Last Lecture

• What is a Uniformly Most Powerful (UMP) Test?
• Does UMP level test always exist for simple hypothesis testing?
• For composite hypothesis, which property makes it possible to construct a UMP level test?
• What is a sufficient condition for an exponential family to have MLR property?
• For one-sided composite hypothesis testing, if a sufficient statistic satisfies MLR property, how can a UMP level test be constructed?
Last Lecture

- What is a Uniformly Most Powerful (UMP) Test?
Last Lecture

- What is a Uniformly Most Powerful (UMP) Test?
- Does UMP level \( \alpha \) test always exist for simple hypothesis testing?
What is a Uniformly Most Powerful (UMP) Test?
Does UMP level $\alpha$ test always exist for simple hypothesis testing?
For composite hypothesis, which property makes it possible to construct a UMP level $\alpha$ test?
Last Lecture

- What is a Uniformly Most Powerful (UMP) Test?
- Does UMP level $\alpha$ test always exist for simple hypothesis testing?
- For composite hypothesis, which property makes it possible to construct a UMP level $\alpha$ test?
- What is a sufficient condition for an exponential family to have MLR property?
Last Lecture

- What is a Uniformly Most Powerful (UMP) Test?
- Does UMP level $\alpha$ test always exist for simple hypothesis testing?
- For composite hypothesis, which property makes it possible to construct a UMP level $\alpha$ test?
- What is a sufficient condition for an exponential family to have MLR property?
- For one-sided composite hypothesis testing, if a sufficient statistic satisfies MLR property, how can a UMP level $\alpha$ test be constructed?
Uniformly Most Powerful Test (UMP)

**Definition**

Let $C$ be a class of tests between $H_0 : \theta \in \Omega$ vs $H_1 : \theta \in \Omega_0^c$. A test in $C$, with power function $\beta(\theta)$ is **uniformly most powerful (UMP) test** in class $C$ if $\beta(\theta) \geq \beta'(\theta)$ for every $\theta \in \Omega_0^c$ and every $\beta'(\theta)$, which is a power function of another test in $C$.

**UMP level $\alpha$ test**

Consider $C$ be the set of all the level $\alpha$ test. The UMP test in this class is called a UMP level $\alpha$ test.
Uniformly Most Powerful Test (UMP)

Definition

Let $C$ be a class of tests between $H_0 : \theta \in \Omega$ vs $H_1 : \theta \in \Omega^c_0$. A test in $C$, with power function $\beta(\theta)$ is uniformly most powerful (UMP) test in class $C$ if $\beta(\theta) \geq \beta'(\theta)$ for every $\theta \in \Omega^c_0$ and every $\beta'(\theta)$, which is a power function of another test in $C$.

UMP level $\alpha$ test

Consider $C$ be the set of all the level $\alpha$ test. The UMP test in this class is called a UMP level $\alpha$ test.

UMP level $\alpha$ test has the smallest type II error probability for any $\theta \in \Omega^c_0$ in this class.
Uniformly Most Powerful Test (UMP)

**Definition**

Let $C$ be a class of tests between $H_0 : \theta \in \Omega$ vs $H_1 : \theta \in \Omega_c^c$. A test in $C$, with power function $\beta(\theta)$ is *uniformly most powerful (UMP) test* in class $C$ if $\beta(\theta) \geq \beta'(\theta)$ for every $\theta \in \Omega_0^c$ and every $\beta'(\theta)$, which is a power function of another test in $C$.

**UMP level $\alpha$ test**

Consider $C$ be the set of all the level $\alpha$ test. The UMP test in this class is called a UMP level $\alpha$ test.

UMP level $\alpha$ test has the smallest type II error probability for any $\theta \in \Omega_0^c$ in this class.
Neyman-Pearson Lemma

Theorem 8.3.12 - Neyman-Pearson Lemma

Consider testing \( H_0 : \theta = \theta_0 \) vs. \( H_1 : \theta = \theta_1 \) where the pdf or pmf corresponding the \( \theta_i \) is \( f(x|\theta_i) \), \( i = 0, 1 \), using a test with rejection region \( R \) that satisfies
Neyman-Pearson Lemma

Theorem 8.3.12 - Neyman-Pearson Lemma

Consider testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ where the pdf or pmf corresponding the $\theta_i$ is $f(x|\theta_i)$, $i = 0, 1$, using a test with rejection region $R$ that satisfies

$$x \in R \quad \text{if} \quad f(x|\theta_1) > kf(x|\theta_0) \quad (8.3.1) \quad \text{and}$$
Neyman-Pearson Lemma

**Theorem 8.3.12 - Neyman-Pearson Lemma**

Consider testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ where the pdf or pmf corresponding the $\theta_i$ is $f(x|\theta_i)$, $i = 0, 1$, using a test with rejection region $R$ that satisfies

\[
\begin{align*}
\mathbf{x} \in R & \quad \text{if } f(x|\theta_1) > kf(x|\theta_0) \quad (8.3.1) \text{ and} \\
\mathbf{x} \in R^c & \quad \text{if } f(x|\theta_1) < kf(x|\theta_0) \quad (8.3.2)
\end{align*}
\]
Neyman-Pearson Lemma

Theorem 8.3.12 - Neyman-Pearson Lemma

Consider testing \( H_0 : \theta = \theta_0 \) vs. \( H_1 : \theta = \theta_1 \) where the pdf or pmf corresponding the \( \theta_i \) is \( f(x|\theta_i) \), \( i = 0, 1 \), using a test with rejection region \( R \) that satisfies

\[
\begin{align*}
x & \in R & \text{if} & \quad f(x|\theta_1) > kf(x|\theta_0) & \quad (8.3.1) \\
x & \in R^c & \text{if} & \quad f(x|\theta_1) < kf(x|\theta_0) & \quad (8.3.2)
\end{align*}
\]

For some \( k \geq 0 \) and \( \alpha = \Pr(X \in R|\theta_0) \), Then,
Neyman-Pearson Lemma

