The Exponential Class of pdfs.

Let \( \{ f(x|\theta), \theta \in \Omega \} \) \( \Omega = \{ \theta : 0 < \theta < \infty \} \)

where \( \Gamma \) and \( S \) are known constants and

\[
f(x|\theta) = e^{(\theta \Gamma(x) + S(x) - q(\theta))} \quad \text{if } x \in S, \text{ zero otherwise}
\]

where

1) \( S \), the support of \( X \), does not depend upon \( \theta \).
2) \( p(\theta) \) is a nontrivial continuous function of \( \theta \in \Omega \).
3) \( c \) if \( X \) is a continuous random variable then \( k(x) \neq 0 \) and \( S(x) \) are continuous functions of \( x \in S \).
4) \( c \) If \( X \) is a discrete random variable then \( k(x) \) is a nontrivial function of \( x \in S \).

The family \( \{ f(x|\theta), \theta \in \Omega \} \) is said to be a regular case of the exponential class of distributions.

Example: Poisson \( (\theta) \). \( f(x|\theta) = \frac{e^{-\theta} \theta^x}{x!} = e^{(\theta x - \theta x! - \theta)} \), \( x \in S \)

\( S = \{ x : x = 0, 1, 2, \ldots \} \), \( \Omega = \{ \theta : 0 < \theta < \infty \} \)

\( p(\theta) = -\theta \), \( k(x) = x \), \( S(x) = -x! \), \( q(\theta) = \theta \)

- Poisson family is a regular exponential family.

Example: \( N(0, \theta) \). \( \Omega = \{ \theta : 0 < \theta < \infty \} \)

\( f(x|\theta) = \frac{1}{\sqrt{2\pi \theta}} e^{-\frac{x^2}{2\theta}} = e^{(\theta x - \theta \sqrt{2\pi} x - \frac{1}{2} \theta)} \), \( x \in S \).

\( S = \{ x : -\infty < x < \infty \} \), \( p(\theta) = -\frac{1}{2\theta} \), \( k(x) = x^2 \), \( S(x) = 0 \), \( q(\theta) = -\theta \sqrt{2\pi} \theta \)

- \( N(0, \theta) \) family is a regular exponential family.
Theorem. Let $X_1, \ldots, X_n$ be a random sample from a regular exponential family. Let $Y_1 = \sum_{i=1}^{n} K(X_i)$.

1) $Y_1$ is a complete sufficient statistic for $\theta$.

2) $E(Y_1) = -n \frac{q'(-\theta)}{p'(-\theta)}$

3) $V(Y_1) = n \frac{1}{[p'(-\theta)]^2} \left( p''(-\theta) q'(-\theta) - q''(-\theta) p'(-\theta) \right)$

A function of $Y_1$, say $\psi(Y_1)$ such that $E[\psi(Y_1)] = \theta$ is the unique MLE of $\theta$.

**Example.** Poisson $(\theta)$.

$f(x|\theta) = \frac{e^{-\theta} \theta^x}{x!} = e^{(\theta x - \theta)}$, $x = 0, 1, 2, \ldots$

$K(x) = x$

$Y_1 = \sum_{i=1}^{n} K(X_i) = \sum_{i=1}^{n} X_i$ is a complete sufficient statistic for $\theta$. $E \left( \frac{Y_1}{n} \right) = \theta$. $\psi(Y_1) = \frac{Y_1}{n} = \bar{x}$ is the unique MLE of $\theta$.

**Functions of a parameter.**

We want to find the MLE of a function of $\theta$.

**Example.** $X \sim \text{Bernoulli}(\theta)$.

$f(x|\theta) = \theta^x (1-\theta)^{1-x} = e^{(x \ln \theta + (1-x) \ln (1-\theta))} = e^{(x \ln \frac{\theta}{\theta + \frac{1}{1-\theta}}) + \ln (1-\theta)}$

$x = 0, 1$
\[ k(x) = x. \] Thus \( Y = \frac{\sum x_i}{n} \) is a complete sufficient statistic for \( \theta \). E(\( Y \)) = \theta, \ V(\( Y \)) = n\theta(1-\theta).

\( \Psi(Y) = \frac{Y}{n} = \bar{x} \) is the unique MVUE of \( \theta \).

What is the MVUE of \( \frac{\theta(1-\theta)}{n} \), the variance of \( \frac{Y}{n} \)?

\[
E\left[ \frac{\bar{Y} (1 - \bar{Y})}{n} \right] = E\left( \frac{\bar{Y}}{n} \right) - E\left( \frac{\bar{Y}^2}{n} \right) = \frac{\theta}{n} - \frac{\theta(1-\theta) + n\theta^2}{n^2} = \left( \frac{n-1}{n} \right) \theta (1 - \theta)
\]

\[
\therefore E\left[ \frac{\frac{Y}{n} (1 - \frac{Y}{n})}{n-1} \right] = \theta (1 - \theta) \quad \text{Thus} \quad \frac{\sum (Y - \bar{Y})}{n-1} \text{is the unique MVUE of} \ \frac{\theta(1-\theta)}{n}.
\]

**Example.** Let \( X \) be a sample of size one from Poisson(\( \theta \)). Find the MVUE of \( e^{-2\theta} \).

\( X \) is a complete sufficient statistic for \( \theta \). Let \( \Psi(X) \) be an unbiased estimator of \( e^{-2\theta} \).

\[
E[\Psi(X)] = \sum_{x=0}^{\infty} \Psi(x) \frac{e^{-\theta} \theta^x}{x!} = e^{-2\theta}
\]

\[
\rightarrow \sum_{x=0}^{\infty} \Psi(x) \frac{\theta^x}{x!} = e^{-\theta}
\]

Since \( \sum_{x=0}^{\infty} (-1)^x \frac{\theta^x}{x!} = e^{-\theta} \), we have

\( \Psi(x) = (-1)^x \begin{cases} -1 & \text{if } x \text{ is odd} \\ 1 & \text{if } x \text{ is even} \end{cases} \)

In this case MVUE of \( e^{-2\theta} \) is a very poor estimator.
Example. Let $X_1, \ldots, X_n$ be a random sample from $\text{Poisson} (\theta)$. Find the \textit{mvue} of $\theta \sum_{i=1}^n \frac{1}{\lambda} \left( 1 + \theta \right) e^{-\theta}$. 

