Testing

Steven Stern
University of Virginia

December 10, 2015

1 Examples

1. Consider \( X_i \sim iid \, N(\mu, \sigma^2) \), \( i = 1, 2, \ldots, n \) with \( \sigma^2 \) known. We want to test if \( \mu \) is a particular value, eg 8. Let \( \hat{\mu} = \bar{X} \) (why?), and construct

\[
\hat{Z} = \frac{\hat{\mu} - 8}{\sigma/\sqrt{n}}.
\]

If \( \mu = 8 \), then \( \hat{Z} \) should be close to zero in the sense that \( \hat{Z} \sim N(0, 1) \). If \( \mu \neq 8 \), then \( \hat{Z} \) should explode. Thus, by constructing \( \hat{Z} \), we can say something about whether \( \mu = 8 \).

2. Consider a drug test where there is a drug being tested against a placebo. Let the outcome of interest for person \( i \) be \( Y_i(t_i) \) where \( t_i \) is a binary measure determining whether \( i \) received the drug (\( t_i = 1 \)) or the placebo (\( t_i = 0 \)). We want to determine whether the treatment is more effective than the placebo. Assume \( Y_i(t_i) \sim ind(\mu(t_i), \sigma^2) \). We want to test if \( \mu(1) > \mu(0) \). Define

\[
\hat{\mu}(t) = \frac{\sum_i 1(t_i = t) Y_i}{\sum_i 1(t_i = t)}, \quad t = 0, 1
\]

(why?). Then construct a test.

3. [Generalization] Continuing with the last example, assume that there are two possible treatments and a placebo, i.e., \( t_i = 0, 1, 2 \). We want to see if \( \mu(2) > \mu(1) > 0 \). How?

2 Definitions

1. Statistical null hypothesis: an assertion about the distribution of random variables; denoted \( H_0 \).

2. Alternative hypothesis: the opposite of the null of interest; denoted \( H_A \).
• One sided tests
• Two sided tests

3. Simple and composite tests

4. Test statistic: a function of data used to decide whether to reject or not reject (accept) the null hypothesis.

5. Critical region: the region of the support of the test statistic where one would reject the null hypothesis.

6. Errors (mistakes):
   • Type I Error: Reject $H_0$ when it is true (false negative)
   • Type II Error: Accept $H_0$ when it is false (false positive)
   • Probability of a type I error is called the size or significance of the test. It is frequently denoted by $\alpha$ and chosen to be small (e.g., 0.05)
   • Probability of rejecting $H_0$ is called the power of the test, sometimes denoted $\beta(\theta)$. Note that the power conditional on $H_0$ is the size; i.e., $\beta(\theta^0) = \alpha$. One wants to construct a test statistic where the power is very small when $H_0$ is true and very large when $H_0$ is false.
   • $p$-value is the largest value of $\alpha$ such that $H_0$ would be rejected.

   • One-sided tests:
     \[
     H_0 : \theta = \theta^0 \text{ vs } H_A : \theta > \theta^0 \text{ or } H_0 : \theta = \theta^0 \text{ vs } H_A : \theta < \theta^0
     \]
     Note: sometimes “against” is used instead of “vs”
   
   • Generalization:
     \[
     H_0 : \theta \in \Theta^0 \text{ vs } H_A : \theta \notin \Theta^0
     \]
     (why is this a generalization?)
   • Two sided tests:
     \[
     H_0 : \theta = \theta^0 \text{ vs } H_A : \theta \neq \theta^0
     \]

3  More Examples

1. Let $X_i \sim iid\text{Binomial}(N, p), i = 1, 2, \ldots, n$ with $p$ unknown. Note that $EX_i = Np \Rightarrow$
   \[
   \hat{p} = \frac{X}{N}
   \]
   is an unbiased and consistent estimator of $p$.
   Test $H_0 : p = 0.4 \text{ vs } H_A : p = 0.6$.
   We want to define a test such that $\hat{p} < c \Rightarrow \text{accept } H_0, \hat{p} > c \Rightarrow \text{reject } H_0$
   [note similarity to one-sided test]. We need to pick the critical value: $c$. 