**Theorem 8.3.12 - Neyman-Pearson Lemma**

Consider testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ where the pdf or pmf corresponding the $\theta_i$ is $f(x|\theta_i)$, $i = 0, 1$, using a test with rejection region $R$ that satisfies

\[ x \in R \quad \text{if} \quad f(x|\theta_1) > kf(x|\theta_0) \quad (8.3.1) \text{ and } \]

\[ x \in R^c \quad \text{if} \quad f(x|\theta_1) < kf(x|\theta_0) \quad (8.3.2) \]

For some $k \geq 0$ and $\alpha = \Pr(X \in R|\theta_0)$, Then,

- (Sufficiency) Any test that satisfies 8.3.1 and 8.3.2 is a UMP level $\alpha$ test
Theorem 8.3.12 - Neyman-Pearson Lemma

Consider testing $H_0: \theta = \theta_0$ vs. $H_1: \theta = \theta_1$ where the pdf or pmf corresponding the $\theta_i$ is $f(x|\theta_i)$, $i = 0, 1$, using a test with rejection region $R$ that satisfies

\[ x \in R \quad \text{if} \quad f(x|\theta_1) > kf(x|\theta_0) \quad (8.3.1) \text{ and } \]
\[ x \in R^c \quad \text{if} \quad f(x|\theta_1) < kf(x|\theta_0) \quad (8.3.2) \]

For some $k \geq 0$ and $\alpha = \Pr(X \in R|\theta_0)$, Then,

- (Sufficiency) Any test that satisfies 8.3.1 and 8.3.2 is a UMP level $\alpha$ test

- (Necessity) if there exist a test satisfying 8.3.1 and 8.3.2 with $k > 0$, then every UMP level $\alpha$ test is a size $\alpha$ test (satisfies 8.3.2), and every UMP level $\alpha$ test satisfies 8.3.1 except perhaps on a set $A$ satisfying $\Pr(X \in A|\theta_0) = \Pr(X \in A|\theta_1) = 0$. 

Neyman-Pearson Lemma
Monotone Likelihood Ratio

**Definition**

A family of pdfs or pmfs \( \{g(t|\theta) : \theta \in \Omega\} \) for a univariate random variable \( T \) with real-valued parameter \( \theta \) have a monotone likelihood ratio if \( \frac{g(t|\theta_2)}{g(t|\theta_1)} \) is an increasing (or non-decreasing) function of \( t \) for every \( \theta_2 > \theta_1 \) on \( \{ t : g(t|\theta_1) > 0 \, \text{or} \, g(t|\theta_2) > 0 \} \).
Monotone Likelihood Ratio

Definition

A family of pdfs or pmfs \( \{g(t|\theta) : \theta \in \Omega\} \) for a univariate random variable \( T \) with real-valued parameter \( \theta \) have a monotone likelihood ratio if

\[
\frac{g(t|\theta_2)}{g(t|\theta_1)}
\]

is an increasing (or non-decreasing) function of \( t \) for every \( \theta_2 > \theta_1 \) on \( \{ t : g(t|\theta_1) > 0 \text{ or } g(t|\theta_2) > 0 \} \).

Note: we may define MLR using decreasing function of \( t \). But all following theorems are stated according to the definition.
Karlin-Rabin Theorem

Theorem 8.1.17
Suppose $T(X)$ is a sufficient statistic for $\theta$ and the family \{\(g(t|\theta) : \theta \in \Omega\)\} is an MLR family. Then
Karlin-Rabin Theorem

**Theorem 8.1.17**

Suppose $T(X)$ is a sufficient statistic for $\theta$ and the family $\{g(t|\theta) : \theta \in \Omega\}$ is an MLR family. Then

1. For testing $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$, the UMP level $\alpha$ test is given by rejecting $H_0$ if and only if $T > t_0$ where $\alpha = \Pr(T > t_0|\theta_0)$. 

Hyun Min Kang  
Biostatistics 602 - Lecture 21  
April 2nd, 2013  
6 / 25
Karlin-Rabin Theorem

Theorem 8.1.17

Suppose $T(X)$ is a sufficient statistic for $\theta$ and the family $\{g(t|\theta) : \theta \in \Omega\}$ is an MLR family. Then

1. For testing $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$, the UMP level $\alpha$ test is given by rejecting $H_0$ if and only if $T > t_0$ where $\alpha = \Pr(T > t_0|\theta_0)$.

2. For testing $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$, the UMP level $\alpha$ test is given by rejecting $H_0$ if and only if $T < t_0$ where $\alpha = \Pr(T < t_0|\theta_0)$. 
Normal Example with Known Mean

\[ X_i \overset{i.i.d.}{\sim} \mathcal{N}(\mu_0, \sigma^2) \] where \( \sigma^2 \) is unknown and \( \mu_0 \) is known. Find the UMP level \( \alpha \) test for testing \( H_0 : \sigma^2 \leq \sigma_0^2 \) vs. \( H_1 : \sigma^2 > \sigma_0^2 \). Let \( T = \sum_{i=1}^{n} (X_i - \mu_0)^2 \) is sufficient for \( \sigma^2 \).
Normal Example with Known Mean

$X_i \sim \text{i.i.d. } \mathcal{N}(\mu_0, \sigma^2)$ where $\sigma^2$ is unknown and $\mu_0$ is known. Find the UMP level $\alpha$ test for testing $H_0 : \sigma^2 \leq \sigma_0^2$ vs. $H_1 : \sigma^2 > \sigma_0^2$. Let $T = \sum_{i=1}^{n} (X_i - \mu_0)^2$ is sufficient for $\sigma^2$. To check whether $T$ has MLR property, we need to find $g(t|\sigma^2)$. 

\[ Y = \frac{T}{\sigma_0^2} = \frac{\sum_{i=1}^{n} (X_i - \mu_0)^2}{\sigma_0^2} \]

\[ f_Y(y) = \frac{1}{(n-2)\sigma_0^2 y^{n/2-1}} e^{-y/n} \]
Normal Example with Known Mean

