# Notes on Statistical Theory

*Benjamin Nutter*

# Contents

# Part I

# Probability Distributions

# The Bernoulli Distribution

A random variable is said to have a Bernoulli Distribution with parameter $p$ if its probability mass function is:

$$p(x) = \begin{cases} p^x(1-p)^{1-x}, & x = 0,1 \\ 0 & \text{otherwise} \end{cases}$$

Where $p$ is the probability of a sucess.

## 1 Cumulative Distribution Function

$$P(x) = \begin{cases} 0 & x < 0 \\ 1-p & x = 0 \\ 1 & 1 \leq x \end{cases}$$
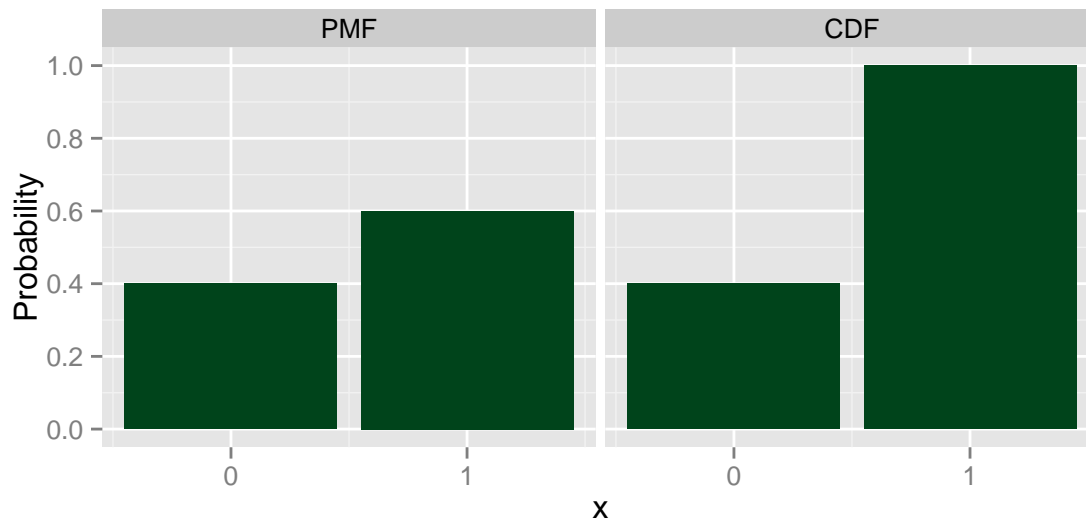


Figure .1: The graphs on the left and right show a Binomial Probability Distribution and Cumulative Distribution Function, respectively, with $p = .4$. Note that this is identical to a Binomial Distribution with parameters $n = 1$ and $p = .4$.

## 2   Expected Values

$$E(X) = \sum_{i=0}^{1} x \cdot p(x)$$
$$= \sum_{i=0}^{1} x \cdot p^x (1-p)^{1-x}$$
$$= 0 \cdot p^0 (1-p)^{1-0} + 1 \cdot p^1 (1-p)^{1-1}$$
$$= 0 + p(1-p)^0$$
$$= p$$

---

$$E(X^2) = \sum_{i=0}^{1} x^2 \cdot p(x)$$
$$= \sum_{i=0}^{1} x^2 \cdot p^x (1-p)^{1-x}$$
$$= \sum_{i=0}^{1} 0^2 \cdot p^0 (1-p)^{1-0} + 1^2 \cdot p^1 (1-p)^{1-1}$$
$$= 0 \cdot 1 \cdot 1 + 1 \cdot p \cdot 1$$
$$= 0 + p$$
$$= p$$

---

$\mu = E(X) = p$

---

$\sigma^2 = E(X^2) - E(X)^2 = p - p^2 = p(1-p)$

---

## 3   Moment Generating Function

$$M_X(t) = E(e^{tX}) = \sum_{i=0}^{1} e^{tx} p(x) = \sum_{i=0}^{1} e^{tx} p^x (1-p)^{1-x}$$
$$= e^{t0} p^0 (1-p)^{1-0} + e^t p^t (1-p)^{1-1} = (1-p) + e^t p = pe^t + (1-p)$$