2
We want to reject only if we are pretty sure that \( p > 0.4 \) (why biased in favor of accepting?) So set \( \alpha = 0.05 : \Pr [\hat{p} > c \mid H_0] = \alpha \):

\[
\Pr [\hat{p} > c \mid H_0] = \Pr \left[ \frac{X}{N} > c \mid p = 0.4 \right] \\
= \Pr \left[ \frac{\bar{X}/N} {\sqrt{p(1-p)/n}} > \frac{c - p} {\sqrt{p(1-p)/n}} \mid p = 0.4 \right] \\
= \Pr \left[ \frac{\bar{X}/N - 0.4} {\sqrt{0.24/n}} > \frac{c - 0.4} {\sqrt{0.24/n}} \right] \\
= \Pr \left[ Z > \frac{c - 0.4} {\sqrt{0.24/n}} \right]
\]

where \( Z \sim N(0,1) \). So solve for

\[
1 - \Phi \left( \frac{c - 0.4} {\sqrt{0.24/n}} \right) = \alpha = 0.05
\]

\[
\Rightarrow \Phi \left( \frac{c - 0.4} {\sqrt{0.24/n}} \right) = 0.95
\]

\[
\Rightarrow \frac{c - 0.4} {\sqrt{0.24/n}} = \Phi^{-1}(0.95)
\]

\[
\Rightarrow c = 0.4 + \sqrt{0.24/n}\Phi^{-1}(0.95),
\]

and look up in a standard normal distribution table. Note what happens to \( c \) as \( n \) gets larger. Why?

What happens to the test and critical value if \( H_A : p = 0.7 \) instead?

What is the power of the test?

\[
\Pr [\text{Reject} \mid p] = \Pr \left[ \frac{X}{N} > c \mid p \right] \\
= \Pr \left[ \frac{\bar{X}/N - p} {\sqrt{p(1-p)/n}} > \frac{c - p} {\sqrt{p(1-p)/n}} \mid p \right] \\
= \Pr \left[ \frac{\bar{X}/N - p} {\sqrt{p(1-p)/n}} > \frac{0.4 + \sqrt{0.24/n}\Phi^{-1}(0.95) - p} {\sqrt{p(1-p)/n}} \mid p \right] \\
= \Pr \left[ Z(p) > \frac{0.4 + \sqrt{0.24/n}\Phi^{-1}(0.95) - p} {\sqrt{p(1-p)/n}} \mid p \right]
\]

where \( Z(p) \sim N(0,1) \mid p \). What does the power look like?
2. Let \( X_i \sim iidN(\mu, \sigma^2), \ i = 1, 2, ..., n \) with both \((\mu, \sigma^2)\) unknown. We want to test \( H_0 : \mu = 4 \) vs \( H_A : \mu \neq 4 \).

Note that \( \hat{\mu} = \bar{X} \) is an unbiased consistent estimate of \( \mu \), and accept \( H_0 \) iff \( a \leq \hat{\mu} \leq b \), i.e. critical region for \( \hat{\mu} \) is \( \{ \hat{\mu} < a \cup \hat{\mu} > b \} \).