$Y = \sum_{i=1}^n X_i$ is a complete sufficient statistic for $\theta$. 

Let $U(X_i) = \begin{cases} 1 & \text{if } X_i \leq 1 \\ 0 & \text{otherwise} \end{cases}$

$U(X_i) \sim \text{Bernoulli} \left( \left( 1 + \theta \right) e^{-\theta} \right)$, \quad $E[U(X_i)] = \left( 1 + \theta \right) e^{-\theta}$

Let $\Phi(y) = E \left[ U(X_i) \mid Y = y \right]$, \quad $E[\Phi(Y)] = E[E[U(X_i) \mid Y = y]] = \left( 1 + \theta \right) e^{-\theta}$

$\Phi(y)$ is the \textit{mvue} of $(1 + \theta) e^{-\theta}$. 

Conditional distribution of $X_1, \ldots, X_n$ given $Y = y$:

$$f(x_1, \ldots, x_n \mid y) = \frac{e^{-\lambda y} \lambda^{x_1} \cdots \lambda^{x_n}}{\lambda^{x_1} \cdots \lambda^{x_n} y!} \cdot \frac{y!/\lambda^{x_1} \cdots \lambda^{x_n}}{y!} \cdot \frac{\lambda^{x_1} \cdots \lambda^{x_n}}{y!} \cdot \frac{1}{\lambda^{x_1} \cdots \lambda^{x_n}}$$

$X_1, \ldots, X_n \mid Y \sim \text{multinomial} \left( y, \frac{1}{\lambda}, \frac{1}{\lambda}, \ldots, \frac{1}{\lambda} \right)$

Marginal distribution of $X_i$ given $Y = y$ is $\text{Bino} \text{mal} \left( y, \frac{1}{\lambda} \right)$

$$\Phi(y) = E \left[ U(X_i) \mid Y = y \right] = \sum_{x_i = 0}^y U(x_i) \left( \frac{y}{\lambda} \right)^{x_i} \left( 1 - \frac{y}{\lambda} \right)^{y-x_i}$$

$$= \frac{y}{\lambda} \left( \frac{y}{\lambda} \right)^{x_i} \left( 1 - \frac{y}{\lambda} \right)^{y-x_i}$$

$$= \left( 1 - \frac{y}{\lambda} \right)^{y-x_i} + \frac{y}{\lambda} \left( 1 - \frac{y}{\lambda} \right)^{y-1}$$

$$= \left( 1 - \frac{y}{\lambda} \right)^y \left( \frac{y}{\lambda-1} \right)$$

is the \textit{umve} of $(1 + \theta) e^{-\theta}$. 

Example. Let \( X_1, \ldots, X_n \) be a random sample from Poisson\((\theta)\). Find the \textit{mvue} of \( e^{-2\theta} \).

Let \( U(X_1, X_2) = \begin{cases} 1 & \text{if } X_1 = 0 \text{ and } X_2 = 0 \\ 0 & \text{otherwise} \end{cases} \)

\( U(X_1, X_2) \sim \text{Bernoulli}(e^{-2\theta}) \). \( E[U(X_1, X_2)] = e^{-2\theta} \)

\( Y = \sum_{i=1}^n X_i \) is a complete sufficient statistic for \( \theta \).

\( X_1, \ldots, X_n \mid Y = y \sim \text{Multinomial}(y; \frac{1}{2}, \ldots, \frac{1}{2}) \)

\( X_1, X_2 \mid Y = y \sim \text{Multinomial}(y; \frac{1}{2}, \frac{1}{2}, 1-\frac{y}{2}) \)

\( f(x_1, x_2 \mid y) = \frac{y!}{x_1! x_2! (y-x_1-x_2)!} \left( \frac{1}{2} \right)^{x_1} \left( \frac{1}{2} \right)^{x_2} \left( 1-\frac{y}{2} \right)^{y-x_1-x_2} \)

Let \( \varphi(y) = E[U(X_1, X_2) \mid Y = y] \).

\( E[\varphi(y)] = E[U(X_1, X_2)] = e^{-2\theta} \). \( \varphi(y) \) is the \textit{mvue} of \( e^{-2\theta} \).

\( \varphi(y) = E[U(X_1, X_2) \mid Y = y] = \sum_{x_1} \sum_{x_2} u(x_1, x_2) f(x_1, x_2 \mid y) \)

\( = u(0, 0) \frac{y!}{0! 0! y!} \left( \frac{1}{2} \right)^y \left( \frac{1}{2} \right)^y \left( 1-\frac{y}{2} \right)^y \), \( u(0, 0) = 1 \)

\( = (1-\frac{y}{2})^y = (1-\frac{2}{n})^\bar{y} \) is the \textit{mvue} of \( e^{-2\theta} \).
The case of several parameters.

Definition: Let \( X = (X_1, \ldots, X_n) \) be a random sample of size \( n \) from a distribution with pdf \( f(x|\theta) \) where \( \theta = (\theta_1, \ldots, \theta_p) \in \mathbb{R}^p \). Let \( S \) denote the support of \( X \). Let \( Y = (Y_1, \ldots, Y_m) \) be \( m \) statistics with joint pdf \( f_Y(y|\theta) \). \( Y_1, \ldots, Y_m \) are called joint sufficient statistics for \( \theta_1, \ldots, \theta_p \) if \( Y \) is jointly sufficient for \( \theta \) if and only if

\[
\prod_{i=1}^{n} \frac{f(x_i|\theta)}{f_Y(y|\theta)} = H(X) \quad \text{for all } x_i \in S
\]

where \( H(X) \) does not depend upon \( \theta \).

In general \( m \neq p \), but in most cases \( m = p \).

Factorization theorem: \( Y \) is jointly sufficient for \( \theta \) if and only if

\[
\prod_{i=1}^{n} f(x_i|\theta) = k_1(y, \theta) \cdot k_2(x) \quad \text{for all } x_i \in S
\]

where \( k_2(x) \) does not depend upon \( \theta \).
Definition: A family of pdf's \( \{ h(z|\theta) \} \) \( z = (z_1, \ldots, z_m) \), \( \theta = (\theta_1, \ldots, \theta_p) \) is said to be complete if for every real valued function \( U \), \( \mathbb{E}[U(z)] = 0 \) for each \( \theta \in \Omega \) implies \( U(z) = 0 \) at each point \( z \) at which at least one member of the family of pdf's is positive.