---

$$M_X^{(1)}(t) = pe^t$$
$$M_X^{(2)}(t) = pe^t$$

$$E(X) = M_X^{(1)}(0) = pe^0 = p \cdot 1 = p$$

$$E(X^2) = M_X^{(2)}(0) = pe^0 = p$$

$$\mu = E(X) = p$$

$$\sigma^2 = E(X^2) - E(X)^2 = p - p^2 = p(1-p)$$

# 4   Theorems for the Bernoulli Distribution

## 4.1   Validity of the Distribution

$$\sum_{x=0}^{1} p^x(1-p)^{1-x} = 1$$

*Proof:*

$$\sum_{x=0}^{1} p^x(1-p)^{1-x} = p^0(1-p)^1 + p^1(1-p)^0 = (1-p) + p = 1$$

∎

## 4.2   Sum of Bernoulli Random Variables

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables from a Bernoulli distribution with parameter $p$. Let $Y = \sum_{i=1}^{n} X_i$.

Then $Y \sim \text{Binomial}(n, p)$

*Proof:*

$$M_Y(t) = E(e^{tY}) = E(e^{tX_1}e^{tX_2}\cdots e^{tX_n}) = E(e^{tX_1})E(e^{tX_2})\cdots E(e^{tX_n})$$

$$= (pe^t + (1-p))(pe^t + (1-p))\cdots(pe^t + (1-p)) = ($$

Which is the mgf of a Binomial random variable with parameters $n$ and $p$.

Thus, $Y \sim \text{Binomial}(n, p)$. ∎

# The Binomial Distribution

A random variable is said to follow a Binomial distribution with parameters $n$ and $p$ if its probability mass function is:\

$$p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

Where $n$ is the number of trials performed and $p$ is the probability of a success on each individual trial.

## 5 Cumulative Distribution Function

$$P(x) = \begin{cases} 0 & x < 0 \\ \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i} & 0 \le x = 0, 1, 2, \ldots, n \\ 1 & n \le x \end{cases}$$

A recursive form of the cdf can be derived and has some usefulness in computer applications. With it, one need only initiate the first value and additional cumulative probabilities can be calculated. It is derived as follows:

$$\begin{aligned}
F(x+1) &= \binom{n}{x+1} p^{x+1} (1-p)^{n-(x+1)} \\
&= \frac{n!}{(x+1)!(n-(x+1))!} p^{x+1} (1-p)^{n-(x+1)} \\
&= \frac{n!}{(x+1)!(n-x-1)!} p^{x+1} (1-p)^{n-x-1} \\
&= \frac{(n-x)n!}{(x+1)x!(n-x)(n-x-1)!} p \cdot p^x \frac{(1-p)^{n-x}}{(1-p)} \\
&= \frac{(n-x)n!}{(x+1)x!(n-x)!} \cdot \frac{p}{1-p} p^x (1-p)^{n-x} \\
&= \frac{p}{1-p} \cdot \frac{n-x}{x+1} \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \frac{p}{1-p} \cdot \frac{n-x}{x+1} \cdot \binom{n}{x} p^x (1-p)^{n-x} \\
&= \frac{p}{1-p} \cdot \frac{n-x}{x+1} \cdot F(x)
\end{aligned}$$

## 6 Expected Values

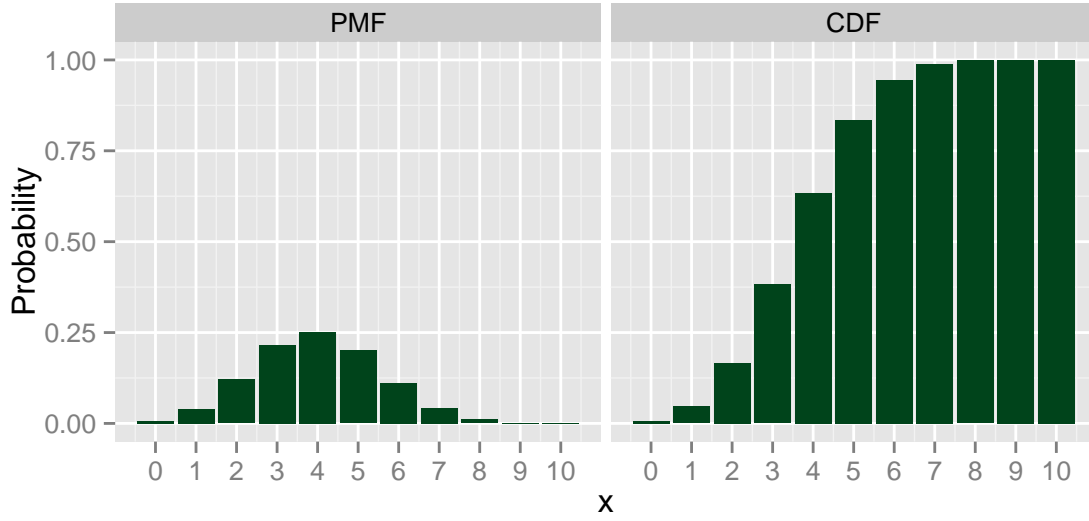Let $X$ be a binomial random variable with parameters $n$ and $p$. The expected value of $X$ is:

Figure .1: plot of chunk unnamed-chunk-7

$$E(X) = \sum_{x=0}^{n} x \cdot p(x)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

For convenience, let $q = (1-p)$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$

$$= 0 \cdot \binom{n}{0} p^0 q^n + 1 \cdot \binom{n}{1} p^1 q^{n-1} + \cdots + n \binom{n}{n} p^n q^{n-n}$$

$$= 0 + 1 \binom{n}{1} p^1 q^{n-1} + 2 \binom{n}{2} p^2 q^{n-2} + \cdots + n \binom{n}{n} p^n q^{n-n}$$

$$= np^1 q^{n-1} + n(n-1) p^2 q^{n-2} + \cdots + n(n-1) p^{n-1} q^{n-(n-1)} + np^n$$

$$= np[q^{n-1} + (n-1)pq^{n-2} + \cdots + p^{n-1}]$$

$$= np\left[ \binom{n-1}{0} p^0 q^{n-1} + \binom{n-1}{1} p^1 q^{(n-1)-1} + \cdots + \binom{n-1}{n-1} p^{n-1} q^{(n-}\right.$$

$$= np\left( \sum_{x=0}^{n-1} \binom{n-1}{x} p^x q^{(n-1)-x} \right)$$

By the Binomial Theorem, $\sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x} = (a+b)^n$

$$= np(p+q)^{n-1}$$

Resubstituting $(1-p)$ for $q$ gives us

$$= np(p + (1-p))^{n-1}$$
$$= np(p + 1 - p)^{n-1}$$
$$= np(1)^{n-1}$$
$$= np(1)$$
$$= np$$

The Expected Value of $X^2$ is:

$$E(X^2) = \sum_{x=0}^{n} x^2 p(x)$$

$$= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}$$

For convenience, let $q = (1-p)$

$$= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$$

$$= 0^2 \frac{n!}{0!(n-0)!} p^0 q^n + 1^2 \frac{n!}{1!(n-1)!} p^1 q^{n-1} + \cdots + n^2 \frac{n!}{n!(n-n)!} p^n q^{n-n}$$

$$= 0 + 1 \frac{n!}{(n-1)!} pq^{n-1} + 2 \frac{n!}{1 \cdot (n-2)!} p^2 q^{n-2} + \cdots + n \frac{n!}{(n-1)!(n-n)!} p^n$$

$$= np \left[ 1 \frac{(n-1)!}{(n-1)!} p^0 q^{n-1} + 2 \frac{(n-1)!}{1(n-2)!} p^2 q^{n-2} + \cdots + n \frac{(n-1)!}{(n-1)!(n-n)!} p^{n-1} \right]$$

$$= np \left[ 1 \frac{(n-1)!}{(1-1)!((n-1)-(-1-1))!} p^{1-1} q^{n-1} + \cdots + n \frac{(n-1)!}{(n-1)!((n-1)-(n-1))!} p^{n-1} q \right.$$

$$= np \sum_{x=1}^{n} x \binom{n-1}{x-1} p^{x-1} 1^{(n-1)-(x-1)}$$

Let $y = x - 1$ and $n = m + 1$
$\Rightarrow x = y + 1$ and $m = n - 1$

$$= \sum_{y=0}^{m} (y+1) \binom{m}{y} p^y q^{m-y}$$

$$= np \left[ \sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y} + \binom{m}{y} p^y q^{m-y} \right]$$

$$= np \left[ \sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y} + \sum_{y=0}^{m} \binom{m}{y} p^y q^{m-y} \right]$$

