It Can Be Shown

Benjamin Nutter

Contents

The content of these pages is intended to make explicit the development of mathematical and statistical theory while avoiding the nebulous "It can be shown." As much as possible, the equations are developed step-by-step so that each step can be followed and understood.

In instances where the changes are not obvious, explanations are provided in red. Originally, these explanations were intended to be footnotes, but footnotes are not well suited to the web page format. For consistency, the PDF version of these pages is produced in an identical format. The PDF may be downloaded from the https://github.com/nutterb/ItCanBeShown GitHub repository.

Additions, corrections, and suggestions are welcome on GitHub.

Part I

Probability Distributions

The Bernoulli Distribution

1 Probability Mass Function

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.

2 Cumulative Distribution Function

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

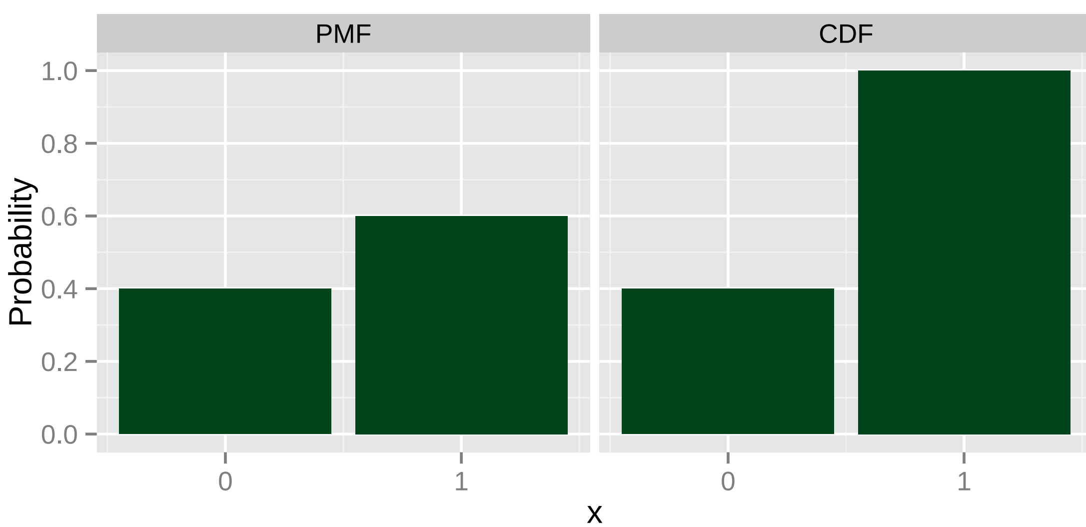

Figure .1: The graphs on the left and right show a Binomial Probability Distribution and Cumulative Distribution Function, respectively, with $p = .4$. Note that this is identical to a Binomial Distribution with parameters $n = 1$ and $p = .4$.

3 Expected Values

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

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

$\mu = E(X) = p$

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

---

4 Moment Generating Function

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

$$\begin{aligned}
M_X^{(1)}(t) &= pe^t \\
M_X^{(2)}(t) &= pe^t
\end{aligned}$$

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

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

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

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

5 Theorems for the Bernoulli Distribution

5.1 Validity of the Distribution

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

Proof:

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

■

5.2 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

1 Probability Mass Function

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.

2 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:

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

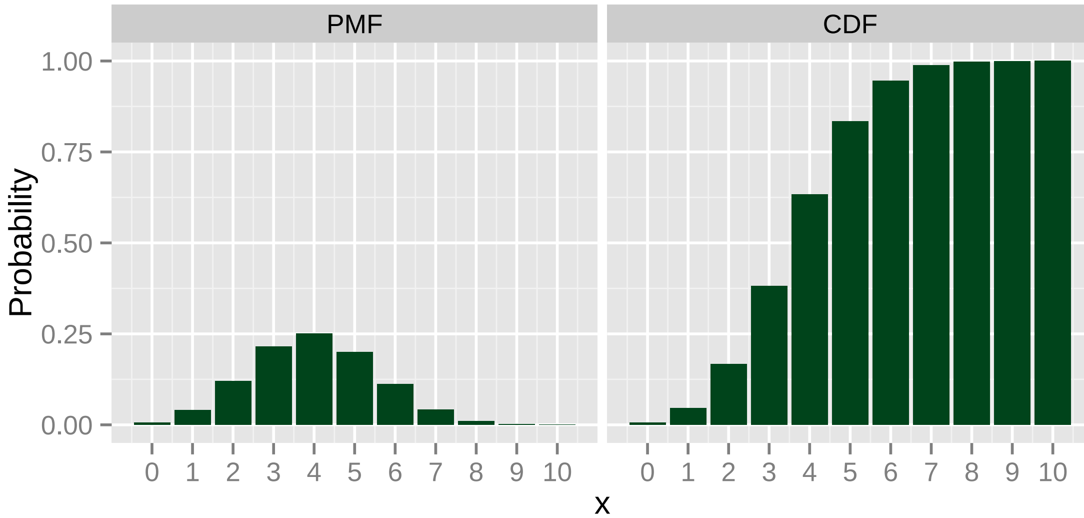

Figure .1: plot of chunk unnamed-chunk-9

3 Expected Values

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

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

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

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$
$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$
$$= 0 \cdot \binom{n}{0} p^0 q^n + 1 \cdot \binom{n}{1} p^1 q^{n-1} + \cdots + n \binom{n}{n} p^n q^{n-n}$$
$$= 0 + 1 \binom{n}{1} p^1 q^{n-1} + 2 \binom{n}{2} p^2 q^{n-2} + \cdots + n \binom{n}{n} p^n q^{n-n}$$
$$= np^1 q^{n-1} + n(n-1) p^2 q^{n-2} + \cdots + n(n-1) p^{n-1} q^{n-(n-1)} + np^n$$
$$= np[q^{n-1} + (n-1) pq^{n-2} + \cdots + p^{n-1}]$$
$$= np\Big[\binom{n-1}{0} p^0 q^{n-1} + \binom{n-1}{1} p^1 q^{(n-1)-1} + \cdots + \binom{n-1}{n-1} p^{n-1} q^{(n-}$$
$$= np(\sum_{x=0}^{n-1} \binom{n-1}{x} p^x q^{(n-1)-x})$$

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:

$$\begin{aligned}
E(X^2) &= \sum_{x=0}^{n} x^2 p(x) \\
&= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}
\end{aligned}$$

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

$$\begin{aligned}
&= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x} \\
&= 0^2 \frac{n!}{0!(n-0)!} p^0 q^n + 1^2 \frac{n!}{1!(n-1)!} p^1 q^{n-1} + \cdots + n^2 \frac{n!}{n!(n-n)!} p^n q^{n-n} \\
&= 0 + 1 \frac{n!}{(n-1)!} p q^{n-1} + 2 \frac{n!}{1 \cdot (n-2)!} p^2 q^{n-2} + \cdots + n \frac{n!}{(n-1)!(n-n)!} p^n \\
&= np\Big[1 \frac{(n-1)!}{(n-1)!} p^0 q^{n-1} + 2 \frac{(n-1)!}{1(n-2)!} p^2 q^{n-2} + \cdots + n \frac{(n-1)!}{(n-1)!(n-n)!} p^{n-1}\Big] \\
&= np\Big[1 \frac{(n-1)!}{(1-1)!((n-1)-(-1-1))!} p^{1-1} q^{n-1} + \cdots + n \frac{(n-1)!}{(n-1)!((n-1)-(n-1))!} p^{n-1} q \\
&= np \sum_{x=1}^{n} x \binom{n-1}{x-1} p^{x-1} 1^{(n-1)-(x-1)}
\end{aligned}$$

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

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

$$\begin{aligned}
&= \sum_{y=0}^{m} (y+1) \binom{m}{y} p^y q^{m-y} \\
&= np\Big[\sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y} + \binom{m}{y} p^y q^{m-y}\Big] \\
&= np\Big[\sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y} + \sum_{y=0}^{m} \binom{m}{y} p^y q^{m-y}\Big]
\end{aligned}$$

$\sum_{y=0}^{m} y \binom{m}{y} p^y q^{m-y}$ 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.

$$\begin{aligned}
&= np(mp + 1) \\
&= np[(n-1)p + 1] \\
&= np(np - p + 1) \\
&= n^2p^2 - np^2 + np
\end{aligned}$$

The mean of X can be calculated as

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

And the variance of X can be calculated by

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

---

4 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

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

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

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

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

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

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

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

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

5 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 $X = \sum_{i=1}^{n} x_i$.

5.1 Likelihood Function

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

5.2 Log-likelihood Function

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

5.3 MLE for p

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

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

6 Theorems for the Binomial Distribution

6.1 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.

$$\begin{aligned}&= big(p + (1-p)\big)^n\\ &= (1)^n\\ &= 1\end{aligned}$$

■

6.2 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 $Y = \sum_{i=1} kX_i$. Then $Y \sim$Binomial($\sum_{i=1}^{k} n_i, p$).