Let \( Y = (Y_1, \ldots, Y_m) \) be jointly sufficient for \( \theta \). Let \( f_Y(y|\theta) \) be the pdf of \( Y \). If \( f_Y(y|\theta) \) is complete then \( Y \) is said to be a complete sufficient statistic for \( \theta \). If \( \mathbb{E}[\Psi(Y)] = g(\theta) \) then \( \Psi(Y) \) is the unique MLE of \( g(\theta) \).

Example. Let \( X_1, \ldots, X_n \) be a random sample from \( \text{Uniform}(\theta_1 - \theta_2, \theta_1 + \theta_2) \).

\[
f(x|\theta) = \frac{1}{2\theta_2}, \quad \theta_1 - \theta_2 < x < \theta_1 + \theta_2, \text{ zero elsewhere.}
\]

\( \theta = (\theta_1, \theta_2) \in \Omega = \{(\theta_1, \theta_2) : -\infty < \theta_1 < \infty, 0 < \theta_2 < \infty \} \)

Let \( Y_1 = \min(X_1, \ldots, X_n) \), \( Y_n = \max(X_1, \ldots, X_n) \)
\[ f_{y_1,y_n}(y_1,y_n | \theta) = \frac{n(n-1)}{(2\theta_x)^n} (y_n - y_i)^{n-1}, \theta_i < y_i < \theta_n < \theta_1 + \theta_x \]

\[ \prod_{i=1}^{n} f(x_i | \theta) = \left( \frac{1}{2\theta_x} \right)^n \prod_{i=1}^{n} I(y_i) I(y_i) I(y_1) I(y_n) (\theta_1, \theta_2, \theta_n) (x_i, \theta_1 + \theta_x) \]

\[ = \frac{1}{n(n-1)(y_n - y_i)^{n-2}} = H(x) \text{ does not depend upon } \theta \]

\( Y_1, Y_n \) are jointly sufficient for \( \theta_1, \theta_x \).

The Exponential class of pdfs.

Let \( \{ f(x | \theta), \theta \in \mathbb{R}^d \} \)

\( \Theta = (\theta_1, ... , \theta_m) \in \mathbb{R} = \{ (\theta_1, ... , \theta_m) : \theta_i < \theta_i < \theta_i \} \)

where \( \theta_i \) and \( \theta_i \) i=1,...,m are known constants and

\[ f(x | \theta) = e^{\sum_{j=1}^{m-2} p_j(\theta) s_j(x) + s(x) + q(\theta)} \]

where

1) \( S \), the support of \( X \), does not depend upon \( \theta \).

2) \( p_j(\theta) j=1,...,m \) are non-trivial, functionally independent, continuous functions of \( \theta \).
3) and

a) if $X$ is a continuous r.v. then $K_j^l(x)$ $j=1,\ldots,m$ are continuous functions of $x \in S$, and no one is a linear homogeneous function of the others and $S(x)$ is a continuous function of $x \in S$.

b) if $X$ is a discrete r.v. then $K_j(x)$ $j=1,\ldots,m$ are non-trivial function of $x \in S$ and no one is a linear homogeneous function of the others.

The family $\{f(x|\theta), \theta \in \Theta\}$ is said to be a regular case of the exponential class of distributions.

**Theorem.** Let $X_1, \ldots, X_n$ be a random sample from a regular exponential family $\{f(x|\theta), \theta \in \Theta\}$ $\Theta = \{\Theta_1, \ldots, \Theta_m\}$. Let $Y_1 = \sum_{i=1}^n K_1(x_i)$, $Y_2 = \sum_{i=1}^n K_2(x_i)$, \ldots, $Y_m = \sum_{i=1}^n K_m(x_i)$. $Y_1, Y_2, \ldots, Y_m$ are joint complete sufficient statistics for $\Theta_1, \Theta_2, \ldots, \Theta_m$.

**Example:** Let $X_1, \ldots, X_n$ $(n > 2)$ be a random sample from $N(\Theta_1, \Theta_2)$. 
\[ \theta = (\theta_1, \theta_2) \in \mathcal{N} = \mathcal{N}(\theta_1, \theta_2), \quad -\infty < \theta_1 < \infty, \quad \alpha < \theta_2 < \infty \]

\[ f(x \mid \theta) = \frac{1}{\sqrt{2\pi \theta_2}} \exp \left( -\frac{(x - \theta_1)^2}{2\theta_2} \right) \]

\[ \begin{align*}
&\quad = \exp \left( -\frac{1}{2\theta_2} x^2 + \frac{\theta_1}{\theta_2} x - \frac{\theta_1^2}{2\theta_2} - \ln(2\pi \theta_2) \right), \\
&\text{where} -\infty < x < \infty
\end{align*} \]

\[ K_1(x) = x^2, \quad K_2(x) = x, \quad S(x) = 0 \]

\[ p_1(\theta) = -\frac{1}{2\theta_2}, \quad p_2(\theta) = \frac{\theta_1}{\theta_2}, \quad q(\theta) = -\frac{\theta_1^2}{2\theta_2^2} - \ln(2\pi \theta_2) \]

\[ Y_1 = \sum_{i=1}^{n} K_1(x_i) = \sum_{i=1}^{n} x_i^2 \]

\[ Y_2 = \sum_{i=1}^{n} K_2(x_i) = \sum_{i=1}^{n} x_i \theta_2 \]

\[ Y_1, Y_2 \text{ are complete sufficient statistics for } \theta_1, \theta_2 \]

Let

\[ \begin{align*}
\phi_1(Y_1, Y_2) &= \frac{Y_1}{n} = \bar{x} \hspace{2cm} \phi_2(Y_1, Y_2) = Y_1 - \frac{Y_2}{n} = \frac{\sum (x_i - \bar{x})^2}{n-1} \\
E \left[ \phi_1(Y_1, Y_2) \right] &= \theta_1 \hspace{2cm} E \left[ \phi_2(Y_1, Y_2) \right] = \theta_2
\end{align*} \]

Thus \( \phi_1(Y_1, Y_2) \approx \bar{x} \), \( \phi_2(Y_1, Y_2) \approx s^2 \)

are the unique minimum variance unbiased estimators of \( \theta_1 \) and \( \theta_2 \) respectively.
Minimal Sufficient Statistics

It is always true that the complete sample \( X = (X_1, \ldots, X_n) \) is a sufficient statistic. Also, any one-to-one function of a sufficient statistic is a sufficient statistic. Because of the numerous sufficient statistics in a problem, we might ask whether one sufficient statistic is any better than another. The purpose of a sufficient statistic is to achieve data reduction without loss of information about the parameter \( \theta \). Thus, a statistic that achieves the most data reduction while still retaining all the information about \( \theta \) might be considered preferable.