We need to pick \((a, b)\):

\[
\Pr [a \leq \bar{X} \leq b \mid H_0] = \alpha
\]

\[
\Pr [a \leq \bar{X} \leq b \mid H_0] = \Pr [a \leq \bar{X} \leq b \mid \mu = 4] = \Pr \left[ \frac{a - 4}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - 4}{\sigma/\sqrt{n}} \leq \frac{b - 4}{\sigma/\sqrt{n}} \right]
\]

But \( \sigma^2 \) is unknown. So replace \( \sigma^2 \) with

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]
\[
\Pr \left[ a \leq \bar{X} \leq b \mid H_0 \right] \\
= \Pr \left[ \frac{a - 4}{s/\sqrt{n}} \leq \frac{X - 4}{s/\sqrt{n}} \leq \frac{b - 4}{s/\sqrt{n}} \right] \\
= \Pr \left[ \sqrt{n} (a - 4) \leq \sqrt{n} (X - 4) \leq \sqrt{n} (b - 4) \right] \\
= \Pr \left[ \frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} \leq \frac{\sqrt{n} (X - 4)}{\sqrt{s^2}} \leq \frac{\sqrt{n} (b - 4)}{\sqrt{s^2}} \right] \\
= \Pr \left[ \frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} \leq \frac{\sqrt{n} (X - 4)}{\sqrt{(n - 1) s^2/\sigma^2 (n - 1)}} \leq \frac{\sqrt{n} (b - 4)}{\sqrt{s^2}} \right] \\
= \Pr \left[ \frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} \leq \frac{N (0, 1)}{\sqrt{\chi_{n-1}/\sigma^2 (n - 1)}} \leq \frac{\sqrt{n} (b - 4)}{\sqrt{s^2}} \right] \\
= \Pr \left[ \frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} \leq t_{n-1} \leq \frac{\sqrt{n} (b - 4)}{\sqrt{s^2}} \right] = \alpha.
\]

Use t-table values for \( \frac{\alpha}{2} \) and \( 1 - \frac{\alpha}{2} \). Note difference from a one-sided test:
(assume \( n \) is big)

\[
\Pr \left[ t_{n-1} < \frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} \right] = \frac{\alpha}{2} \\
\Phi \left( \frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} \right) = \frac{\alpha}{2} \\
\frac{\sqrt{n} (a - 4)}{\sqrt{s^2}} = \Phi^{-1} \left( \frac{\alpha}{2} \right)
\]

\[
a = 4 + \frac{\sqrt{s^2}}{\sqrt{n}} \Phi^{-1} \left( \frac{\alpha}{2} \right)
\]

\[
a = 4 - \frac{\sqrt{s^2}}{\sqrt{n}} \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right)
\]

\[
\Pr \left[ t_{n-1} < \frac{\sqrt{n} (b - 4)}{\sqrt{s^2}} \right] = 1 - \frac{\alpha}{2} \\
\Phi \left( \frac{\sqrt{n} (b - 4)}{\sqrt{s^2}} \right) = 1 - \frac{\alpha}{2} \\
b = 4 + \frac{\sqrt{s^2}}{\sqrt{n}} \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right)
\]
What does the power look like?

\[
\Pr [\text{Reject } | \mu] = \Pr [a \leq \bar{X} \leq b | \mu] = \Pr \left[ \frac{a - \mu}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{b - \mu}{\sigma/\sqrt{n}} \right] = \Pr \left[ \frac{4 - \frac{\chi^2}{n}}{\sqrt{n}} \Phi^{-1} \left( 1 - \frac{c}{2} \right) - \mu}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{4 + \frac{\chi^2}{n}}{\sqrt{n}} \Phi^{-1} \left( 1 - \frac{c}{2} \right) - \mu} \right]
\]

Power

![Power Graph](power.png)

3. Let \( X_i \sim iidN (\mu, \sigma^2), i = 1, 2, ..., n \) with \((\mu, \sigma^2)\) unknown.

Test \( H_0: \sigma^2 = 9 \) vs \( H_A: \sigma^2 < 9 \).

Use

\[
S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2
\]

as a test statistic because \( S^2 \) is a consistent estimate of \( \sigma^2 \) and \((n - 1) S^2/\sigma^2 \sim \chi^2_{n-1}\). We want to reject \( H_0 \) iff \( S^2 < c \) (why?).