Proof:

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

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

Thus Y ~Binomial($\sum\limits_{i=1}^{k} n_i, p$).■

6.3 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 $Y = \sum\limits_{i=1}^{n} X_i$.

Then Y ~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

1 Probability Distribution Function

A random variable X is said to have a Chi-Square Distribution with parameter ν if its probability distribution function is

$$f(x) = \begin{cases} \frac{x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} & 0 < x,\ 0 < \nu \\ 0 & otherwise \end{cases}$$

ν is commonly referred to as the *degrees of freedom*.

2 Cumulative Distribution Function

The cumulative distribution function for the Chi-Square Distribution cannot be written in closed form. It's integral form is expressed as

$$F(x) = \begin{cases} \int_0^x \frac{t^{\frac{\nu}{2}-1}e^{-\frac{t}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} dt & 0 < x,\ 0 < \nu \\ \\ 0 & otherwise \end{cases}$$

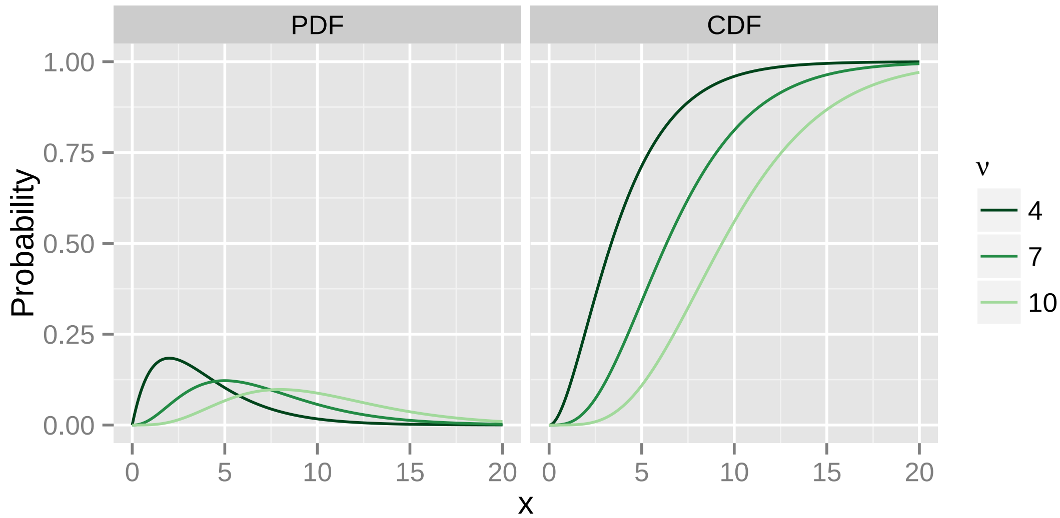

Figure .1: The graphs on the left and right depict the Chi-Square probability distribution and cumulative distribution functions, respectively, for $\nu = 4, 7, 10$. As ν gets larger, the distribution becomes flatter with thicker tails.

3 Expected Values

$$
\begin{aligned}
E(X) &= \int_0^\infty x \frac{x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x \cdot x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}} e^{-\frac{x}{2}} dx
\end{aligned}
$$

$$\int_0^\infty x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \beta^\alpha \Gamma(\alpha)$$

$$
\begin{aligned}
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \left[\Gamma\left(\frac{\nu}{2} + 1\right) 2^{\frac{\nu}{2}+1} \right] \\
&= \frac{\Gamma(\frac{\nu}{2} + 1) 2^{\frac{\nu}{2}+1}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \\
&= \frac{\frac{\nu}{2} \Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2}+1}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \\
&= \frac{2\nu}{2} \\
&= \nu
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \int_0^\infty x^2 \frac{x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}dx \\
&= \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\int_0^\infty x^2 \cdot x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}dx \\
&= \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\int_0^\infty x^{\frac{\nu}{2}+1}e^{-\frac{x}{2}}dx
\end{aligned}
$$

$$\int_0^\infty x^{\alpha-1}e^{-\frac{x}{\beta}}dx = \beta^\alpha\Gamma(\alpha)$$

$$
\begin{aligned}
&= \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\left[\Gamma(\frac{\nu}{2}+2)2^{\frac{\nu}{2}+2}\right] \\
&= \frac{\Gamma\left(\frac{\nu}{2}+2\right)2^{\frac{\nu}{2}+2}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \\
&= \frac{(\frac{\nu}{2}+1)\Gamma(\frac{\nu}{2}+1)2^{\frac{\nu}{2}+2}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \\
&= \frac{\left(\frac{\nu}{2}+1\right)\frac{\nu}{2}\Gamma(\frac{\nu}{2})2^{\frac{\nu}{2}+2}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \\
&= \left(\frac{\nu}{2}+1\right)\frac{\nu}{2}\cdot 2^2 = 2\left(\frac{\nu}{2}+1\right)\nu \\
&= (\nu+2)\nu = \nu^2+2\nu
\end{aligned}
$$