**Definition.** Minimal sufficient statistics are those that are sufficient for the parameters and are functions of every other set of sufficient statistics for those same parameters.

Using the above definition to find a minimal sufficient statistic is impractical. The following result gives an easier way to find a minimal sufficient statistic.
Theorem: Let \( X = (X_1, \ldots, X_n) \) and \( \bar{Z} = (Z_1, \ldots, Z_n) \) denote random samples of size \( n \) from a distribution with pdf \( f(x | \Theta) \) where \( \Theta = (\Theta_1, \ldots, \Theta_p) \in \mathbb{R}^p \). Suppose there exists functions \( U_1, \ldots, U_m \) such that the ratio
\[
\prod_{i=1}^{\bar{n}} f(x_i | \Theta) / \prod_{i=1}^{\bar{n}} f(z_i | \Theta)
\]
is constant as a function of \( \Theta \) if
\[
\prod_{i=1}^{\bar{n}} f(x_i | \Theta) / \prod_{i=1}^{\bar{n}} f(z_i | \Theta)
\]
and only if \( U_i(x) = U_i(z) \) for \( i = 1, \ldots, m \). Then \( U_1(x), \ldots, U_m \) are minimal sufficient statistics for \( \Theta_1, \ldots, \Theta_p \).

Example: \( X \sim N(\Theta_1, \Theta_2) \) both \( \Theta_1 \) and \( \Theta_2 \) are unknown.

\[
\prod_{i=1}^{\bar{n}} f(x_i | \Theta) / \prod_{i=1}^{\bar{n}} f(z_i | \Theta) = \left( \frac{1}{\sqrt{2\pi \Theta_2}} \right)^n e^{-\frac{1}{2\Theta_2} \sum (x_i - \Theta_1)^2}
\]

\[
\prod_{i=1}^{\bar{n}} f(z_i | \Theta) / \prod_{i=1}^{\bar{n}} f(x_i | \Theta) = \left( \frac{1}{\sqrt{2\pi \Theta_2}} \right)^n e^{-\frac{1}{2\Theta_2} \sum (z_i - \Theta_1)^2}
\]

\[
\left( \frac{1}{\sqrt{2\pi \Theta_2}} \right)^n e^{-\frac{1}{2\Theta_2} \sum (x_i - \Theta_1)^2}
\]

\[
= \frac{1}{2\Theta_2} \left[ n (\bar{x} - \Theta_1)^2 + (n-1) s_x^2 \right]
\]

where \( \bar{x} = \frac{\sum x_i}{n}, \bar{z} = \frac{\sum z_i}{n}, s_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}, s_z^2 = \frac{\sum (z_i - \bar{z})^2}{n-1} \)
The ratio will be constant as a function of \( \theta_1 \) and \( \theta_2 \) if and only if \( \bar{X} = \bar{Z} \) and \( S^2_x = S^2_z \). Thus \((\bar{X}, S^2)\) is a minimal sufficient statistic for \((\theta_1, \theta_2)\).

**Example.** \( X \sim \text{Uniform}(\theta, \theta+1), -\infty < \theta < \infty \).

\[
\frac{\prod_{i=1}^{n} f(x_i; \theta)}{\prod_{i=1}^{n} f(z_i; \theta)} = \frac{\prod_{i=1}^{n} I(x_i; \theta, \theta+1)}{\prod_{i=1}^{n} I(z_i; \theta, \theta+1)} = \frac{\prod_{i=1}^{n} I(\theta)}{\prod_{i=1}^{n} I(\text{max}(x_i)-1, \text{min}(x_i))}
\]

The ratio is constant and in fact equals 1 if \( \text{min}(x_i) = \text{min}(z_i) \), \( \text{max}(x_i) = \text{max}(z_i) \). Thus \((\text{min}(X_i), \text{max}(X_i))\) are a minimal sufficient statistics for \( \theta \).

Often if there are \( k \) parameters, we can find \( k \) joint sufficient statistics that are minimal. In particular, if there is one parameter, we can often find a single sufficient statistic which is minimal but this is not always the case, as the above example shows.

A minimal sufficient statistic is not unique. Any one-to-one function of a minimal sufficient statistic is also a minimal sufficient statistic.
Example: $X \sim \text{Poisson}(\theta)$

\[
\prod_{i=1}^{n} f_X(x_i;\theta) \overset{\prod_{i=1}^{n} e^{-\theta} \frac{\theta^{x_i}}{x_i!}}{=} \prod_{i=1}^{n} e^{-\theta} \frac{\theta^{z_i}}{z_i!} \overset{\prod_{i=1}^{n} e^{-\theta} \frac{\theta^{z_i}}{z_i!}}{=} \frac{\theta^{\sum x_i/n}}{n^{\sum z_i/n}}
\]

This ratio will be constant as a function of $\theta$ if and only if $\sum x_i = \sum z_i$. Thus $\sum x_i$ is a minimal sufficient statistic for $\theta$. $\bar{X}$ is also a minimal sufficient statistic for $\theta$.

Theorem: If the MLE of $\theta$ exists and is unique then it is a function of a minimal sufficient set of statistics.

Ancillary Statistics

Sufficient statistics, in a sense, contain all the information about $\theta$ that is available in the sample. We now introduce a different sort of statistic, one that has a complementary purpose.

Definition: A statistic $U(\bar{X})$ whose distribution does not depend on the parameter $\theta$ is called an ancillary statistic.

Example: $N(\theta, 1)$. Sample variance $S^2$ is an ancillary statistic for $\theta$. 
Alone, an ancillary statistic contains no information about $\theta$. Paradoxically, an ancillary statistic, when used in conjunction with other statistics, sometimes does contain valuable information for inferences about $\theta$.

