

## Binomial Distribution

The Binomial Distribution is a generalization of the Bernoulli distribution to a distribution over integers. In particular, the Binomial can be used to describe the probability of observing ' $m$ ' occurrence of  $X = 1$  in a set of ' $N$ ' samples from a '*Bernoulli Distribution*' where

$$p = (X = 1) = \mu \in [0, 1]$$

The Binomial distribution  $Bin(N, \mu)$  is defined as:

$$p(m|N, \mu) = \binom{N}{m} \mu^m (1 - \mu)^{N-m}$$

Where,

$p$  = binomial probability,

$m$  = no. of occurrence at  $(X = 1)$  within  $N$ , samples,

$N$  = Set of Samples,

$\mu$  = probability of success,

$1 - \mu$  = probability of failures,

$\binom{N}{m}$  = number of combinations.

Hence, above equation can also be written as:

$$p(m|N, \mu) = N C_m \mu^m (1 - \mu)^{N-m}$$

or,

$$p(m|N, \mu) = \frac{N!}{(N - m)m!} \mu^m (1 - \mu)^{N-m}$$

And,

$$E[m] = N\mu,$$

$$V[m] = N\mu(1 - \mu),$$

Where  $E[m]$  is mean and  $V[m]$  is variance.

## Exponential Family

$$f(x|\theta) = h(x)\exp(\eta(\theta)T(x) - A(\theta))$$

Now, take the Binomial distribution series:

$$p(m|N, \mu) = \binom{N}{m} \mu^m (1 - \mu)^{N-m}$$

Where  $m = 0, 1, 2, 3, 4, \dots, n$ .

Note:

$$f(m|N, \mu) = \begin{cases} \mu, & \text{if } m = 1 \\ 1 - \mu, & \text{if } m = 0 \end{cases}$$

*i. e.  $\mu$  (success) when  $m = 1$  i. e. there is an occurrence and  $1 - \mu$ , (failures) if  $m = 0$*

*i. e. there is no occurrence in  $N$  no. of samples and when there is occurrence at  $X$  is always 1 and  $P$  is probability .*

Therefore,

Now let us expand the Binomial's distribution,

$$\begin{aligned} p(m|N, \mu) &= \binom{N}{m} \mu^m (1 - \mu)^{N-m} \\ &= \exp \left( \log \left( \binom{N}{m} \mu^m (1 - \mu)^{N-m} \right) \right) \end{aligned}$$

*Note,  $\binom{N}{m}$  cannot be in log as log only be applied in  $a^b$  format.*

$$\begin{aligned} &= \binom{N}{m} \exp \left( \log(\mu^m (1 - \mu)^{N-m}) \right) \\ &= \binom{N}{m} \exp(m \log(\mu) + (N - m) \log(1 - \mu)) \\ &= \binom{N}{m} \exp(m \log(\mu) + N \log(1 - \mu) - m \log(1 - \mu)) \\ &= \binom{N}{m} \exp(m \log(\mu) - m \log(1 - \mu) + N \log(1 - \mu)) \\ &= \binom{N}{m} \exp \left( m \log \frac{\mu}{1 - \mu} + N \log(1 - \mu) \right) \\ &= \binom{N}{m} \exp \left( m \log \frac{\mu}{1 - \mu} - (-N \log(1 - \mu)) \right) \end{aligned}$$

From exponential family we get,

$$f(x|\theta) = h(x) \exp(\eta(\theta)T(x) - A(\theta))$$

We get,

$$\begin{aligned} h(x) &= \binom{N}{m}, \\ \eta(\theta) &= \log \frac{\mu}{1 - \mu}, \text{ or } \theta = \log \frac{\mu}{1 - \mu}, \end{aligned}$$

**$T(m) = m$  , or  $\theta(m) = m$  as  $x = m$  here,**

$$A(\theta) = -N \log(1 - \mu).$$

Now, note the relationship between  $\mu$  and  $\theta$  are invertible:

$$\mu = \frac{1}{1 + \exp^{-\theta}} - eqn(i)$$

And,

$$\exp^{-\theta} = \frac{1 - \mu}{\mu} - eqn(a)$$

Or,

$$\begin{aligned} \mu &= \frac{1}{1 + \frac{1}{\exp^{\theta}}} \\ \Rightarrow \mu &= \frac{1}{\frac{\exp^{\theta} + 1}{\exp^{\theta}}} \end{aligned}$$

$$\Rightarrow \mu = \frac{\exp^{\theta}}{\exp^{\theta} + 1} - eqn(ii)$$

Similarly:

$$\exp^{\theta} = \frac{\mu}{1 - \mu} - eqn(b)$$

From the above equation(i) we get :

$$\begin{aligned}\mu &= \frac{1}{1 + \exp^{-\theta}} \\ \Rightarrow \mu(1 + \exp^{-\theta}) &= 1 \\ \Rightarrow (1 + \exp^{-\theta}) &= \frac{1}{\mu} \\ \Rightarrow (\exp^{-\theta}) &= \frac{1}{\mu} - 1 \\ \Rightarrow (\exp^{-\theta}) &= \frac{1 - \mu}{\mu} \\ \Rightarrow \mu(\exp^{-\theta}) &= 1 - \mu\end{aligned}$$