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

$$\sigma^2 = E(X^2) - E(X)^2 = \nu^2 + 2\nu - \nu^2 = 2\nu$$

4 Moment Generating Function

$$
\begin{aligned}
M_X(t) &= E(e^{tX}) \\
&= \int_0^\infty e^{tx} \frac{x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty e^{tx} \cdot x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{tx} e^{-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{tx-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{\frac{2tx}{2}-\frac{x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{-\frac{2tx-x}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{-x\frac{-2t+1}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{-x\frac{1-2t}{2}} dx \\
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^\infty x^{\frac{\nu}{2}-1} e^{\frac{\frac{-x}{2}}{1-2t}} dx
\end{aligned}
$$

$$
\int_0^\infty x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \beta^\alpha \Gamma(\alpha)
$$

$$
\begin{aligned}
&= \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \left[\left(\frac{2}{1-2t} \right)^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2}) \right] \\
&= \frac{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2}) (1-2t)^{\frac{\nu}{2}}} \\
&= \frac{1}{(1-2t)^{\frac{\nu}{2}}} \\
&= (1-2t)^{-\frac{\nu}{2}}
\end{aligned}
$$

$$
\begin{aligned}
M_X^{(1)}(t) &= -\frac{\nu}{2}(1-2t)^{-\frac{\nu}{2}-1}(-2) \\
&= \frac{2\nu}{2}(1-2t)^{-\frac{\nu}{2}-1} \\
&= \nu(1-2t)^{-\frac{\nu}{2}-1}
\end{aligned}
$$

$$
\begin{aligned}
M_X^{(2)}(t) &= (-\frac{\nu}{2} - 1)\nu(1-2t)^{-\frac{\nu}{2}-2}(-2) \\
&= (\frac{2\nu}{2} + 2)\nu(1-2t)^{-\frac{\nu}{2}-2} \\
&= (\nu + 2)\nu)(1-2t)^{-\frac{\nu}{2}-2} \\
&= (\nu^2 + 2\nu)(1-2t)^{-\frac{\nu}{2}-2}
\end{aligned}
$$

$$
\begin{aligned}
M_X^{(1)}(0) &= \nu(1-2\cdot 0)^{-\frac{\nu}{2}-1} \\
&= \nu(1-0)^{-\frac{\nu}{2}-1} \\
&= \nu(1)^{-\frac{\nu}{2}-1} \\
&= \nu
\end{aligned}
$$

$$
\begin{aligned}
M_X^{(2)}(0) &= (\nu^2 + 2\nu)(1-2\cdot 0)^{-\frac{\nu}{2}-2} \\
&= (\nu^2 + 2\nu)(1-0)^{-\frac{\nu}{2}-2} \\
&= (\nu^2 + 2\nu)(1)^{-\frac{\nu}{2}-2} \\
&= (\nu^2 + 2\nu)
\end{aligned}
$$

$$E(X) = M_X^{(1)}(0) = \nu$$

$$E(X^2) = M_X^{(2)}(0) = (\nu^2 + 2\nu)$$

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

$$\sigma^2 = E(X^2) - E(X)^2 = \nu^2 + 2\nu - \nu^2 = 2\nu$$

5 Maximum Likelihood Function

Let $x_1, x_2, \ldots, x_n$ be a random sample from a Chi-square distribution with parameter ν.