To find \( c \), set

\[
\Pr [S^2 < c | H_0] = \alpha
\]

\[
\Pr [S^2 < c | H_0] = \Pr \left[ \frac{(n-1) S^2}{\sigma^2} < \frac{(n-1)c}{\sigma^2} | \sigma^2 = 9 \right] = \Pr \left[ \chi^2_{n-1} < \frac{(n-1)c}{9} \right]
\]
4 Best Critical Regions and Likelihood Ratio Tests

Let $L(X \mid \theta^0)$ be the likelihood function at the $H_0$ value of $\theta$, $\theta^0$, and let $L(X \mid \theta^1)$ be the likelihood function at some other value of $\theta$, $\theta^1$. Define

$$
\lambda = \frac{L(X \mid \theta^0)}{L(X \mid \theta^1)}
$$

The goal is to pick a critical region $C$ so that

$$
\lambda \begin{cases} 
\leq k & \text{if } X \in C \\
> k & \text{if } X \notin C
\end{cases}.
$$

(1)

More intuitively, the goal is to minimize Type II error given a particular size.

4.1 Examples

1. Let $X_i \sim iid N(\mu, \sigma^2)$, $i = 1, 2, ..., n$ with $\sigma^2$ known.
   Test $H_0 : \mu = 4$ vs $H_0 : \mu > 4$.

   $$
   L(X \mid \mu) = \prod_{i=1}^{n} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{X_i - \mu}{\sigma} \right)^2 \right\}.
   $$

   We know that $\hat{\mu} = \overline{X}$ maximizes $L(X \mid \mu)$.
   If $\overline{X} \leq 4$, accept $H_0$ (because this is a one-sided test);
   If $\overline{X} > 4$, then

   $$
   \lambda = \frac{\prod_{i=1}^{n} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{X_i - 4}{\sigma} \right)^2 \right\}}{\prod_{i=1}^{n} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{X_i - \overline{X}}{\sigma} \right)^2 \right\}}.
   $$

   Now pick a critical region that satisfies equation (1):

   $$
   \lambda \leq k \text{ if } H_0 : \mu > 4.
   $$
2. Let \( X_i \sim \text{iid} e^{-\lambda x}, i = 1, 2, \ldots, n, \) with \( \lambda \) unknown. Test \( H_0 : \lambda = 1 \) vs \( H_A : \lambda \neq 1. \)

\[
L (X \mid \lambda) = \lambda^n e^{-\lambda \bar{X}} = \lambda^n e^{-\lambda \overline{X}} \Rightarrow \bar{\lambda} = \overline{X}^{-1}
\]
$$\delta = \frac{L(X | \lambda^0)}{L(X | \lambda)} = \frac{e^{-nX}}{\lambda^n e^{-n\lambda X}} = \frac{X e^{-nX}}{e^{-n}} = X^n e^{-n(X-1)}$$

Now pick a critical region that satisfies equation (1):

$$\delta \leq k \text{ iff } H_0 : \log \delta = n \log \bar{X} - n(\bar{X} - 1)$$

So pick a critical region in terms of $\bar{X}$.
We can find the distribution of $\bar{X}$ using moment generating functions:

$$M_{\bar{X}}(t) = \frac{1}{1 - \lambda t}$$

for the exponential density. $\Rightarrow$

$$M_{\bar{X}}(t) = \left( \frac{1}{1 - \frac{\lambda}{n}} \right)^n$$

which is the moment generating function for a gamma density:

$$f(\bar{x}) = \frac{1}{\Gamma(n) \left( \frac{\lambda}{n} \right)^n} \bar{x}^{n-1} \exp \left\{ -\frac{n\bar{x}}{\lambda} \right\}.$$ 