**Rules for finding ancillary statistics:**

1. **Location-Invariant Statistic.**

   Let $f(x)$ be any pdf. Then the family of pdfs
   
   \[ f(x|\theta) = f(x-\theta) \]
   
   is called a location family with standard pdf $f(x)$ and $\theta$ is called a location parameter of the family.

   Let $X_1, \ldots, X_n$ be a random sample from a distribution with pdf of the form $f(x-\theta)$, that is $\theta$ is a location parameter. If $Z = U(X_1, \ldots, X_n)$ is a statistic such that
   
   \[ U(X_1+d, \ldots, X_n+d) = U(X_1, \ldots, X_n) \]
   
   for every $d$, then the distribution of $Z$ does not depend upon $\theta$ (i.e., $Z$ is an ancillary statistic) and $Z$ is called a location-invariant statistic.

**Example.**

\[ f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad -\infty < x < \infty \]

\[ f_x(x|\theta) = f(x-\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} \quad -\infty < x < \infty \]
\[ s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}, \quad \bar{x} = \frac{\max(X_1) - \min(X_1)}{n}, \quad X_1 + 2X_2 - 3X_3 \]

are location-invariant statistics.

2) Scale-Invariant Statistics.

Let \( f(x) \) be any pdf. Then for \( \theta > 0 \), the family of pdfs \( f_x(x|\theta) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right) \) indexed by the parameter \( \theta \) is called a scale family with standard pdf \( f(x) \) and \( \theta \) is called the scale parameter of the family.

Let \( X_1, \ldots, X_n \) be a random sample from a distribution with pdf of the form \( \frac{1}{\theta} f\left(\frac{x}{\theta}\right), \theta > 0 \), that is \( \theta \) is a scale parameter. If \( Z = U(X_1, \ldots, X_n) \) is a statistic such that

\[ U(C_1, X_1, \ldots, C \cdot X_n) = U(X_1, \ldots, X_n) \quad \text{for all } C > 0 \]

then the distribution of \( Z \) does not depend upon \( \theta \) (i.e., \( Z \) is an ancillary statistic) and \( Z \) is called a scale-invariant statistic.

Example: \( f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < \infty \).

\[ f_x(x|\theta) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right) = \frac{1}{\sqrt{2\pi} \theta} e^{-\frac{x^2}{2\theta^2}}, -\infty < x < \infty \]

\( \frac{X_1}{X_1 + X_2}, \quad \frac{X_1^2}{\sum_{i=1}^{n} X_i^2}, \quad \min(X_i), \quad \max(X_i) \) are scale-invariant statistics.
3) Location and Scale Invariant Statistics.

Let \( f(x) \) be any pdf. Then for \( \Theta_1, \Theta_2 > 0 \) and \( \Theta_2 > 0 \) the family of pdfs
\[
\frac{1}{\Theta_2} f\left( \frac{x - \Theta_1}{\Theta_2} \right)
\]
indexed by the parameters \( \Theta_1, \Theta_2 \) is called location-scale family with standard pdf \( f(x) \). \( \Theta_1 \) is called the location parameter and \( \Theta_2 \) is called the scale parameter.

Let \( X_1, \ldots, X_n \) be a random sample from a distribution with pdf of the form \( \frac{1}{\Theta_2} f\left( \frac{x - \Theta_1}{\Theta_2} \right) \) where \(-\infty < \Theta_1 < \infty, \Theta_2 > 0\). If \( Z = U(X_1, \ldots, X_n) \) is a statistic such that
\[
U(cX_1 + d, \ldots, cX_n + d) = U(X_1, \ldots, X_n) \quad \text{for} \quad -\infty < d < \infty, \ c > 0
\]
then the distribution of \( Z \) does not depend upon \( \Theta_1 \) and \( \Theta_2 \). The statistic is called a location and scale invariant statistic.

Example: \( f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \) for \(-\infty < x < \infty\)

\[
f(x; \Theta_1, \Theta_2) = \frac{1}{\Theta_2} f\left( \frac{x - \Theta_1}{\Theta_2} \right) = \frac{1}{\sqrt{2\pi} \Theta_2} e^{-\frac{(x - \Theta_1)^2}{2\Theta_2^2}} \quad \text{for} \quad -\infty < x < \infty
\]

\(-\infty < \Theta_1 < \infty, \Theta_2 > 0\).
\[ \frac{\text{max}(X_i) - \text{min}(X_i)}{s}, \quad \frac{X_i - \overline{X}}{s}, \quad \frac{|X_i - X_j|}{s} \] are location and scale invariant statistics.

**Theorem.** Let \( Y \) be a complete sufficient statistic for \( \Theta \). Let \( Z \) be any other statistic (not a function of \( Y \) alone). If the distribution of \( Z \) does not depend upon \( \Theta \), then \( Z \) is independent of \( Y \).

**Example.** \( f(x) = e^{-x}, x > 0 \)

\[ f(x|\Theta) = f(x-\Theta) = e^{-(x-\Theta)}, x > \Theta. \]

\( Y_1 = \text{min}(X_i) \) is a complete sufficient statistic for \( \Theta \). Here, \( \Theta \) is a location parameter then any statistic \( Z = U(X_1, \ldots, X_n) \) such that

\[ U(X_1+d_1, \ldots, X_n+d_n) = U(X_1, \ldots, X_n) \] for all \( d \),

will be independent of \( Y_1 \). Thus,

\( Z = \text{max}(X_i) - \text{min}(X_i) \) and \( Y_1 \) are independent.

\( Z = \frac{\sum(X_i - \overline{X})}{\sqrt{n}} \) and \( Y_1 \) are independent.

\( Z = \frac{\sum(X_i - \text{min}(X_i))}{\sqrt{n}} \) and \( Y_1 \) are independent.

\( Z = X_1 + 3X_2 - X_3 - X_4 - 2X_5 \) and \( Y_1 \) are independent.
Example. \( f(x) = e^{-x} \), \( x > 0 \)

\[
f_X(x|\theta) = \frac{1}{\theta} f\left(\frac{x}{\theta}\right) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x > 0 .
\]

i.e. \( X \sim \text{Exponential}(\theta) \).

\( Y = \sum X_i \) is a complete sufficient statistic for \( \theta \).