\( X_i \overset{i.i.d.}{\sim} \mathcal{N}(\mu_0, \sigma^2) \) where \( \sigma^2 \) is unknown and \( \mu_0 \) is known. Find the UMP level \( \alpha \) test for testing \( H_0 : \sigma^2 \leq \sigma_0^2 \) vs. \( H_1 : \sigma^2 > \sigma_0^2 \). Let \( T = \sum_{i=1}^{n} (X_i - \mu_0)^2 \) is sufficient for \( \sigma^2 \). To check whether \( T \) has MLR property, we need to find \( g(t|\sigma^2) \).

\[
\frac{X_i - \mu_0}{\sigma} \sim \mathcal{N}(0, 1)
\]
Normal Example with Known Mean

\(X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_0, \sigma^2)\) where \(\sigma^2\) is unknown and \(\mu_0\) is known. Find the UMP level \(\alpha\) test for testing \(H_0 : \sigma^2 \leq \sigma_0^2\) vs. \(H_1 : \sigma^2 > \sigma_0^2\). Let
\(T = \sum_{i=1}^{n} (X_i - \mu_0)^2\) is sufficient for \(\sigma^2\). To check whether \(T\) has MLR property, we need to find \(g(t|\sigma^2)\).

\[
\frac{X_i - \mu_0}{\sigma} \sim \mathcal{N}(0, 1)
\]

\[
\left(\frac{X_i - \mu_0}{\sigma}\right)^2 \sim \chi_1^2
\]
Normal Example with Known Mean

\( X_i \sim i.i.d. \mathcal{N}(\mu_0, \sigma^2) \) where \( \sigma^2 \) is unknown and \( \mu_0 \) is known. Find the UMP level \( \alpha \) test for testing \( H_0 : \sigma^2 \leq \sigma_0^2 \) vs. \( H_1 : \sigma^2 > \sigma_0^2 \). Let \( T = \sum_{i=1}^{n} (X_i - \mu_0)^2 \) is sufficient for \( \sigma^2 \). To check whether \( T \) has MLR property, we need to find \( g(t|\sigma^2) \).

\[
\frac{X_i - \mu_0}{\sigma} \sim \mathcal{N}(0, 1)
\]

\[
\left( \frac{X_i - \mu_0}{\sigma} \right)^2 \sim \chi_1^2
\]

\[
Y = \frac{T}{\sigma^2} = \sum_{i=1}^{n} \left( \frac{X_i - \mu_0}{\sigma} \right)^2 \sim \chi_n^2
\]
Normal Example with Known Mean

Let $X_i \sim \mathcal{N}(\mu_0, \sigma^2)$ where $\sigma^2$ is unknown and $\mu_0$ is known. Find the UMP level $\alpha$ test for testing $H_0 : \sigma^2 \leq \sigma_0^2$ vs. $H_1 : \sigma^2 > \sigma_0^2$. Let $T = \sum_{i=1}^{n} (X_i - \mu_0)^2$ is sufficient for $\sigma^2$. To check whether $T$ has MLR property, we need to find $g(t|\sigma^2)$.

\[
\frac{X_i - \mu_0}{\sigma} \sim \mathcal{N}(0, 1)
\]
\[
\left( \frac{X_i - \mu_0}{\sigma} \right)^2 \sim \chi_1^2
\]

\[
Y = \frac{T}{\sigma^2} = \sum_{i=1}^{n} \left( \frac{X_i - \mu_0}{\sigma} \right)^2 \sim \chi_n^2
\]

\[
f_Y(y) = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} y^{n/2 - 1} e^{-y/2}
\]
Normal Example with Known Mean (cont’d)

\[ f_T(t) = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2} \left( \frac{t}{\sigma^2} \right)^{n/2-1}} e^{-\frac{t}{2\sigma^2}} \left| \frac{dy}{dt} \right| dt \]
Normal Example with Known Mean (cont’d)

\[ f_T(t) = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{t}{\sigma^2} \right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \left| \frac{dy}{dt} \right| dt \]

\[ = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{t}{\sigma^2} \right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \frac{1}{\sigma^2} dt \]
Normal Example with Known Mean (cont’d)

\[
f_T(t) = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{t}{\sigma^2} \right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \left| \frac{dy}{dt} \right| dt
\]

\[
= \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{t}{\sigma^2} \right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \frac{1}{\sigma^2} dt
\]

\[
= \frac{t^{n/2-1}}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{1}{\sigma^2} \right)^{n/2} e^{-\frac{t}{2\sigma^2}} dt
\]
Normal Example with Known Mean (cont’d)

\[ f_T(t) = \frac{1}{\Gamma \left(\frac{n}{2}\right) 2^{n/2}} \left(\frac{t}{\sigma^2}\right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \left| \frac{dy}{dt} \right| dt \]

\[ = \frac{1}{\Gamma \left(\frac{n}{2}\right) 2^{n/2}} \left(\frac{t}{\sigma^2}\right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \frac{1}{\sigma^2} dt \]

\[ = \frac{t^{n/2-1}}{\Gamma \left(\frac{n}{2}\right) 2^{n/2}} \left(\frac{1}{\sigma^2}\right)^{n/2} e^{-\frac{t}{2\sigma^2}} dt \]

\[ = h(t) c(\sigma^2) \exp[w(\sigma^2) t] \]

where \( w(\sigma^2) = -\frac{1}{2\sigma^2} \) is an increasing function in \( \sigma^2 \).
Normal Example with Known Mean (cont’d)

\[ f_T(t) = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{t}{\sigma^2} \right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \left| \frac{dy}{dt} \right| dt \]

\[ = \frac{1}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{t}{\sigma^2} \right)^{n/2-1} e^{-\frac{t}{2\sigma^2}} \frac{1}{\sigma^2} dt \]

\[ = \frac{t^{n/2-1}}{\Gamma \left( \frac{n}{2} \right) 2^{n/2}} \left( \frac{1}{\sigma^2} \right)^{n/2} e^{-\frac{t}{2\sigma^2}} dt \]

\[ = h(t) c(\sigma^2) \exp[w(\sigma^2)t] \]

where \( w(\sigma^2) = -\frac{1}{2\sigma^2} \) is an increasing function in \( \sigma^2 \). Therefore, \( T = \sum_{i=1}^{n} (X_i - \mu)^2 \) has the MLR property.
Normal Example with Known Mean (cont’d)