5.1 Likelihood Function

$$
\begin{aligned}
L(\theta) &= f(x_1|\theta)f(x_2|\theta)\cdots f(x_n|\theta) \\
&= \frac{x_1^{\nu/2-1}e^{-x_1/2}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{x_2^{\nu/2-1}e^{-x_2/2}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} \cdots \frac{x_n^{\nu/2-1}e^{-x_n/2}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} \\
&= \prod_{i=1}^{n} \frac{x_i^{\nu/2-1}e^{-x_i/2}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} \\
&= \left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right) \prod_{i=1}^{n} x_i^{\nu/2-1}e^{-x_i/2} \\
&= \left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right) \cdot \exp\left\{\sum_{i=1}^{n}\frac{x_i}{2}\right\} \cdot \prod_{i=1}^{n} x_i^{\nu/2-1} \\
&= \left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right) \cdot \exp\left\{\frac{1}{2}\sum_{i=1}^{n} x_i\right\} \cdot \prod_{i=1}^{n} x_i^{\nu/2-1}
\end{aligned}
$$

5.2 Log-likelihood Function

$$
\begin{aligned}
\ell(\theta) &= \ln\big(L(\theta)\big) \\
&= \ln\left[\left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right) \cdot \exp\left\{\frac{1}{2}\sum_{i=1}^{n} x_i\right\} \cdot \prod_{i=1}^{n} x_i^{\nu/2-1}\right] \\
&= \ln\left[\left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right)\right] + \ln\left(\exp\left\{\frac{1}{2}\sum_{i=1}^{n} x_i\right\}\right) + \ln\left(\prod_{i=1}^{n} x_i^{\nu/2-1}\right) \\
&= -n\ln\left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right) + \frac{1}{2}\sum_{i=1}^{n} x_i + \left(\frac{\nu}{2}-1\right)\ln\left(\prod_{i=1}^{n} x_i\right) \\
&= -n\left(\ln(2^{\nu/2}) + \Gamma\left(\frac{\nu}{2}\right)\right) + \frac{1}{2}\sum_{i=1}^{n} x_i + \left(\frac{\nu}{2}-1\right)\sum_{i=1}^{n}\ln x_i \\
&= -n\left(\frac{\nu}{2}\ln 2 + \ln\Gamma\left(\frac{\nu}{2}\right)\right) + \frac{1}{2}\sum_{i=1}^{n} x_i + \left(\frac{\nu}{2}-1\right)\sum_{i=1}^{n}\ln x_i \\
&= -\frac{n\nu}{2}\ln 2 - n\ln\Gamma\left(\frac{\nu}{2}\right) + \frac{1}{2}\sum_{i=1}^{n} x_i + \left(\frac{\nu}{2}-1\right)\sum_{i=1}^{n}\ln x_i
\end{aligned}
$$

5.3 MLE for ν

$$
\begin{aligned}
\frac{d\ell}{d\nu} &= -\frac{n}{2}\ln 2 - \frac{n}{\Gamma\left(\frac{\nu}{2}\right)}\Gamma'\left(\frac{\nu}{2}\right)\cdot\frac{1}{2} + 0 + \frac{1}{2}\sum_{i=1}^{n}\ln x_i \\
&= -\frac{n}{2}\ln 2 - \frac{n}{2\Gamma\left(\frac{\nu}{2}\right)}\Gamma'\left(\frac{\nu}{2}\right) + \frac{1}{2}\sum_{i=1}^{n}\ln x_i
\end{aligned}
$$

$$
\begin{aligned}
0 &= -\frac{n}{2}\ln 2 - \frac{n}{2\Gamma\left(\frac{\nu}{2}\right)}\Gamma'\left(\frac{\nu}{2}\right) + \frac{1}{2}\sum_{i=1}^{n}\ln x_i \\
\frac{n}{2}\ln 2 - \frac{1}{2}\sum_{i=1}^{n}\ln x_i &= -\frac{n}{2\Gamma\left(\frac{\nu}{2}\right)}\Gamma'\left(\frac{\nu}{2}\right) \\
n\ln 2 - \sum_{i=1}^{n}\ln x_i &= -\frac{n}{\Gamma\left(\frac{\nu}{2}\right)}\Gamma'\left(\frac{\nu}{2}\right) \\
\frac{\sum_{i=1}^{n}\ln x_i - n\ln 2}{n} &= \frac{\Gamma'\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}
\end{aligned}
$$

Due to the complexity of the Gamma function in this equation, no solution can be developed for ν in closed form. Thus, we have to rely on numerical methods to obtain a solution to the equation and find the maximum likelihood estimator.

6 Theorems for the Chi-Square Distribution

6.1 Validity of the Distribution

$$
\int_0^\infty \frac{x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}dx = 1
$$

Proof:

$$
\int_0^\infty \frac{x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}dx = \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\int_0^\infty x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}dx
$$

$$
\int_0^\infty x^{\alpha-1}e^{-\frac{x}{\beta}}dx = \beta^\alpha\Gamma(\alpha)
$$

$$
\begin{aligned}
&= \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\left[2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})\right] \\
&= \frac{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \\
&= 1
\end{aligned}
$$

■

6.2 Sum of Chi-Square Random Variables

Let $X_1, X_2, \ldots, X_n$ be independent Chi-Square random variables with parameter ν_i, that is $X_i \sim \chi^2(\nu_i),\ i = 1, 2, \ldots, n$.