Here, \( \theta \) is a scale parameter. Any statistic \( Z = U(X_1, \ldots, X_n) \) such that

\[
U(c, X_1, \ldots, cX_n) = U(X_1, \ldots, X_n), \quad c > 0
\]

will be independent of \( Y \). Thus,

\[
Z = \frac{X_1}{X_n} \quad \text{and} \quad Y \text{ are independent}
\]

\[
Z = \frac{X_1}{X_1 + \ldots + X_n} \quad \text{and} \quad Y \text{ are independent}
\]

\[
Z = \frac{\sum X_i^2}{\sum X_i} \quad \text{and} \quad Y \text{ are independent}
\]

\[
Z = \frac{\min(X_i)}{\max(X_i)} \quad \text{and} \quad Y \text{ are independent}
\]

**Definition.** Let $C$ denote a subset of the sample space. $C$ is called a best critical region of size $\alpha$ for testing the simple hypothesis $H_0: \theta = \theta'$ against the alternative simple hypothesis $H_1: \theta = \theta''$ if for every subset $A$ of the sample space for which $P[C \cap A \mid H_0] = \alpha$

a) $P[C \cap \{X \in A \mid H_0\}] = \alpha$

b) $P[C \cap A \mid H_1] \geq P[C \cap A \mid H_0]$

Neyman-Pearson theorem provides a method of determining a best critical region.

**Neyman-Pearson theorem.** Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample of size $n$ from a distribution with pdf $f(x \mid \theta)$. Then the likelihood function is $L(\theta; \mathbf{X}) = \prod_{i=1}^{n} f(x_i \mid \theta)$. Let $k$ be a positive number. Let $C$ be a subset of the sample space such that

a) $\frac{L(\theta'; \mathbf{X})}{L(\theta''; \mathbf{X})} \leq k$ for each point $X \in C$

b) $\frac{L(\theta'; \mathbf{X})}{L(\theta''; \mathbf{X})} \geq k$ for each point $X \in C^c$

c) $\alpha = P[C \mid H_0]$
Then $C$ is a best critical region of size $\alpha$ for testing the simple hypothesis $H_0: \Theta = \Theta'$ against the alternative simple hypothesis $H_1: \Theta = \Theta''$, where $\Theta'$ and $\Theta''$ are distinct fixed values of $\Theta$ so that $\mathcal{R} = \{ \Theta: \Theta = \Theta', \Theta'' \}$.

The conditions (a), (b) and (c) are necessary and sufficient for region $C$ to be a best critical region of size $\alpha$.

If we let $C = \{ x : \frac{L(\Theta'; x)}{L(\Theta'', x)} \leq k, \ k > 0 \}$

then $C$ will be a best critical region.

**Example.** Let $X = (X_1, \ldots, X_n)$ be a random sample from $N(\Theta, 1)$.

- $H_0: \Theta = 0 \quad \text{or} \quad \Theta' = 0, \ \Theta'' = 1$.
- $H_1: \Theta = 1$

$$
\frac{L(\Theta'; x)}{L(\Theta'', x)} = \left( \frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{\sum x_i^2}{2}} \leq k \rightarrow e^{-\sum x_i^2 + \frac{n}{2}} \leq k
$$

Thus $C' = \{ x : \sum x_i \geq c \} = \{ x : x_1 \geq c_1 \} \quad \text{where} \quad c_1 = \frac{c}{n}$.

The constant $c_1$ can be determined so that the size of the critical region is $\alpha$.

$$
\alpha = \mathbb{P} \left( X \in C \mid H_0 \right) = \mathbb{P} \left( x \geq c_1, \Theta = 0 \right) = \mathbb{P} \left( z \geq \frac{c_1}{\sqrt{n}} \right)
$$

where $Z \sim N(0, 1)$ under $H_0$. 
Thus \( \frac{c_1}{\frac{1}{\sqrt{n}}} = \frac{2}{\sqrt{n}} \) \( \therefore \) \( c_1 = \frac{2}{\sqrt{n}} \frac{1}{\sqrt{n}} \)

We reject \( H_0 \) if \( X \geq \frac{Z}{\frac{1}{\sqrt{n}}} \), that is we reject \( H_0 \) if \( Z = \frac{X}{\frac{1}{\sqrt{n}}} \geq Z_\alpha \).

Let \( X_1, \ldots, X_n \) be random variables. Let \( H_0 \) and \( H_1 \) be two simple hypotheses.

\( H_0 \): Joint pdf of \( X_1, \ldots, X_n \) is \( g(x_1, \ldots, x_n) \)

\( H_1 \): Joint pdf of \( X_1, \ldots, X_n \) is \( h(x_1, \ldots, x_n) \).

\( C \) is a best critical region of size \( \alpha \) for testing \( H_0 \) against \( H_1 \) if for \( k > k_0 \),

a) \( \frac{g(x_1, \ldots, x_n)}{h(x_1, \ldots, x_n)} < k \) for \( (x_1, \ldots, x_n) \in C \)

b) \( \frac{g(x_1, \ldots, x_n)}{h(x_1, \ldots, x_n)} > k \) for \( (x_1, \ldots, x_n) \in C^C \)

c) \( P[(x_1, \ldots, x_n) \in C \mid H_0] = \alpha \).
Uniformly Most Powerful (UMP) Tests:

Definition: The critical region $C$ is a UMP critical region of size $\alpha$ for testing simple $H_0$ against $H_1$ if $C$ is a best critical region of size $\alpha$ for testing $H_0$ against each simple hypothesis in $H_1$.

Note: UMP tests do not always exist. When they exist, the NP theorem provides a technique for finding them.

Example: Let $X = (X_1, \ldots, X_n)$ be a r.v. from $N(0, \theta)$

$H_0: \theta = \theta^*$

$H_1: \theta > \theta^*, \ (\theta^* > 0$ is a given constant$)$

Thus $\Omega = \{ \theta: \theta \geq \theta^* \}$

$$L(\theta, x) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{\sum X_i^2}{2\theta}}$$

Let $\theta'' > \theta^*$ and consider testing simple $H_0: \theta = \theta^*$ against simple $H_0^*: \theta = \theta''$. By NP lemma,

$$C = \{ x: \frac{L(\theta', x)}{L(\theta'', x)} \leq k \}$$

is a best critical region for testing $H_0$ against $H_0^*$.