So pick critical values:

$$F(a) = 1 - F(b) = \frac{\alpha}{2}.$$

9
3. Let \( X_i \sim iidN(\mu, \sigma^2) \), \( i = 1, 2, ..., n \), with \( (\mu, \sigma^2) \) unknown. Test \( H_0 : \mu = 0, \sigma^2 = 1 \) vs \( H_A : \mu \neq 0, \sigma^2 \neq 1 \). Then

\[
\lambda = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{X_i^2}{2} \right\} \prod_{i=1}^{n} \frac{1}{\hat{\sigma}\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\hat{\sigma}^2} (X_i - \hat{\mu})^2 \right\}
\]

\[
\log \lambda = \sum_{i=1}^{n} \left[ -\frac{X_i^2}{2} + \frac{1}{2\hat{\sigma}^2} (X_i - \hat{\mu})^2 + \log \hat{\sigma} \right]
\]

\[
= -0.5 \sum_{i=1}^{n} X_i^2 + \frac{n}{2} + n \log \hat{\sigma}.
\]

It is very difficult, if not impossible, to derive the joint distribution of \( \left( \sum_{i=1}^{n} X_i^2, \log \hat{\sigma} \right) \). Instead, rely on asymptotic result:

\[-2 \log \lambda \sim \chi^2_k\]

where \( k \) is the number of restrictions imposed by the null hypothesis; in this case, \( k = 2 \).

5 Wald Tests, Lagrange Multiplier Tests, and Likelihood Ratio Tests

Consider \( H_0 : \theta = \theta^0 \) vs \( H_A : \theta \neq \theta^0 \). Let \( L(X | \theta) \) be the likelihood function.

5.1 Likelihood Ratio Tests

Define

\[
\lambda(\theta) = \frac{L(X | \theta^0)}{L(X | \theta)},
\]

and choose \( C \) and \( k \) such that \( \lambda(\hat{\theta}) \leq k \) iff \( X \in C \). Then

\[-2 \log \lambda(\hat{\theta}) \sim \chi^2_m\]

under \( H_0 \) where \( m \) is the number of restrictions implied by \( H_0 \) vs \( H_A \).

5.2 Lagrange Multiplier Tests

At the maximum, \( \frac{\partial}{\partial \theta} L(X | \theta) = 0 \). This suggests \( H_0 : \frac{\partial}{\partial \theta} L(X | \theta^0) = 0 \) vs \( H_A : \frac{\partial}{\partial \theta} L(X | \theta^0) \neq 0 \).

Define

\[
S(\theta) = \frac{\partial}{\partial \theta} \log L(X | \theta)
\]
as the score statistic. Then, under $H_0$,
\[ S (\theta^0)' D^{-1} \begin{bmatrix} S (\theta^0) \\ S (\theta^0) \end{bmatrix} S (\theta^0) \sim \chi^2_m \]
where $m$ is the number of restrictions.

### 5.3 Wald Test

Under $H_0$, $\hat{\theta} - \theta^0$ should be close to zero. This suggests a Wald test,
\[ W = \left( \hat{\theta} - \theta^0 \right)' D^{-1} \begin{bmatrix} \hat{\theta} \\ \hat{\theta} \end{bmatrix} \left( \hat{\theta} - \theta^0 \right) \sim \chi^2_m. \]

### 5.4 Example 1

Let $X_i \sim iidN (\mu, \sigma^2)$, $i = 1, 2, .., n$, with $\sigma^2$ known.
Test $H_0 : \mu = 0$ vs $H_A : \mu \neq 0$.

\[
L (X \mid \mu) = \prod_{i=1}^{n} \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ - \frac{1}{2\sigma^2} (X_i - \mu)^2 \right\}.
\]

- For LR test, reject iff $\overline{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$ does not include 0 (from previous example).
- For LM test,
\[
\log L (X \mid \mu) = - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2
\]
\[
\frac{\partial}{\partial \mu} \log L (X \mid \mu) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu).
\]