Suppose $Y = \sum_{i=1}^{n} X_i$.

Then $Y \sim \chi^2(\sum_{i=1}^{n}\nu_i)$.

_Proof:

$$
\begin{aligned}
M_Y(t) &= E(e^{tY} = E(e^{t(X_1+X_2+\cdots+X_n})\\
&= E(e^{tX_1}e^{tX_2}\cdots e^{tX_n})\\
&= E(e^{tX_1})E(e^{tX_2})\cdots E(e^{tX_n})\\
&= (1-2t)^{-\frac{\nu_1}{2}}(1-2t)^{-\frac{\nu_2}{2}}\cdots(1-2t)^{-\frac{\nu_n}{2}}\\
&= (1-2t)^{\sum\limits_{i=1}^{n}\nu_i}
\end{aligned}
$$

Which is the mgf of a Chi-Square random variable with parameter $\sum\limits_{i=1}^{n}\nu_i$.

Thus $Y \sim \chi^2\left(\sum\limits_{i=1}^{n}\nu_i\right)$. ∎

6.3 Square of a Standard Normal Random Variable

If $Z \sim N(0,1)$, then $Z^2 \sim \chi^2(1)$.

Proof:

$$
\begin{aligned}
M_{Z^2}(t) &= E(e^{tZ^2}) \\
&= \int_{-\infty}^{\infty} e^{tz^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz^2} e^{-\frac{z^2}{2}} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}(-2t+1)} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}(1-2t)} dz
\end{aligned}
$$

$$\int_{-\infty}^{\infty} f(x)dx = 2\int_{0}^{\infty} f(x)dx$$

when $f(x)$ is an even function (**??**)

$$= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{z^2}{2}(1-2t)} dz$$

$$
\begin{aligned}
\text{Let } u &= \frac{z^2}{2}(1-2t) \\
\Rightarrow z &= \frac{\sqrt{2}u^{\frac{1}{2}}}{(1-2t)^{\frac{1}{2}}} \\
\text{So } dz &= \frac{\sqrt{2}u^{-\frac{1}{2}}}{2(1-2t)^{\frac{1}{2}}}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-u} \frac{\sqrt{2}u^{-\frac{1}{2}}}{2(1-2t)^{\frac{1}{2}}} du \\
&= \frac{2\sqrt{2}}{2\sqrt{2\pi}(1-2t)^{\frac{1}{2}}} \int_{0}^{\infty} e^{-u} u^{-\frac{1}{2}} du \\
&= \frac{2\sqrt{2}}{2\sqrt{2\pi}(1-2t)^{\frac{1}{2}}} \int_{0}^{\infty} u^{\frac{1}{2}-1} e^{-u} du
\end{aligned}
$$

$$\int_{0}^{\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \beta^{\alpha}\Gamma(\alpha)$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{\pi}(1-2t)^{\frac{1}{2}}} \Gamma(\frac{1}{2}) \\
&= \frac{\sqrt{\pi}}{\sqrt{\pi}(1-2t)^{\frac{1}{2}}} \\
&= \frac{1}{(1-2t)^{\frac{1}{2}}} = (1-2t)^{-\frac{1}{2}}
\end{aligned}
$$

Which is the mgf of a Chi-Square random variable with 1 degree of freedom. Thus $Z^2 \sim \chi^2(1)$. ∎

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

1 Fundamental Concepts and Definitions

Much of this chapter is taken from the lectures of Dr. John Brunette, University of Southern Maine.

A *function* is a collection of ordered pairs in which no two pairs have the same first element.

The set of all *first* members of the pairs is called the *domain*.

The set of all *second* members of the pairs is called the *range*.

Suppose now that for any function f we have two items x and y such that $x \in dom(f)$ and $y \in ran(x)$ where $dom(f)$ and $ran(f)$ denote the domaain and range of f, respectively. It is said that f maps x onto y, written

$$f:\ x \mapsto y$$

It is common to write the $ran(f)$ as some expression of x. For example, $f : x \mapsto x^2$ takes each element in the domain, and pairs it with it's square. The common shorthand for this is $f(x) = x^2$, meaning that whatever appears between the parentheses following the f is to be squared.