Now

$$\frac{L(\theta', x)}{L(\theta'', x)} \leq k \Rightarrow \left(\frac{1}{\sqrt{2\pi\theta'}}\right)^n e^{-\frac{\sum X_i^2}{2\theta'}} \leq k$$

$$\Rightarrow \left(\frac{\theta''}{\theta'}\right)^{\frac{n}{2}} e^{-\frac{\theta''}{2\theta''} \sum X_i^2} \leq k$$

$$\Rightarrow \frac{1}{2} \log \left(\frac{\theta''}{\theta'}\right) - \left(\frac{\theta'' - \theta'}{2\theta''}\right) \sum X_i^2 \leq \log k$$
\[ \sum x_i = \frac{2 \theta' \theta''}{\theta'' - \theta'} \left[ \frac{n}{2} \log \left( \frac{\theta''}{\theta'} \right) - \log \frac{\theta''}{\theta'} \right] = c \]

If \( \{ \sum x_i = c \} \) is a best critical region
for testing \( H_0: \theta = \theta' \) against \( H_0: \theta = \theta'' \).

Since the foregoing argument holds for each simple
alternative \( H_0: \theta = \theta'' \) (where \( \theta'' > \theta' \)) in \( H_1 \), \( T \) is a
UMP critical region for testing \( H_0: \theta = \theta' \) against
\( H_1: \theta > \theta' \).

Determination of \( c \):

\[ \alpha = \Pr \left[ \sum x_i \geq c \mid \theta = \theta' \right] \]
\[ = \Pr \left[ \frac{\sum x_i}{\theta'} \geq \frac{c}{\theta'} \mid \theta = \theta' \right] = \Pr \left[ \chi^2_n \geq \frac{c}{\theta'} \right] \]
\[ \therefore \quad \frac{c}{\theta'} = \chi^2_n \Rightarrow c = \theta' \chi^2_n \]

**Example:** Let \( x = (x_1, \ldots, x_n) \) be a rs from \( N(\theta, 1) \)

\( H_0: \theta = \theta' \)

\( H_1: \theta > \theta' \)

\( \theta' \) is a given constant.

Thus \( \delta = \delta_1: -\infty < \delta < \infty \)

Let \( \delta'' \) be a number not equal to \( \delta' \).

Consider testing \( H_0: \theta = \theta'' \)

\( H_0: \theta = \theta'' \quad \sum (x_i - \theta'')^2 \)

\[ \frac{L(\theta' | x)}{L(\theta'', x)} \leq k \rightarrow \left( \frac{1}{2\pi} \right)^{n/2} e^{-\frac{\sum (x_i - \theta'')^2}{2}} \leq k \]
\[ -e^{-(\theta^"\prime - \theta^\prime)} \sum_{i} x_i + \frac{n}{2} \left[ \theta^"\prime - \theta^\prime \right]^2 < k \]

\[ -e^{-(\theta^"\prime - \theta^\prime)} \sum_{i} x_i \leq \log k - \frac{n}{2} \left[ \theta^"\prime - \theta^\prime \right]^2 \]

\[ (\theta^"\prime - \theta^\prime) \sum_{i} x_i \geq \frac{n}{2} \left[ (\theta^" - \theta^\prime) (\theta^" + \theta^\prime) \right] - \log k \]

Thus,

\[ \sum_{i} x_i \geq \frac{n}{2} (\theta^" + \theta^\prime) - \frac{\log k}{\theta^" - \theta^\prime} = C_1 \quad \text{if} \quad \theta^" > \theta^\prime \]

\[ \sum_{i} x_i \leq \frac{n}{2} (\theta^" + \theta^\prime) - \frac{\log k}{\theta^" - \theta^\prime} = C_2 \quad \text{if} \quad \theta^" < \theta^\prime \]

Thus \( C_1 \geq x_i \sum_{i} x_i \geq C_1 \) is a best critical region for testing \( H_0 : \theta = \theta^\prime \) against \( \overline{H}_0(+) : \theta = \theta^\prime \) where \( \theta^" > \theta^\prime \).

and \( C_2 \leq x_i \sum_{i} x_i \leq C_2 \) is a best critical region for testing \( H_0 : \theta = \theta^\prime \) against \( \overline{H}_0(-) : \theta = \theta^\prime \) where \( \theta^" < \theta^\prime \).

Testing \( H_0 : \theta = \theta^\prime \) against \( \overline{H}_0(+) \) and \( \overline{H}_0(-) \) are both simple hypotheses in \( H_1 \)

But \( \overline{H}_0(+) \) and \( \overline{H}_0(-) \) are both simple hypotheses in \( H_1 \)

Hence \( C_1 \) is the best critical region for testing \( H_0 \) against \( \overline{H}_0(+) \).

Hence \( C_1 \) is the best critical region for testing \( H_0 \) against \( \overline{H}_0(-) \).

By definition, there is no ump test for testing \( H_0 \) against \( H_1 \).

Note that had the alternative hypothesis been \( H_1 : \theta > \theta^\prime \) or \( H_1 : \theta < \theta^\prime \), a ump test would exist in each instance and they will be given by \( C_1 \) and \( C_2 \) respectively.
We know from the previous example that UMP tests may not exist. On the other hand, if they do exist, they provide the best available procedures. It is therefore natural to ask for conditions under which UMP tests exist. The answer is given by the next theorem, which we shall precede by some necessary definitions.

**Definition.** We say that the family \( \{ f(x; \theta), \theta \in \Theta \} \) of distributions has a **monotone likelihood ratio** in statistic \( Y = U(x_1, \ldots, x_n) \) if for any two values \( \theta_1, \theta_2 \in \Theta \) with \( \theta_1 < \theta_2 \), the likelihood ratio

\[
\frac{L(\theta_2, x)}{L(\theta_1, x)} = \frac{\prod_{i=1}^{n} f(x_i; \theta_2)}{\prod_{i=1}^{n} f(x_i; \theta_1)}
\]

depends on the random sample \( x=(x_1, \ldots, x_n) \) only through \( Y \) and this ratio is an increasing function of \( Y \).