$\sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y}$ is of the form

of the expected value of $Y$,
and $E(Y) = mp = (n-1)p$

$\sum_{y=0}^{m} \binom{m}{y} p^y q^{m-y}$ is the sum of all

probabilities over the domain of $Y$,
which is 1.

$$= np(mp + 1)$$
$$= np[(n-1)p + 1]$$
$$= np(np - p + 1)$$
$$= n^2 p^2 - np^2 + np$$

The mean of $X$ can be calculated as

$$\mu = E(X) = np$$

And the variance of $X$ can be calculated by

$$
\begin{aligned}
\sigma^2 &= E(X^2) - E(X)^2 \\
&= n^2p^2 - np^2 + np - n^2p^2 \\
&= -np^2 + np \\
&= np(-p - 1) \\
&= np(1 - p)
\end{aligned}
$$

# 7   Moment Generating Function

$$
\begin{aligned}
M_X(t) = E(e^{tX}) &= \sum_{x=0}^{n} e^{tx} p(x) \\
&= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} e^{tx} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (pe^{tx})^x (1-p)^{n-x}
\end{aligned}
$$

By Binomial Theorem REF

$$\sum_{x=0}^{n} \binom{n}{x} b^x a^{n-x} = (a+b)^n$$

$$= [(1-p) + pe^t]^n$$

$$M_X^{(1)}(t) = n[(1-p) + pe^t]^{n-1} pe^t$$

$$
\begin{aligned}
M_X^{(2)}(t) &= n[(1-p) + pe^t]^{n-1} pe^t + n(n-1)[(1-p) + pe^t]^{n-2}(pe^t)^2 \\
&= npe^t[(1-p) + pe^t]^{n-1} + n(n-1)pe^{2t}[(1-p) + pe^t]^{n-2}
\end{aligned}
$$

$$\begin{aligned}
E(X) &= M_X^{(1)}(0) \\
&= n[(1-p)+pe^0]^{n-1}pe^0 \\
&= n[1-p+p^{n-1}p \\
&= n(1)^{n-1}p = np
\end{aligned}$$

$$\begin{aligned}
E(X^2) &= M_X^{(2)}(0) = npe^0[(1-p)+pe^0]^{n-1} + n(n-2)pe^{2\cdot 0}[(1-p)+pe^0]^{n-2} \\
&= np(1-p+p)^{n-2} + n(n-1)p^2(1-p+p^{n-2} \\
&= np(1)^{n-1} + n(n-1)p^2(1)^{n-2} = np + n(n-1)p^2 = np + (n^2-n)p^2 \\
&= np + n^2 + n^2p^2 - np^2
\end{aligned}$$

$$\mu = E(X) = np$$

$$\begin{aligned}
\sigma^2 &= E(X^2) - E(X)^2 \\
&= np + n^2p^2 - np^2 - n^2p^2 \\
&= np - np^2 \\
&= np(1-p)
\end{aligned}$$

# 8   Maximum Likelihood Estimator

Since $n$ is fixed in each Binomial experiment, and must therefore be given, it is unnecessary to develop an estimator for $n$. The mean and variance can both be estimated from the single parameter $p$.

Let $X$ be a Binomial random variable with parameter $p$ and $n$ outcomes $(x_1, x_2, \ldots, x_n)$. Let $x_i = 0$ for a failure and $x_i = 1$ for a success. In other words, $X$ is the sum of $n$ Bernoulli trials with equal probability of success and $X = \sum_{i=1}^{n} x_i$.