By Karlin-Rabin Theorem, UMP level $\alpha$ rejects $s H_0$ if and only if $T > t_0$ where $t_0$ is chosen such that $\alpha = \Pr(T > t_0|\sigma^2_0)$. 
Normal Example with Known Mean (cont’d)

By Karlin-Rabin Theorem, UMP level $\alpha$ rejects $H_0$ if and only if $T > t_0$ where $t_0$ is chosen such that $\alpha = \Pr(T > t_0 | \sigma_0^2)$.

Note that $\frac{T}{\sigma^2} \sim \chi_n^2$

$$\Pr(T > t_0 | \sigma_0^2) = \Pr \left( \frac{T}{\sigma_0^2} > \frac{t_0}{\sigma_0^2} \bigg| \sigma_0^2 \right)$$
Normal Example with Known Mean (cont’d)

By Karlin-Rabin Theorem, UMP level $\alpha$ rejects $H_0$ if and only if $T > t_0$ where $t_0$ is chosen such that $\alpha = \Pr(T > t_0|\sigma_0^2)$.

Note that $\frac{T}{\sigma^2} \sim \chi_n^2$

$$\Pr(T > t_0|\sigma_0^2) = \Pr\left(\frac{T}{\sigma_0^2} > \frac{t_0}{\sigma_0^2} \bigg| \sigma_0^2\right)$$

$$\frac{T}{\sigma_0^2} \sim \chi_n^2$$
Normal Example with Known Mean (cont’d)

By Karlin-Rabin Theorem, UMP level $\alpha$ rejects $s H_0$ if and only if $T > t_0$ where $t_0$ is chosen such that $\alpha = \Pr(T > t_0 | \sigma_0^2)$.

Note that $\frac{T}{\sigma^2} \sim \chi^2_n$

\[
\Pr(T > t_0 | \sigma_0^2) = \Pr\left(\frac{T}{\sigma_0^2} > \frac{t_0}{\sigma_0^2} \bigg| \sigma_0^2\right)
\]

\[
\frac{T}{\sigma_0^2} \sim \chi^2_n
\]

\[
\Pr\left(\chi^2_n > \frac{t_0}{\sigma_0^2}\right) = \alpha
\]
Normal Example with Known Mean (cont’d)

By Karlin-Rabin Theorem, UMP level $\alpha$ rejects $s H_0$ if and only if $T > t_0$ where $t_0$ is chosen such that $\alpha = \Pr(T > t_0 | \sigma_0^2)$.

Note that $\frac{T}{\sigma^2} \sim \chi^2_n$

\[
\Pr(T > t_0 | \sigma_0^2) = \Pr \left( \frac{T}{\sigma_0^2} > \frac{t_0}{\sigma_0^2} \middle| \sigma_0^2 \right)
\]

\[
\frac{T}{\sigma_0^2} \sim \chi^2_n
\]

\[
\Pr \left( \chi_n^2 > \frac{t_0}{\sigma_0^2} \right) = \alpha
\]

\[
\frac{t_0}{\sigma_0^2} = \chi_{n, \alpha}^2
\]
Normal Example with Known Mean (cont’d)

By Karlin-Rabin Theorem, UMP level \( \alpha \) rejects \( s H_0 \) if and only if \( T > t_0 \) where \( t_0 \) is chosen such that \( \alpha = \Pr(T > t_0|\sigma_0^2) \).

Note that \( \frac{T}{\sigma^2} \sim \chi_n^2 \)

\[
\Pr(T > t_0|\sigma_0^2) = \Pr\left(\frac{T}{\sigma_0^2} > \frac{t_0}{\sigma_0^2} \bigg| \sigma_0^2 \right)
\]

\[
\frac{T}{\sigma_0^2} \sim \chi_n^2
\]

\[
\Pr\left(\frac{\chi_n^2}{\sigma_0^2} > \frac{t_0}{\sigma_0^2}\right) = \alpha
\]

\[
\frac{t_0}{\sigma_0^2} = \chi_{n,\alpha}^2
\]

\[
t_0 = \sigma_0^2 \chi_{n,\alpha}^2
\]

where \( \chi_{n,\alpha}^2 \) satisfies \( \int_{\chi_{n,\alpha}^2}^{\infty} f_{\chi_n^2}(x) \, dx = \alpha \).
Remarks

- For many problems, UMP level $\alpha$ test does not exist (Example 8.3.19).
Remarks

- For many problems, UMP level $\alpha$ test does not exist (Example 8.3.19).
- In such cases, we can restrict our search among a subset of tests, for example, all unbiased tests.
Distribution of LRT

\[ \lambda(x) = \frac{\sup_{\Omega_0} L(\theta|x)}{\sup_\Omega L(\theta|x)} \]
Distribution of LRT

\[ \lambda(x) = \frac{\sup_{\Omega_0} L(\theta|x)}{\sup_{\Omega} L(\theta|x)} \]

LRT level \( \alpha \) test procedure rejects \( H_0 \) if and only if \( \lambda(x) \leq c \). \( c \) is chosen such that
Distribution of LRT

\[ \lambda(x) = \frac{\sup_{\Omega_0} L(\theta|x)}{\sup_{\Omega} L(\theta|x)} \]

LRT level \( \alpha \) test procedure rejects \( H_0 \) if and only if \( \lambda(x) \leq c \). \( c \) is chosen such that

\[ \sup_{\theta \in \Omega_0} \Pr(\lambda(x) \leq c) \leq \alpha \]
Distribution of LRT

\[ \lambda(x) = \frac{\sup_{\Omega_0} L(\theta|x)}{\sup_{\Omega} L(\theta|x)} \]