**Example.** Consider \( N(\mu, \sigma^2) \) with \( \sigma^2 \) known.

Let \( \mu_1 < \mu_2 \)

\[
\frac{L(\mu_2, x)}{L(\mu_1, x)} = e^{-\frac{1}{2\sigma^2} \left[ \frac{\sum_{i=1}^{n}(x_i - \mu_2)^2}{\sigma^2} - \frac{\sum_{i=1}^{n}(x_i - \mu_1)^2}{\sigma^2} \right]}
\]

\[
= e^{-\frac{1}{2\sigma^2} \left[ -2(\mu_2 - \mu_1)\sum_{i=1}^{n} x_i + n(\mu_2^2 - \mu_1^2) \right]}
\]

\[
= e^{-\frac{1}{2\sigma^2} \left[ -2(\mu_2 - \mu_1)\sum_{i=1}^{n} x_i + n(\mu_2^2 - \mu_1^2) \right]}
\]
\[ D \exp \left( \frac{\bar{X} - \mu}{\sigma^2} \right) \sum X_i \]

where \( D > 0 \). Consequently, normal distributions for fixed \( \sigma^2 \) have monotone likelihood ratios in \( Y = \sum X_i \), or equivalently in \( \bar{X} \).

**Example:** Consider now the \( N(\mu, \sigma^2) \) distribution with known \( \mu \). Let \( \sigma^1_1 < \sigma^2_2 \). Then

\[
\begin{align*}
L(\sigma^2_2, \bar{X}) &= D \exp \left\{ \frac{1}{2\sigma^2_2} \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{1}{2\sigma^2_2} \sum_{i=1}^n (X_i - \mu)^2 \right\} \\
L(\sigma^1_1, \bar{X}) &= D \exp \left( \frac{1}{\sigma^2_1} \sum_{i=1}^n (X_i - \mu)^2 \right) \\
&= D \exp \left( \frac{1}{\sigma^2_1} \sum_{i=1}^n (1 - \frac{1}{\sigma^2_1}) (X_i - \mu)^2 \right)
\end{align*}
\]

Since \( D > 0 \) and \( \frac{1}{\sigma^2_1} - \frac{1}{\sigma^2_2} > 0 \), we see that the normal family has monotone likelihood ratios in \( Y = \sum_{i=1}^n (X_i - \mu)^2 \).

**Example:** Consider Bernoulli \( (p) \). Let \( p_1 < p_2 \).

\[
\begin{align*}
L(p_2, \bar{X}) &= \frac{\bar{X}^n (1 - p_2)^{n - \bar{X}}}{\bar{X}^n (1 - p_2)^{n - \bar{X}}} = \left[ \frac{p_2 (1 - p_1)}{p_1 (1 - p_2)} \right]^{\sum_{i=1}^n (1 - p_2)} \\
L(p_1, \bar{X}) &= \frac{\bar{X}^n (1 - p_1)^{n - \bar{X}}}{\bar{X}^n (1 - p_1)^{n - \bar{X}}} = \left[ \frac{p_1 (1 - p_2)}{p_2 (1 - p_1)} \right]^{\sum_{i=1}^n (1 - p_1)} \\
&= \left[ \frac{p_2 (1 - p_1)}{p_1 (1 - p_2)} \right]^{\sum_{i=1}^n (1 - p_2)} \\
&= \left[ \frac{p_2 (1 - p_1)}{p_1 (1 - p_2)} \right]^{\sum_{i=1}^n (1 - p_2)}
\end{align*}
\]

Since \( \frac{p_2 (1 - p_1)}{p_1 (1 - p_2)} > 1 \), the Bernoulli distributions have monotone likelihood ratio in \( Y = \sum_{i=1}^n X_i \).

It turns out that most of the known families of distributions have monotone likelihood ratio in some statistics.
The role of monotone likelihood ratios for UMP tests is explained by the following theorems.

**Theorem.** Let \( \{ f(x|\theta), \theta \in \mathbb{R} \} \) be a family indexed by one-dimensional parameter \( \theta \). Consider testing

\( \text{H}_0: \theta \leq \theta' \) versus \( \text{H}_1: \theta > \theta' \). Suppose the family has a monotone likelihood ratio in statistic \( Y \). Then for any \( c \), the test that rejects \( \text{H}_0 \) if and only if \( Y > c \) is a UMP level \( \alpha \) test where

\[ \alpha = P[ Y > c | \theta = \theta'] \]

**Theorem.** Let \( \{ f(x|\theta), \theta \in \mathbb{R} \} \) be a family indexed by one-dimensional parameter \( \theta \). Consider testing

\( \text{H}_0: \theta \geq \theta' \) versus \( \text{H}_1: \theta < \theta' \). Suppose the family has a monotone likelihood ratio in statistic \( Y \). Then for any \( c \), the test that rejects \( \text{H}_0 \) if and only if \( Y < c \) is a UMP level \( \alpha \) test where

\[ \alpha = P[ Y < c | \theta = \theta'] \]
Let \( f(x|\theta), \theta \in \mathcal{R} \) be the regular exponential family.
\[
f(x|\theta) = e^{p(\theta) k(x) + S(x) + q(\theta)} \quad x \in \mathcal{X}
\]

Let \( \theta_1 < \theta_2 \).
\[
\frac{L(\theta_2, x)}{L(\theta_1, x)} = \frac{e^{p(\theta_2) \sum k(x_i) + \sum s(x_i) + nq(\theta_2)}}{e^{p(\theta_1) \sum k(x_i) + \sum s(x_i) + nq(\theta_1)}}
\]

\[
= e^{\left[ p(\theta_2) - p(\theta_1) \right] \frac{\sum k(x_i)}{n} + nq(\theta_2) - nq(\theta_1)}
\]

Thus, if \( p(\theta) \) is an increasing function of \( \theta \), then the family has a monotone likelihood ratio in \( Y = \sum k(x_i) \).

For \( H_0: \theta \leq \theta' \),
\[ H_1: \theta > \theta' \]

The UMP test is: Reject \( H_0 \) if \( Y = \sum k(x_i) > c \)
where \( \alpha = P[Y > c| \theta = \theta'] \)

For \( H_0: \theta \geq \theta' \),
\[ H_1: \theta < \theta' \]

The UMP test is: Reject \( H_0 \) if \( Y = \sum k(x_i) < c \)
where \( \alpha = P[Y < c| \theta = \theta'] \)