we know ,

$$\mu = \frac{1}{1 + \exp^{-\theta}}$$

Therefore,

$$\begin{aligned}\Rightarrow \frac{1}{1 + \exp^{-\theta}} \times (\exp^{-\theta}) &= 1 - \mu \\ \Rightarrow \frac{\exp^{-\theta}}{1 + \exp^{-\theta}} &= 1 - \mu \\ \Rightarrow \frac{\frac{1}{\exp^{\theta}}}{1 + \frac{1}{\exp^{\theta}}} &= 1 - \mu\end{aligned}$$

$$\Rightarrow \frac{\frac{1}{\exp^{\theta}}}{\frac{\exp^{\theta} + 1}{\exp^{\theta}}} = 1 - \mu$$

$$\Rightarrow \frac{1}{\exp^{\theta}} \times \frac{\exp^{\theta}}{\exp^{\theta} + 1} = 1 - \mu$$

$$\Rightarrow \frac{1}{\exp^{\theta} + 1} = 1 - \mu$$

Therefore we got,

$$1 - \mu = \frac{1}{1 + \exp^{\theta}}$$

i.e.

**In addition, after computing it we get:**

$$\begin{aligned} A(\theta) &= -N \log \left( \frac{1}{1 + \exp^{\theta}} \right) \\ &= -N \left( \log(1) - \log(1 + \exp^{\theta}) \right) \\ &= -N \left( 0 - \log(1 + \exp^{\theta}) \right) \\ &= -N \left( -\log(1 + \exp^{\theta}) \right) \\ &= N \left( \log(1 + \exp^{\theta}) \right) \\ &= N \log(1 + \exp^{\theta}) \end{aligned}$$

Therefore, now we have:

$$p(m|N, \mu) = \binom{N}{m} \exp \left( m \log \frac{\mu}{1-\mu} - (N \log(1 + \exp^\theta)) \right)$$

From exponential family we get,

$$f(x|\theta) = h(x) \exp(\eta(\theta)T(x) - A(\theta))$$

We get,

$$h(x) = \binom{N}{m},$$

$$\eta(\theta) = \log \frac{\mu}{1-\mu}, \text{ or } \theta = \log \frac{\mu}{1-\mu},$$

$$T(m) = m, \text{ or } \theta(m) = m \text{ as } x = m \text{ here,}$$

$$A(\theta) = N \log(1 + \exp^\theta).$$

**Mean:**

$$E(T(x)) = E[x]$$

$$= \frac{\partial}{\partial \theta} A(\theta)$$

$$= \frac{\partial}{\partial \theta} (N \log(1 + \exp^\theta))$$

$$= N \left( \frac{\partial}{\partial \theta} (\log(1 + \exp^\theta)) \right) + \log(1 + \exp^\theta) \frac{\partial}{\partial \theta} \times N$$

$$= N \left( \frac{\partial}{\partial \theta} (\log(1 + \exp^\theta)) \right) + \log(1 + \exp^\theta) \times 0$$

$$= N \left( \frac{\partial}{\partial \theta} (\log(1 + \exp^\theta)) \right) + 0$$

$$= N \left( \frac{\partial}{\partial \theta} (\log(1 + \exp^\theta)) \right)$$

$$= N \left( \frac{1}{1 + \exp^\theta} \times \frac{\partial}{\partial \theta} (1 + \exp^\theta) \right)$$

$$= N \left( \frac{1}{1 + \exp^\theta} \times \left( 0 + \frac{\partial}{\partial \theta} \times \exp^\theta \right) \right)$$

Note, '**exp**' represents as '**e**' and '**exp**' stands for '**exponent**'.

Re-writing the equation:

$$= N \left( \frac{1}{1 + e^\theta} \times \left( 0 + \frac{\partial}{\partial \theta} \times e^\theta \right) \right)$$



$$\begin{aligned}
&= N \left( \frac{1}{1 + e^\theta} \times (0 + e^\theta) \right) \\
&= N \left( \frac{1}{1 + e^\theta} \times (e^\theta) \right) \\
&= N \left( \frac{e^\theta}{1 + e^\theta} \right)
\end{aligned}$$

We know,

$$\theta = \log \left( \frac{\mu}{1 - \mu} \right),$$

Or,

$$\eta(\theta) = \log \left( \frac{\mu}{1 - \mu} \right)$$

Hence putting the value of  $\theta$  in the equation:

$$= N \left( \frac{e^{\left(\log\left(\frac{\mu}{1-\mu}\right)\right)}}{1 + e^{\left(\log\left(\frac{\mu}{1-\mu}\right)\right)}} \right)$$