LRT level \( \alpha \) test procedure rejects \( H_0 \) if and only if \( \lambda(x) \leq c \). \( c \) is chosen such that

\[ \sup_{\theta \in \Omega_0} \Pr(\lambda(x) \leq c) \leq \alpha \]

Usually, it is difficult to derive the distribution of \( \lambda(x) \) and to solve the equation of \( c \).
Asymptotics of LRT

Theorem 10.3.1
Consider testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$. Suppose $X_1, \cdots, X_n$ are iid samples from $f(x|\theta)$, and $\hat{\theta}$ is the MLE of $\theta$, and $f(x|\theta)$ satisfies certain "regularity conditions" (e.g. see misc 10.6.2), then under $H_0$: 

$$2 \log \left(\frac{f(x|\theta)}{f(x|\hat{\theta})}\right) \xrightarrow{d} \chi^2_1 as n \to \infty.$$
Asymptotics of LRT

Theorem 10.3.1

Consider testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$. Suppose $X_1, \cdots, X_n$ are iid samples from $f(x|\theta)$, and $\hat{\theta}$ is the MLE of $\theta$, and $f(x|\theta)$ satisfies certain "regularity conditions" (e.g. see misc 10.6.2), then under $H_0$:

$$-2 \log \lambda(x) \xrightarrow{d} \chi^2_1$$

as $n \rightarrow \infty$. 
Proof

\[ \lambda(x) = \frac{\sup_{\theta \in \Omega_0} L(\theta|x)}{\sup_{\theta \in \Omega} L(\theta|x)} = \frac{L(\theta_0|x)}{L(\hat{\theta}|x)} \]
Proof

\[
\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Omega_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Omega} L(\theta|\mathbf{x})} = \frac{L(\theta_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})}
\]

\[
-2\lambda(\mathbf{x}) = -2 \log \left( \frac{L(\theta_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})} \right)
\]
Proof

\[ \lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Omega_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Omega} L(\theta|\mathbf{x})} = \frac{L(\theta_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})} \]

\[ -2\lambda(\mathbf{x}) = -2 \log \left( \frac{L(\theta_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})} \right) \]

\[ = -2 \log L(\theta_0|\mathbf{x}) + 2 \log L(\hat{\theta}|\mathbf{x}) \]
Proof

\[ \lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Omega_0} L(\theta | \mathbf{x})}{\sup_{\theta \in \Omega} L(\theta | \mathbf{x})} = \frac{L(\theta_0 | \mathbf{x})}{L(\hat{\theta} | \mathbf{x})} \]

\[ -2\lambda(\mathbf{x}) = -2 \log \left( \frac{L(\theta_0 | \mathbf{x})}{L(\hat{\theta} | \mathbf{x})} \right) \]

\[ = -2 \log L(\theta_0 | \mathbf{x}) + 2 \log L(\hat{\theta} | \mathbf{x}) \]

\[ = -2l(\theta_0 | \mathbf{x}) + 2l(\hat{\theta} | \mathbf{x}) \]
Proof (cont’d)

Expanding $l(\theta|x)$ around $\hat{\theta}$,
Proof (cont’d)

Expanding $l(\theta|x)$ around $\hat{\theta}$,

$$l(\theta|x) = l(\hat{\theta}|x) + l'(\hat{\theta}|x)(\theta - \hat{\theta}) + \frac{l''(\hat{\theta}|x)(\theta - \hat{\theta})^2}{2} + \ldots$$
Proof (cont’d)

Expanding $l(\theta|x)$ around $\hat{\theta}$,

$$l(\theta|x) = l(\hat{\theta}|x) + l'(\hat{\theta}|x)(\theta - \hat{\theta}) + l''(\hat{\theta}|x)\frac{(\theta - \hat{\theta})^2}{2} + \cdots$$

$$l'(\hat{\theta}|x) = 0 \quad \text{(assuming regularity condition)}$$
Proof (cont’d)

Expanding \( l(\theta|\mathbf{x}) \) around \( \hat{\theta} \),

\[
    l(\theta|\mathbf{x}) = l(\hat{\theta}|\mathbf{x}) + l'(\hat{\theta}|\mathbf{x})(\theta - \hat{\theta}) + l''(\hat{\theta}|\mathbf{x}) \frac{(\theta - \hat{\theta})^2}{2} + \cdots \\
    l'(\hat{\theta}|\mathbf{x}) = 0 \quad \text{(assuming regularity condition)} \\
    l(\theta_0|\mathbf{x}) \approx l(\hat{\theta}|\mathbf{x}) + l''(\hat{\theta}|\mathbf{x}) \frac{(\theta_0 - \hat{\theta})^2}{2}
\]
Proof (cont’d)

Expanding $l(\theta|x)$ around $\hat{\theta}$,

\[
l(\theta|x) = l(\hat{\theta}|x) + l'(\hat{\theta}|x)(\theta - \hat{\theta}) + l''(\hat{\theta}|x)\frac{(\theta - \hat{\theta})^2}{2} + \cdots
\]

\[
l'(\hat{\theta}|x) = 0 \quad \text{(assuming regularity condition)}
\]

\[
l(\theta_0|x) \approx l(\hat{\theta}|x) + l''(\hat{\theta}|x)\frac{(\theta_0 - \hat{\theta})^2}{2}
\]

\[
-2 \log \lambda(x) = -2l(\theta_0|x) + 2l(\hat{\theta}|x)
\]
Proof (cont’d)

