisson(\lambda)} \tag{7}$$

This is another special case result of a Binomial random variable, for when p is close to zero and n is very large[Pap91].

Take some random variable Y such that $Y \sim Bin(n,p)$ and fix some parameter $\lambda \doteq np$ where n is known to be large and p is known to be small. Such a case is intuitively a rare event since the probability of success p is small for a large number of trials n. This is know as "the law of rare events"[Sem]. We continue with the assumtion that such events are sufficiently independent. Then, we have a limiting case of the Binomial random variable Y with parameter λ.

Hence if,

$$Y \sim Bin(n,p)$$

Substituting for λ and by introduction of an auxiliary random variable X

$$X \sim Bin(n, \frac{\lambda}{n})$$

we derive the probability mass function as follows,

$$\begin{aligned} f_X(x;\lambda) &= \frac{n!}{(n-x)!x!}(\frac{\lambda}{n})^x(1-\frac{\lambda}{n})^{n-x} \\ &= \frac{n!}{n^x(n-x)!}\frac{1}{(1-\frac{\lambda}{n})^x}\frac{\lambda^x}{x!}(1-\frac{\lambda}{n})^n \end{aligned}$$

Taking the limit as $n \to \infty$ and noting the definition of $e^{-\lambda} \doteq \lim_{n\to\infty}(1-\frac{\lambda}{n})^n$

$$\begin{aligned} &\frac{\lambda^x}{x!}\lim_{n\to\infty}\left(\frac{n!}{n^x(n-x)!}\right)\left(\frac{1}{(1-\frac{\lambda}{n})^x}\right)\left(1-\frac{\lambda}{n}\right)^n \\ &= \frac{\lambda^x}{x!}(1)(1)(e^{-\lambda}) \end{aligned}$$

hence, we have the probability mass function to be:

$$\Rightarrow f_X(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} : x \in \{0,1,2,\dots\} \text{ for some fixed } \lambda.$$

Thus, Given a random variable X such that $X \sim Poisson(\lambda)$ then X has probability mass function:

$$\boxed{f_X(x;\lambda) \triangleq \frac{e^{-\lambda}\lambda^x}{x!} : x \in \{0,1,2,\dots\}} \tag{8}$$

5. Hypergeometric Distribution.

$$\boxed{X \sim Hypergeom(M,N,n) : \max(0, M+n-N) \leq k \leq \min(M,n)} \tag{9}$$

Bernoulli trials without replacement (Binomial is actually a special case of the hypergeometric). Probability mass function is defined as:

$$\boxed{\begin{aligned} f_X(x;M,N,n) \triangleq \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} :&N \in \{1,2,\dots\} \\ &m \in \{0,1,2,\dots,N\} \\ &n \in \{1,2,\dots,N\} \end{aligned}} \tag{10}$$

References

[Pap91] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, NY, USA, 1991.

[Sem] Ladislaus Semali. Home page.