## 8.1   Likelihood Function

$$\begin{aligned}
L(\theta) &= L(x_1, x_2, \ldots, x_n | \theta) \\
&= P(x_1|\theta)P(x_2|\theta)\cdots P(x_n|\theta) \\
&= [\theta^{x_1}(1-\theta)^{1-x_1}][\theta^{x_2}(1-\theta)^{1-x_2}]\cdots[\theta^{x_n}(1-\theta)^{1-x_n}] \\
&= \exp_\theta\left\{\sum_{i=1}^{n} x_i\right\} \exp_{(1-\theta)}\left\{n - \sum_{i=1}^{n} x_i\right\} \\
&= \theta^X(1-\theta)^{n-X}
\end{aligned}$$

## 8.2 Log-likelihood Function

$$\ell(\theta) = \ln L(\theta)$$
$$= \ln \left( \theta^X (1-\theta)^{n-X} \right)$$
$$= X \ln(\theta) + (n-X) \ln(1-\theta)$$

## 8.3 MLE for $p$

$$\frac{d\ell(p)}{dp} = \frac{X}{p} - \frac{n-X}{1-p}$$
$$0 = \frac{X}{p} - \frac{n-X}{1-p}$$
$$\frac{X}{p} = \frac{n-X}{1-p}$$
$$(1-p)X = p(n-X)$$
$$X - pX = np - pX$$
$$X = np$$
$$\frac{X}{n} = p$$

So $\hat{p} = \frac{X}{n} = \frac{1}{n} \sum\limits_{i=1}^{n} x_i$ is the maximum likelihood estimator for $p$.

# 9 Theorems for the Binomial Distribution

## 9.1 Validity of the Distribution

$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = 1$$

*Proof:*

$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} =$$
$$\sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x} (a+b)^n.$$
See Binomial Theorem REF.
$$= big(p + (1-p))^n$$
$$= (1)^n$$
$$= 1$$

∎

## 9.2   Sum of Binomial Random Variables

Let $X_1, X_2, \ldots, X_k$ be independent random variables where $X_i$ comes from a Binomial distribution with parameters $n_i$ and $p$. That is $X_i \sim (n_i, p)$.\ Let $Y = \sum_{i=1}^{k} kX_i$. Then $Y \sim$Binomial$(\sum_{i=1}^{k} n_i, p)$.

*Proof:*

$$
\begin{aligned}
M_Y(t) &= E(e^{tY}) \\
&= E(e^{t(x_1 + X_2 + \cdots + X_k)}) \\
&= E(e^{tX_1} e^{tX_2} \cdots e^{tX_k}) \\
&= E(e^{tX_1}) E(e^{tX_2}) \cdots E(e^{tX_k}) \\
&= \prod_{i=1}^{k} [(1-p) + pe^t]^{n_i} \\
&= [(1-p) + pe^t]^{\sum_{i=1}^{k} n_i}
\end{aligned}
$$

\ \ Which is the mgf of a Binomial random variable with parameters $\sum_{i=1}^{k} n_i$ and $p$.

Thus $Y \sim$Binomial$(\sum_{i=1}^{k} n_i, p)$.∎

## 9.3   Sum of Bernoulli Random Variables

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables from a Bernoulli distribution with parameter $p$. Let $Y = \sum_{i=1}^{n} X_i$.

Then $Y \sim$Binomial$(n, p)$

*Proof:*

$$
\begin{aligned}
M_Y(t) &= E(e^{tY}) \\
&= E(e^{tX_1} e^{tX_2} \cdots e^{tX_n}) \\
&= E(e^{tX_1}) E(e^{tX_2}) \cdots E(e^{tX_n}) \\
&= (pe^t + (1-p))(pe^t + (1-p)) \cdots (pe^t + (1-p)) \\
&= (pe^t + (1-p))^n
\end{aligned}
$$

Which is the mgf of a Binomial random variable with parameters $n$ and $p$. Thus, $Y \sim$ Binomial$(n, p)$. ∎

# The Chi-square Distribution

# The Exponential Distribution

# The Gamma Distribution

# The Geometric Distribution

# The Hypergeometric Distribution

# The Multinomial Distribution

# The Normal Distribution

# The Poisson Distribution

# The Skew-Normal Distribution

# The Uniform Distribution

# The Weibull Distribution

# Part II

# Frequentist Hypothesis Testing

# Mantel-Haenszel Test

# Part III

# Supplemental Subjects

# Chebychev's Theorem

# Combinations

# The Correlation Coefficient

# Covariance

# Experimental Designs

# Moments and Moment Generating Functions

# Summation

# The Method of Transformations

# Variance Paramter

# Part IV

# Non-Statistical Proofs

# The Binomial Theorem

# Functions

# The Geometric Series

# Integraion: Techniques and Theorems

# Logarithmic and Exponential Functions

# The Real Number System