Expanding $l(\theta|x)$ around $\hat{\theta}$,

$$l(\theta|x) = l(\hat{\theta}|x) + l'(\hat{\theta}|x)(\theta - \hat{\theta}) + l''(\hat{\theta}|x)(\theta - \hat{\theta})^2 + \cdots$$

$$l'(\hat{\theta}|x) = 0 \quad \text{(assuming regularity condition)}$$

$$l(\theta_0|x) \approx l(\hat{\theta}|x) + l''(\hat{\theta}|x)\frac{(\theta_0 - \hat{\theta})^2}{2} - 2\log \lambda(x) = -2l(\theta_0|x) + 2l(\hat{\theta}|x)$$

$$\approx -(\theta_0 - \hat{\theta})^2 l''(\hat{\theta}|x)$$
Proof (cont’d)

Because $\hat{\theta}$ is MLE, under $H_0$,

$$\hat{\theta} \sim \mathcal{N}(\theta_0, \frac{1}{I_n(\theta_0)})$$
Proof (cont’d)

Because \( \hat{\theta} \) is MLE, under \( H_0 \),

\[
\hat{\theta} \sim \mathcal{AN} \left( \theta_0, \frac{1}{I_n(\theta_0)} \right)
\]

\[
(\hat{\theta} - \theta_0) \sqrt{I_n(\theta_0)} \xrightarrow{d} \mathcal{N}(0, 1)
\]
Proof (cont’d)

Because \( \hat{\theta} \) is MLE, under \( H_0 \),

\[
\hat{\theta} \sim \mathcal{N} \left( \theta_0, \frac{1}{I_n(\theta_0)} \right)
\]

\[
(\hat{\theta} - \theta_0) \sqrt{I_n(\theta_0)} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\[
(\hat{\theta} - \theta_0)^2 I_n(\theta_0) \xrightarrow{d} \chi^2_1
\]
Proof (cont’d)

Because $\hat{\theta}$ is MLE, under $H_0,$

\[
\hat{\theta} \sim \mathcal{N}(\theta_0, \frac{1}{I_n(\theta_0)})
\]

\[
(\hat{\theta} - \theta_0) \sqrt{I_n(\theta_0)} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\[
(\hat{\theta} - \theta_0)^2 I_n(\theta_0) \xrightarrow{d} \chi_1^2
\]

Therefore,

\[
-2 \log \lambda(x) \approx -(\theta_0 - \hat{\theta})^2 l''(\hat{\theta} | x)
\]
Proof (cont’d)

Because \( \hat{\theta} \) is MLE, under \( H_0 \),

\[
\hat{\theta} \sim \mathcal{N} \left( \theta_0, \frac{1}{I_n(\theta_0)} \right)
\]

\[
(\hat{\theta} - \theta_0) \sqrt{I_n(\theta_0)} \xrightarrow{d} \mathcal{N}(0, 1)
\]

\[
(\hat{\theta} - \theta_0)^2 I_n(\theta_0) \xrightarrow{d} \chi^2_1
\]

Therefore,

\[
-2 \log \lambda(\mathbf{x}) \approx -(\theta_0 - \hat{\theta})^2 l''(\hat{\theta}|\mathbf{x})
\]

\[
= (\hat{\theta} - \theta_0)^2 I_n(\theta_0) \frac{-\frac{1}{n} l''(\hat{\theta}|\mathbf{x})}{\frac{1}{n} I_n(\theta_0)}
\]
Proof (cont’d)

\[- \frac{1}{n} l''(\hat{\theta}|x) = - \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} f(x_i|\theta) \bigg|_{\theta=\hat{\theta}} \]
Proof (cont’d)

\[- \frac{1}{n} l''(\hat{\theta} | x) = - \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} f(x_i | \theta) \bigg|_{\theta = \hat{\theta}} \]

\[\xrightarrow{P} - E \left( \frac{\partial^2}{\partial \theta^2} f(x | \theta) \right) \bigg|_{\theta = \theta_0} = I(\theta_0)\]
Proof (cont’d)

\[- \frac{1}{n} l''(\hat{\theta}|x) = - \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} f(x_i|\theta) \bigg|_{\theta=\hat{\theta}} \]

\[\xrightarrow{P} - E \left( \frac{\partial^2}{\partial \theta^2} f(x|\theta) \right) \bigg|_{\theta=\theta_0} = I(\theta_0)\]

\[- \frac{1}{n} l''(\hat{\theta}|x) = - \frac{1}{n} l''(\hat{\theta}|x) \xrightarrow{P} 1\]
Proof (cont’d)

\[-\frac{1}{n} l''(\hat{\theta}|x) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} f(x_i|\theta) \bigg|_{\theta = \hat{\theta}} \]

\[\xrightarrow{P} -E \left( \frac{\partial^2}{\partial \theta^2} f(x|\theta) \right) \bigg|_{\theta = \theta_0} = I(\theta_0)\]

\[-\frac{1}{n} l''(\hat{\theta}|x) = -\frac{1}{n} l''(\hat{\theta}|x) \xrightarrow{P} 1\]

\[\frac{1}{n} I_n(\theta_0) = \frac{1}{n} I(\theta_0)\]

By Slutsky’s Theorem, under \(H_0\)

\[-(\hat{\theta} - \theta_0)^2 l''(\hat{\theta}|x) \xrightarrow{d} \chi^2_1\]
Proof (cont’d)

\[-\frac{1}{n} l''(\hat{\theta}|\mathbf{x}) \quad = \quad -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} f(x_i|\theta) \bigg|_{\theta=\hat{\theta}} \]

\[\xrightarrow{P} -E \left( \frac{\partial^2}{\partial \theta^2} f(x|\theta) \right) \bigg|_{\theta=\theta_0} = I(\theta_0)\]

\[-\frac{1}{n} l''(\hat{\theta}|\mathbf{x}) = -\frac{1}{n} l''(\hat{\theta}|\mathbf{x}) \quad \xrightarrow{P} \quad 1\]

\[\frac{1}{n} I_n(\theta_0) = \frac{1}{I(\theta_0)} \]

By Slutsky’s Theorem, under $H_0$

\[-(\hat{\theta} - \theta_0)^2 l''(\hat{\theta}|\mathbf{X}) \xrightarrow{d} \chi^2_1\]

\[-2 \log \lambda(\mathbf{X}) \xrightarrow{d} \chi^2_1\]
Example

$X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. Consider testing $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$. 
Example

$X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. Consider testing $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$. Using LRT,

$$
\lambda(x) = \frac{L(\lambda_0|x)}{\sup_{\lambda} L(\lambda|x)}
$$
Example

$X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. Consider testing $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$. Using LRT,

$$\lambda(x) = \frac{L(\lambda_0|x)}{\sup_{\lambda} L(\lambda|x)}$$

MLE of $\lambda$ is $\hat{\lambda} = \bar{X} = \frac{1}{n} \sum X_i$. 