1.1 Function Operations

The three basic operations that can be performed on functions are addition, multipilication, and composition. For any two functions f and g these operations are defined as:\

Addition	$[f+g](x) =: \big\{ \big(x, f(x) + g(x)\big) \vert x \in dom(f) \cap dom(g) \big\}$
Multiplication	$[f \cdot g](x) := \big\{ \big(x, f(x) \cdot g(x)\big) \vert x \in dom(f) \cap dom(g) \big\}$
Composition	$[f \circ g](x) = \big\{ \big(x, f\big(g(x)\big) \vert x \in dom(g) \text{ and } g(x) \in dom(f) \big\}$

Notice that the composition $[f \circ g](x) = f \circ g : g(x) \mapsto f(x)$. In other words, the result of g is then applied to f to produce the result of the composition.

2 Identities and Inverses

Recall that addition and multiplication have identity properties. Specifically, for any real number x, applying one of these identities returns the value x, i.e. $x + 0 = x$ and $x \cdot 1$=x. Functions also have an identity, denoted $id(x)$, that is defined as

$$id: \ x \mapsto x$$

Furthermore, the composition of id with f behaves in this way:

$$id \circ f = f \circ id = f$$

Functions may also exhibit the property of inverses that are exhibited by addition and multiplication. In the latter two, combining any real number x and its inverse returns the identity of that operation, i.e. $x + -x = 0$ and $x \cdot x^{-1} = 1, \ x \neq 0$. Likewise, some functions have an inverse function. If a function f has an inverse f^{-1}, then

$f \circ f^{-1} = f^{-1} \circ f = id.$

On closer observation, we see $f^{-1} \circ f\big(dom(x)\big) = f^{-1}\Big(f\big(dom(x)\big)\Big) = f^{-1}\big(ran(x)\big) = dom(x)$. So f^{-1} must be the set of all ordered pairs (y, x) where $x \in dom(x)$ and $y \in ran(x)$, i.e. $f^{-1}(x) = \{(y, x) | x \in dom(x)$ and $y \in ran(x)\}$. By the definition of functions, no two first elements in f^{-1} can be the same. But the first elements in f^{-1} are the second elements in f. So f^{-1} only exists if no two second elements in f are the same. We thus make the following definition:

A function f is called a **one-to-one** function if it has no two ordered pairs with the same second element.

For any one-to-one function f, no two of the first elements are the same, and no two of the second elements are the same. Thus, f^{-1} is a function, because no two of its first elements are the same, and because the range of f^{-1} is the domain of f, no two second elements in f^{-1} are the same, and f^{-1} is a one-to-one function. Thus, every one-to-one function has an inverse.

If a function f is not one-to-one, however, then there exist two pairs in f that have the same second element. The inverse f^{-1} therefore has two pairs where the first element is the same. When such is the case, the definition of a function is violated, and f^{-1} cannot be a function. Thus, if a function is invertible, it must be one-to-one.

3 Odd and Even Functions

A function is said to be *even* if for any real number $x, \ f(-x) = f(x)$.

A function is said to be *odd* if for any real number $x, \ f(-x) = -f(x)$.

If neither of these criteria are met, the function is simply said to be neither odd nor even.

4 Theorems for Functions

4.1 Operations on Even Functions

Let f and g both be even functions. Then:

i. $[f + g](x)$ is an even function
ii. $[f \cdot g](x)$ is an even function
iii. $[f \circ g](x)$ is an even function.

Proof:

i. $[f+g](x)$ is an even function

$$\begin{aligned}(f+g)(-x) &= f(-x)+g(-x)\\ &= f(x)+g(x)\\ &= [f+g](x)\end{aligned}$$

---

so $[f+g](x)$ is an even function.

ii. $[f \cdot g](x)$ is an even function

$$\begin{aligned}(f \cdot g)(-x) &= f(-x)\cdot g(-x)\\ &= f(x)\cdot g(x)\\ &= [f \cdot g](x)\end{aligned}$$

so $[f \cdot g](x)$ is an even function.

---

iii. $[f \circ g](x)$ is an even function.

$$\begin{aligned}(f \circ g)(-x) &= f\big(g(-x)\big)\\ &= f\big(g(x)\big)\\ &= [f \circ g](x)\end{aligned}$$

so $[f \circ g](x)$ is an even function. ■

4.2 Operations on Odd Functions

Let f and g both be even functions. Then:

i. $[f+g](x)$ is an odd function
ii. $[f \cdot g](x)$ is an even function
iii. $[f \circ g](x)$ is an odd function.

