## 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**

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

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

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

**Where,**

$$p = binomial\ probability,$$

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

$$N = Set\ of\ Samples,$$

$$\mu = probabity\ of\ success,$$

$$1 - \mu = probabity\ of\ failures,$$

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

**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

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

**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,..., n.**

**Note:**

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

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

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

Therefore,

**Now let us expand the Binomial's distribution,**

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

$$= exp\left(log\left(\binom{N}{m}\mu^m(1-\mu)^{N-m}\right)\right)$$

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

$$= \binom{N}{m} exp\left(log(\mu^m(1-\mu)^{N-m})\right)$$

$$= \binom{N}{m} exp\big(mlog(\mu) + (N-m)log(1-\mu)\big)$$

$$= \binom{N}{m} exp(mlog(\mu) + Nlog(1-\mu) - mlog(1-\mu))$$

$$= \binom{N}{m} exp(mlog(\mu) - mlog(1-\mu) + Nlog(1-\mu))$$

$$= \binom{N}{m} exp\left(mlog\frac{\mu}{1-\mu} + Nlog(1-\mu)\right)$$

$$= \binom{N}{m} exp\left(mlog\frac{\mu}{1-\mu} - \big(-Nlog(1-\mu)\big)\right)$$

From exponential family we get,

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

We get,

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

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

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

$$\boldsymbol{A(\theta) = -Nlog(1-\mu)}.$$

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,

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

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

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

Similarly:

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

From the above equation(i) we get :

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

$$=> \mu\left(1 + \exp^{-\theta}\right) = 1$$

$$=> \left(1 + \exp^{-\theta}\right) = \frac{1}{\mu}$$

$$=> \left(\exp^{-\theta}\right) = \frac{1}{\mu} - 1$$

$$=> \left(\exp^{-\theta}\right) = \frac{1 - \mu}{\mu}$$

$$=> \mu\left(\exp^{-\theta}\right) = 1 - \mu$$

we know ,

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

Therefore,

$$=> \frac{1}{1 + \exp^{-\theta}} \times \left(\exp^{-\theta}\right) = 1 - \mu$$

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

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

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

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

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

Therefore we got,

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

i.e.

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

$$\boldsymbol{A(\theta) = -Nlog\left(\frac{1}{1+exp^{\theta}}\right)}$$

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

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

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

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

$$\boldsymbol{= Nlog\left(1+exp^{\theta}\right)}$$

**Therefore, now we have:**

$$\boldsymbol{p(m|N,\mu)}$$

$$= \binom{N}{m} \boldsymbol{exp}\left(\boldsymbol{mlog}\frac{\boldsymbol{\mu}}{\boldsymbol{1-\mu}} - \left(\boldsymbol{Nlog}(\boldsymbol{1+exp^{\theta}})\right)\right)$$

**From exponential family we get,**

$$\boldsymbol{f(x|\theta) = h(x)exp\big(\eta(\theta)T(x) - A(\theta)\big)}$$

**We get,**

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

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

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

$$\boldsymbol{A(\theta) = Nlog\big(1 + exp^{\theta}\big)}.$$

# Mean:

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

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

$$= \frac{\partial}{\partial\theta}\big(\boldsymbol{Nlog(1 + exp^{\theta})}\big)$$

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

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

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

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

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

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

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

Re-writing the equation:

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

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

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

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

We know,

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

**Or,**

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

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

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

We can also represent it as:

$$= N\left(\frac{\mathrm{e}^{\left(\log_{\mathrm{e}}\left(\frac{\boldsymbol{\mu}}{\mathbf{1}-\boldsymbol{\mu}}\right)\right)}}{1+\mathrm{e}^{\left(\log_{\mathrm{e}}\left(\frac{\boldsymbol{\mu}}{\mathbf{1}-\boldsymbol{\mu}}\right)\right)}}\right)$$

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

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

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

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

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

$$= \boldsymbol{N\mu}$$

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

# Variance

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

We know variance = $\sigma^2$

$$Var[x] = \frac{\partial^2}{\partial\theta^2} \times A(\theta)$$

$$= \frac{\partial^2}{\partial\theta^2} \times \boldsymbol{Nlog}(\exp^{\theta} + 1)$$

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

$$\boldsymbol{We\ know}\ \frac{\partial}{\partial\theta} \times \left(\boldsymbol{Nlog}(\exp^{\theta} + 1)\right) = \frac{\mathrm{e}^{\theta}}{1 + \mathrm{e}^{\theta}}$$

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

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

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

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

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

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

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

We know the division calculation of differential calculus: $\frac{\partial}{\partial x} \times \frac{f(x)}{g(x)} = \frac{g(x)\times\frac{\partial}{\partial x}\times f(x) - f(x)\times\frac{\partial}{\partial x}\times g(x)}{\left(g(x)\right)^2}$

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

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

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

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

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

We know,

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

**Therefore,**

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

We can write it as:

$$= N\left(\frac{e^{\log_{\mathrm{e}}\left(\frac{\boldsymbol{\mu}}{\mathbf{1}-\boldsymbol{\mu}}\right)}}{\left(1+e^{\log_{\mathrm{e}}\left(\frac{\boldsymbol{\mu}}{\mathbf{1}-\boldsymbol{\mu}}\right)}\right)^{2}}\right)$$

$$we\ know, a^{\log_{\mathrm{a}} b} = b$$

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

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

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

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

$$= \boldsymbol{N} \times \boldsymbol{\mu} \times (\mathbf{1}-\boldsymbol{\mu})$$

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

**Now if we represent $\boldsymbol{\mu}$ $\boldsymbol{as}$ $\boldsymbol{p}$**

**Then,**

Mean = $E[x] = n\boldsymbol{p}$

$Var[x] = n\boldsymbol{p}(\mathbf{1}-\boldsymbol{p})$