---

Recap

Karlin-Rabin

Asymptotics of LRT

Wald Test

Summary
Example

\( X_i \text{ i.i.d. Poisson}(\lambda) \). Consider testing \( H_0 : \lambda = \lambda_0 \) vs \( H_1 : \lambda \neq \lambda_0 \). Using LRT,

\[
\lambda(x) = \frac{L(\lambda_0 | x)}{\sup_\lambda L(\lambda | x)}
\]

MLE of \( \lambda \) is \( \hat{\lambda} = \bar{X} = \frac{1}{n} \sum X_i \).

\[
\lambda(x) = \frac{\prod_{i=1}^{n} \frac{e^{-\lambda_0} \lambda_0^{x_i}}{x_i!}}{\prod_{i=1}^{n} \frac{e^{-\bar{x}} \bar{x}^{x_i}}{x_i!}}
\]
Example

\( X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda) \). Consider testing \( H_0 : \lambda = \lambda_0 \) vs \( H_1 : \lambda \neq \lambda_0 \). Using LRT,

\[
\lambda(x) = \frac{L(\lambda_0|x)}{\sup_{\lambda} L(\lambda|x)}
\]

MLE of \( \lambda \) is \( \hat{\lambda} = \bar{X} = \frac{1}{n} \sum X_i \).

\[
\lambda(x) = \frac{\prod_{i=1}^{n} \frac{e^{-\lambda_0 \lambda_0^{x_i}}}{x_i!}}{\prod_{i=1}^{n} \frac{e^{-\bar{x} \lambda_0^{x_i}}}{x_i!}} = \frac{e^{-n \lambda_0} \lambda_0^{\sum x_i}}{e^{-n \bar{x} \lambda_0} \sum x_i}
\]
Example

\( X_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda) \). Consider testing \( H_0 : \lambda = \lambda_0 \) vs \( H_1 : \lambda \neq \lambda_0 \). Using LRT,

\[
\lambda(\mathbf{x}) = \frac{L(\lambda_0|\mathbf{x})}{\sup_{\lambda} L(\lambda|\mathbf{x})}
\]

MLE of \( \lambda \) is \( \hat{\lambda} = \bar{X} = \frac{1}{n} \sum X_i \).

\[
\lambda(\mathbf{x}) = \frac{\prod_{i=1}^{n} \frac{e^{-\lambda_0 \lambda_i x_i}}{x_i!}}{\prod_{i=1}^{n} \frac{e^{-\bar{X} \lambda_i x_i}}{x_i!}} = \frac{e^{-n\lambda_0 \sum x_i}}{e^{-n\bar{X} \sum x_i}} = e^{-n(\lambda_0 - \bar{X})} \left( \frac{\lambda_0}{\bar{X}} \right)^{\sum x_i}
\]
Example (cont’d)

LRT is to reject $H_0$ when $\lambda(x) \leq c$

$$\alpha = \Pr(\lambda(X) \leq c | \lambda_0)$$
Example (cont’d)

LRT is to reject $H_0$ when $\lambda(x) \leq c$

\[
\alpha = \Pr(\lambda(X) \leq c | \lambda_0)
\]

\[
-2 \log \lambda(X) = -2 \left[ -n(\lambda_0 - \bar{X}) + \sum X_i(\log \lambda_0 - \log \bar{X}) \right]
\]
Example (cont’d)

LRT is to reject \( H_0 \) when \( \lambda(x) \leq c \)

\[
\alpha = \Pr(\lambda(X) \leq c \mid \lambda_0) \\
-2 \log \lambda(X) = -2 \left[ -n(\lambda_0 - \bar{X}) + \sum X_i(\log \lambda_0 - \log \bar{X}) \right] \\
= 2n \left( \lambda_0 - \bar{X} - \bar{X} \log \left( \frac{\lambda_0}{\bar{X}} \right) \right) \xrightarrow{d} \chi^2_1
\]

under \( H_0 \), (by Theorem 10.3.1).
Example (cont’d)

Therefore, asymptotic size $\alpha$ test is

$$\Pr(\lambda(X) \leq c | \lambda_0) = \alpha$$
Example (cont’d)

Therefore, asymptotic size $\alpha$ test is

$$\Pr(\lambda(X) \leq c | \lambda_0) = \alpha$$

$$\Pr(-2 \log \lambda(X) \leq c^* | \lambda_0) = \alpha$$
Example (cont’d)

Therefore, asymptotic size \( \alpha \) test is

\[
\Pr(\lambda(\mathbf{X}) \leq c|\lambda_0) = \alpha
\]

\[
\Pr(-2 \log \lambda(\mathbf{X}) \leq c^*|\lambda_0) = \alpha
\]

\[
\Pr(\chi_1^2 \geq c^*) \approx \alpha
\]
Example (cont’d)

Therefore, asymptotic size $\alpha$ test is

\[
\Pr(\lambda(X) \leq c | \lambda_0) = \alpha
\]
\[
\Pr(-2 \log \lambda(X) \leq c^* | \lambda_0) = \alpha
\]
\[
\Pr(\chi_1^2 \geq c^*) \approx \alpha
\]
\[
c^* = \chi_{1,\alpha}^2
\]

rejects $H_0$ if and only if $-2 \log \lambda(x) \geq \chi_{1,\alpha}^2$
Wald Test

Wald test relates point estimator of \( \theta \) to hypothesis testing about \( \theta \).