Proof:

_i. $[f+g](x)$ is an odd function

$$\begin{aligned}[f+g](-x) &= f(-x)+g(-x)\\ &= -f(x)-g(x)\\ &= -[f+g](x)\end{aligned}$$

so $[f+g](x)$ is an odd function.

---

ii. $[f \cdot g](x)$ is an even function

$$\begin{aligned}[f \cdot g](-x) &= f(-x)\cdot g(-x)\\ &= -f(x)\cdot -g(x)\\ &= f(x)\cdot g(x)\\ &= [f \cdot g](x)\end{aligned}$$

so $[f \cdot g](x)$ is an even function.

---

iii. iii. $[f \circ g](x)$ *is an odd function.*

$$\begin{aligned}[f \circ g](-x) &= f\big(g(-x)\big) \\ &= f\big(-g(x)\big) \\ &= -f\big(g(x)\big) \\ &= -[f \circ g](x)\end{aligned}$$

so $[f \circ g](x)$ is an odd function. ∎ \end{itemize}

4.3 Operations on an Odd and Even Function

Let f be an even function and let g both be an odd function. Then:

i. $[f+g](x)$ is neither an odd nor an even function
ii. $[f \cdot g](x)$ is an odd function
iii. $[f \circ g](x)$ is an even function
iv. $[g \circ f](x)$ is an even function.

Proof:

i. $[f+g](x)$ *is neither an odd nor an even function*

$$\begin{aligned}[f+g](-x) &= f(-x) + g(-x) \\ &= -f(x) - g(x)\end{aligned}$$

so $[f+g](x)$ is neither an odd nor an even function.

---

ii. $[f \cdot g](x)$ *is an odd function*

$$\begin{aligned}[f \cdot g](-x) &= f(-x) \cdot g(-x) \\ &= f(x) \cdot -g(x) \\ &= -\big(f(x) \cdot g(x)\big) \\ &= -[f \cdot g](x)\end{aligned}$$

so $[f \cdot g](x)$ is an odd function.

---

iii. $[f \circ g](x)$ *is an even function*

$$\begin{aligned}[f \circ g](-x) &= f\big(g(-x)\big) \\ &= f\big(-g(x)\big) \\ &= f\big(g(x)\big) \\ &= [f \circ g](x)\end{aligned}$$

so $[f \circ g](x)$ is an even function. ■

iv. $[g \circ f](x)$ *is an even function.*

$$\begin{aligned}[g \circ f](-x) &= g\big(f(-x)\big) \\ &= g\big(f(x)\big) \\ &= [g \circ f](x)\end{aligned}$$

so $[f \circ g](x)$ is an even function. ■

4.4 Derivatives and Anti-derivatives of Odd Functions

Let f be an odd function and let f' and F denote the derivative and anti-derivative of f, respectively. Then f' and F are both even functions.

Proof:

$$\begin{aligned}f(-x) &= -f(x) \\ \frac{d}{dx}\big[f(-x)\big] &= \frac{d}{dx}\big[-f(x)\big] \\ f'(-x)\cdot -1 &= -f'(x) \\ -f'(-x) &= -f'(x) \\ f'(-x) &= f'(x)\end{aligned}$$

So f' is an even function.

$$\begin{aligned}f(-x) &= -f(x) \\ \int f(-x) &= \int -f(x) \\ F(-x)\cdot -1 &= -F(x) \\ -F(-x) &= -F(x) \\ F(-x) &= F(x)\end{aligned}$$

So F is also an even function. ■

4.5 Derivatives and Anti-derivatives of Even Functions

Let g be an even function, and let g' and G denote the derivative and anti-derivative of g, respectively. Then g' and G are both odd functions.

Proof:

$$\begin{aligned}g(-x) &= g(x) \\ \frac{d}{dx}\big[g(-x)\big] &= \frac{d}{dx}\big[g(x)\big] \\ g'(-x)\cdot -1* &= g'(x) \\ -g'(-x) &= g'(x) \\ g'(-x) &= -g'(x)\end{aligned}$$

So g' is an odd function.

$$\begin{aligned} g(-x) &= g(x) \\ \int g(-x) &= \int g(x) \\ G(-x) \cdot -1 &= G(x) \\ -G(-x) &= G(x) \\ G(-x) &= -G(x) \end{aligned}$$

So G is also an odd function. ■

The Geometric Series

Integraion: Techniques and Theorems

Logarithmic and Exponential Functions

The Real Number System