We can also represent it as:

$$= N \left( \frac{e^{\left(\log_e\left(\frac{\mu}{1-\mu}\right)\right)}}{1 + e^{\left(\log_e\left(\frac{\mu}{1-\mu}\right)\right)}} \right)$$

we know,  $a^{\log_a b} = b$

$$= N \left( \frac{\frac{\mu}{1-\mu}}{1 + \frac{\mu}{1-\mu}} \right)$$

$$= N \left( \frac{\frac{\mu}{1-\mu}}{\frac{1-\mu + \mu}{1-\mu}} \right)$$

$$= N \left( \frac{\frac{\mu}{1-\mu}}{\frac{1}{1-\mu}} \right)$$

$$= N \left( \frac{\mu}{1-\mu} \times \frac{1-\mu}{1} \right)$$

$$= N\mu$$

Hence, Mean =  $E[x] = N\mu$

## Variance

$$\text{Var}(T(x)) = \text{Var}[x]$$

We know variance =  $\sigma^2$

$$\begin{aligned}\text{Var}[x] &= \frac{\partial^2}{\partial \theta^2} \times A(\theta) \\ &= \frac{\partial^2}{\partial \theta^2} \times N \log(\exp^\theta + 1) \\ &= \frac{\partial}{\partial \theta} \times \frac{\partial}{\partial \theta} \times (N \log(\exp^\theta + 1))\end{aligned}$$

$$\text{We know } \frac{\partial}{\partial \theta} \times (N \log(\exp^\theta + 1)) = \frac{e^\theta}{1 + e^\theta}$$

*as we got it while doing calculation of mean, therefore:*

$$\begin{aligned}&= \frac{\partial}{\partial \theta} \times \left( N \left( \frac{e^\theta}{1 + e^\theta} \right) \right) \\ &= N \times \frac{\partial}{\partial \theta} \times \left( \left( \frac{e^\theta}{1 + e^\theta} \right) \right) + \left( \frac{e^\theta}{1 + e^\theta} \right) \times \frac{\partial}{\partial \theta} \times N \\ &= N \times \frac{\partial}{\partial \theta} \times \left( \left( \frac{e^\theta}{1 + e^\theta} \right) \right) + \left( \frac{e^\theta}{1 + e^\theta} \right) \times 0 \\ &= N \times \frac{\partial}{\partial \theta} \times \left( \left( \frac{e^\theta}{1 + e^\theta} \right) \right) + 0\end{aligned}$$

$$= N \times \frac{\partial}{\partial \theta} \times \left( \left( \frac{e^{\theta}}{1 + e^{\theta}} \right) \right)$$

$$= N \frac{\partial}{\partial \theta} \left( \frac{e^{\theta}}{1 + e^{\theta}} \right)$$

We know the division calculation of differential

$$\text{calculus: } \frac{\partial}{\partial x} \times \frac{f(x)}{g(x)} = \frac{g(x) \times \frac{\partial}{\partial x} \times f(x) - f(x) \times \frac{\partial}{\partial x} \times g(x)}{(g(x))^2}$$

$$= N \left( \frac{(1 + e^{\theta}) \times \frac{\partial}{\partial \theta} \times e^{\theta} - \left( e^{\theta} \times \frac{\partial}{\partial \theta} \times (1 + e^{\theta}) \right)}{(1 + e^{\theta})^2} \right)$$

$$= N \left( \frac{(1 + e^{\theta}) \times e^{\theta} - \left( e^{\theta} \times \left( 0 + \frac{\partial}{\partial \theta} \times e^{\theta} \right) \right)}{(1 + e^{\theta})^2} \right)$$

$$= N \left( \frac{e^{\theta} + e^{2\theta} - (e^{\theta} \times (0 + e^{\theta}))}{(1 + e^{\theta})^2} \right)$$

$$= N \left( \frac{e^{\theta} + e^{2\theta} - e^{2\theta}}{(1 + e^{\theta})^2} \right)$$

$$= N \left( \frac{e^{\theta}}{(1 + e^{\theta})^2} \right)$$

We know,

$$\theta = \log \left( \frac{\mu}{1 - \mu} \right),$$

**Therefore,**

$$= N \left( \frac{e^{\log\left(\frac{\mu}{1-\mu}\right)}}{\left(1 + e^{\log\left(\frac{\mu}{1-\mu}\right)}\right)^2} \right)$$

We can write it as:

$$= N \left( \frac{e^{\log_e \left( \frac{\mu}{1-\mu} \right)}}{\left( 1 + e^{\log_e \left( \frac{\mu}{1-\mu} \right)} \right)^2} \right)$$

$$we\ know, a^{\log_a b} = b$$

$$= N \left( \frac{\frac{\mu}{1-\mu}}{\left( 1 + \frac{\mu}{1-\mu} \right)^2} \right)$$

$$= N \left( \frac{\frac{\mu}{1-\mu}}{\left( \frac{1-\mu+\mu}{1-\mu} \right)^2} \right)$$

$$= N \left( \frac{\frac{\mu}{1-\mu}}{\left( \frac{1}{1-\mu} \right)^2} \right)$$

$$= N \left( \frac{\mu}{1 - \mu} \times (1 - \mu)^2 \right)$$

$$= N \times \mu \times (1 - \mu)$$

Therefore, **Variance** =  $Var[x] = N\mu(1 - \mu)$

**Now if we represent  $\mu$  as  $p$**

**Then,**

$$\text{Mean} = E[x] = np$$

$$Var[x] = np(1 - p)$$