Note that
\[
E \left[ \frac{\partial}{\partial \mu} \log L (X \mid \mu) \right] = 0,
\]
\[
E \left[ \left( \frac{\partial}{\partial \mu} \log L (X \mid \mu) \right)^2 \right] = \frac{1}{\sigma^4} E \sum_{i=1}^{n} (X_i - \mu)^2
\]
Why?
\[
= \frac{n}{\sigma^2}.
\]

\[
LM = \left[ \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu^0) \right] \left( \frac{\sigma^2}{n} \right) \left[ \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu^0) \right]
\]
\[
= \left[ \sum_{i=1}^{n} (X_i - \mu^0) \right]^2 / n\sigma^2 = n\overline{X^2} / \sigma^2.
\]

What is the distribution? Why?
For Wald test,

\[ \hat{\mu} - \mu^0 = X, \]

\[ \text{Var} X = \frac{\sigma^2}{n} \]

\[ \Rightarrow \]

\[ W = n\overline{X^2}/\sigma^2 \]

which has the same distribution as the LM test.

5.5 Example 2

\[ y_i^* = \alpha + \beta x_i + u_i, \ i = 1, 2, .., n \]
\[ u_i \sim iidN(0,1) \]
\[ y_i = 1(y_i^* > 0) \]

\[ H_0 : \beta = 0 \ vs \ H_A : \beta \neq 0 \]

\[ \log L = \sum_i y_i \log \Phi (\alpha + \beta x_i) + (1 - y_i) \log [1 - \Phi (\alpha + \beta x_i)] \]

\[ \frac{\partial}{\partial \beta} \log L = \sum_i y_i \frac{\phi (\alpha + \beta x_i) x_i}{\Phi (\alpha + \beta x_i)} - (1 - y_i) \frac{\phi (\alpha + \beta x_i) x_i}{1 - \Phi (\alpha + \beta x_i)} \]
\[ = \sum_i \frac{\phi (\alpha + \beta x_i) x_i}{\Phi (\alpha + \beta x_i) [1 - \Phi (\alpha + \beta x_i)]} [y_i - \Phi (\alpha + \beta x_i)] \]
\[ = \sum_i z_i (\alpha, \beta, x_i) [y_i - \Phi (\alpha + \beta x_i)] \]

LM test statistic:

\[ S = \frac{1}{n} \sum_i z_i (\hat{\alpha}, 0, x_i) [y_i - \Phi (\hat{\alpha})] \]