**Definition**

Suppose \( W_n \) is an estimator of \( \theta \) and \( W_n \sim \mathcal{N}(\theta, \sigma^2_W) \). Then Wald test statistic is defined as
Wald Test

Wald test relates point estimator of $\theta$ to hypothesis testing about $\theta$.

**Definition**

Suppose $W_n$ is an estimator of $\theta$ and $W_n \sim \mathcal{N}(\theta, \sigma_W^2)$. Then Wald test statistic is defined as

$$Z_n = \frac{W_n - \theta_0}{S_n}$$
Wald Test

Wald test relates point estimator of $\theta$ to hypothesis testing about $\theta$.

**Definition**

Suppose $W_n$ is an estimator of $\theta$ and $W_n \sim \mathcal{N}(\theta, \sigma^2_W)$. Then Wald test statistic is defined as

$$Z_n = \frac{W_n - \theta_0}{S_n}$$

where $\theta_0$ is the value of $\theta$ under $H_0$ and $S_n$ is a consistent estimator of $\sigma^W$.
Two-sided Wald Test

$H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$, then Wald asymptotic level $\alpha$ test is to reject $H_0$ if and only if

$$|Z_n| > z_{\alpha/2}$$
Examples of Wald Test

Two-sided Wald Test

\[ H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta \neq \theta_0, \text{ then Wald asymptotic level } \alpha \text{ test is to reject } H_0 \text{ if and only if} \]
\[ |Z_n| > z_{\alpha/2} \]

One-sided Wald Test

\[ H_0 : \theta \leq \theta_0 \text{ vs. } H_1 : \theta > \theta_0, \text{ then Wald asymptotic level } \alpha \text{ test is to reject } H_0 \text{ if and only if} \]
\[ Z_n > z_{\alpha} \]
Remarks

- Different estimators of $\theta$ leads to different testing procedures.
Remarks

- Different estimators of $\theta$ leads to different testing procedures.
- One choice of $W_n$ is MLE and we may choose $S_n' = \frac{1}{I_n(W_n)}$ or $\frac{1}{I_n(\hat{\theta})}$ (observed information number) when $\sigma^2_W = \frac{1}{I_n(\theta)}$. 
Example of Wald Test

Suppose $X_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p)$, and consider testing $H_0: p = p_0$ vs $H_1: p \neq p_0$. 

The MLE of $p$ is $\hat{X} = \frac{1}{n} \sum X_i$, which follows $\mathcal{N}(p, p(1-p)/n)$ by the Central Limit Theorem. The Wald test statistic is $Z_n = \frac{X - p_0}{\sqrt{p_0(1-p_0)/n}}$, where $S_n$ is a consistent estimator of $\sqrt{p(1-p)/n}$, whose MLE is $S_n = \sqrt{\hat{X}(1-\hat{X})/n}$ by the invariance property of MLE.
Example of Wald Test

Suppose $X_i \overset{	ext{i.i.d.}}{\sim} \text{Bernoulli}(p)$, and consider testing $H_0 : p = p_0$ vs $H_1 : p \neq p_0$. MLE of $p$ is $\bar{X}$, which follows

$$\bar{X} \sim \mathcal{N} \left( p, \frac{p(1-p)}{n} \right)$$
Example of Wald Test

Suppose $X_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p)$, and consider testing $H_0 : p = p_0$ vs $H_1 : p \neq p_0$. MLE of $p$ is $\bar{X}$, which follows

$$\bar{X} \sim \mathcal{N} \left( p, \frac{p(1-p)}{n} \right)$$

by the Central Limit Theorem. The Wald test statistic is

$$Z_n = \frac{\bar{X} - p_0}{S_n}$$
Example of Wald Test

Suppose $X_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p)$, and consider testing $H_0 : p = p_0$ vs $H_1 : p \neq p_0$. MLE of $p$ is $\overline{X}$, which follows

$$\overline{X} \sim \mathcal{N} \left( p, \frac{p(1-p)}{n} \right)$$

by the Central Limit Theorem. The Wald test statistic is

$$Z_n = \frac{\overline{X} - p_0}{S_n}$$

where $S_n$ is a consistent estimator of $\sqrt{\frac{p(1-p)}{n}}$,.
Example of Wald Test

Suppose $X_i \sim \text{i.i.d. Bernoulli}(p)$, and consider testing $H_0 : p = p_0$ vs $H_1 : p \neq p_0$. MLE of $p$ is $\overline{X}$, which follows $\overline{X} \sim \mathcal{N} \left( p, \frac{p(1-p)}{n} \right)$ by the Central Limit Theorem. The Wald test statistic is

$$Z_n = \frac{\overline{X} - p_0}{S_n}$$

where $S_n$ is a consistent estimator of $\sqrt{\frac{p(1-p)}{n}}$, whose MLE is

$$S_n = \sqrt{\frac{\overline{X}(1 - \overline{X})}{n}}$$

by the invariance property of MLE.
Therefore, $S_n$ is consistent for $\sqrt{\frac{p(1-p)}{n}}$. The Wald statistic is
Example of Wald Test (cont’d)

Therefore, $S_n$ is consistent for $\sqrt{\frac{p(1-p)}{n}}$. The Wald statistic is

$$Z_n = \frac{\bar{X} - p_0}{\sqrt{\bar{X}(1 - \bar{X})/n}}$$
Example of Wald Test (cont’d)

Therefore, $S_n$ is consistent for $\sqrt{\frac{p(1-p)}{n}}$. The Wald statistic is

$$Z_n = \frac{\bar{X} - p_0}{\sqrt{\frac{\bar{X}(1 - \bar{X})}{n}}}$$

An asymptotic level $\alpha$ Wald test rejects $H_0$ if and only if

$$\left| \frac{\bar{X} - p_0}{\sqrt{\frac{\bar{X}(1 - \bar{X})}{n}}} \right| > \frac{z_\alpha}{2}$$
Summary

Today
- Asymptotics of LRT
- Wald Test

Next Week
- p-Values