\[ \hat{\alpha} \] solves

\[ 0 = \sum_i \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} [y_i - \Phi (\alpha)] \]
\[ = \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} \sum_i [y_i - \Phi (\alpha)] \]
\[ = \sum_i [y_i - \Phi (\alpha)] \]
\[ \overline{y} = \Phi (\alpha) \Rightarrow \hat{\alpha} = \Phi^{-1} (\overline{y}) \]
LM test statistic is
\[
\frac{1}{n} \sum_i z_i (\hat{\alpha}, 0, x_i) [y_i - \Phi (\hat{\alpha})] \bigg( \frac{1}{n} \sum_i \Phi (\alpha) [1 - \Phi (\alpha)] \bigg) = \frac{1}{n} \sum_i \frac{\phi (\alpha) x_i}{\Phi (\alpha) [1 - \Phi (\alpha)]} \bigg( \frac{1}{n} \sum_i \Phi (\alpha) [1 - \Phi (\alpha)] \bigg) \nonumber
\]
\[
= \frac{1}{n} \left( \frac{\phi (\alpha) x_i}{\Phi (\alpha) [1 - \Phi (\alpha)]} \bigg( \frac{1}{n} \sum_i \Phi (\alpha) [1 - \Phi (\alpha)] \bigg) \right) \nonumber
\]
This conditions on \( \alpha = \hat{\alpha} \). Without conditioning, the score statistic (evaluated under \( H_0 \)) is
\[
S = \left( \frac{1}{n} \sum_i \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} \bigg( \frac{1}{n} \sum_i [y_i - \Phi (\hat{\alpha})] \bigg) \right) \bigg( \frac{1}{n} \sum_i [y_i - \Phi (\hat{\alpha})] \bigg) \nonumber
\]
with covariance matrix,
\[
ESS' = V
\]
with
\[
V_{11} = E \left( \frac{1}{n} \sum_i \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} \bigg( \frac{1}{n} \sum_i [y_i - \Phi (\alpha)] \bigg) \right)^2 \nonumber
\]
\[
= \frac{1}{n^2} \left( \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} \right)^2 \sum_i E [y_i - \Phi (\alpha)]^2 \nonumber
\]
\[
= \frac{1}{n^2} \left( \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} \right)^2 \sum_i \Phi (\alpha) [1 - \Phi (\alpha)] \nonumber
\]
\[
= \frac{1}{n} \frac{(\phi (\alpha))^2}{\Phi (\alpha) [1 - \Phi (\alpha)]}; \nonumber
\]
\[
V_{22} = E \left( \frac{1}{n} \sum_i \frac{\phi (\alpha)}{\Phi (\alpha) [1 - \Phi (\alpha)]} \bigg( \frac{1}{n} \sum_i x_i [y_i - \Phi (\alpha)] \bigg) \right)^2 \nonumber
\]
\[
= \frac{1}{n^2} \left( \frac{\phi (\alpha))^2}{\Phi (\alpha) [1 - \Phi (\alpha)]} \right) \sum_i x_i^2; \nonumber
\]
\[
V_{12} = V_{21} = E \left( \frac{1}{n} \Phi(\alpha) \left[ 1 - \Phi(\alpha) \right] \sum_i [y_i - \Phi(\alpha)] \right) \left( \frac{1}{n} \Phi(\alpha) \left[ 1 - \Phi(\alpha) \right] \sum_i x_i [y_i - \Phi(\alpha)] \right) \\
= \frac{1}{n^2} \left( \frac{\phi(\alpha)}{\Phi(\alpha) \left[ 1 - \Phi(\alpha) \right]} \right)^2 E \left( \sum_i [y_i - \Phi(\alpha)] \right) \left( \sum_i x_i [y_i - \Phi(\alpha)] \right) \\
= \frac{1}{n^2} \left( \frac{\phi(\alpha)}{\Phi(\alpha) \left[ 1 - \Phi(\alpha) \right]} \right)^2 \left( \sum_i x_i E [y_i - \Phi(\alpha)]^2 \right) \\
= \frac{1}{n^2 \Phi(\alpha) \left[ 1 - \Phi(\alpha) \right]} \sum_i x_i.
\]

\[
\Rightarrow
V = \frac{1}{n^2} \frac{(\phi(\alpha))^2}{\Phi(\alpha) \left[ 1 - \Phi(\alpha) \right]} \left( \frac{n}{\sum_i x_i} \frac{\sum_i x_i}{\sum_i x_i^2} \right) \\
V^{-1} = \frac{1}{n^2} \frac{(\phi(\alpha))^2}{\Phi(\alpha) \left[ 1 - \Phi(\alpha) \right]} \left( \frac{\sum_i x_i^2 - (\sum_i x_i)^2}{n^2} \right).
\]

LM test statistic is

\[
\begin{align*}
&\left( \frac{1}{n} \frac{\phi(\hat{\alpha})}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \sum_i [y_i - \Phi(\hat{\alpha})] \right)' \left( \frac{1}{n} \frac{\phi(\hat{\alpha})}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \sum_i [y_i - \Phi(\hat{\alpha})] \right) \\
&\frac{1}{n^2} \frac{(\phi(\hat{\alpha}))^2}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \left( n \sum_i x_i^2 - (\sum_i x_i)^2 \right)
\end{align*}
\]

\[
= \frac{1}{n} \frac{\phi(\hat{\alpha})}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \sum_i x_i [y_i - \Phi(\hat{\alpha})] \left( \frac{0}{n} \frac{\phi(\hat{\alpha})}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \sum_i x_i [y_i - \Phi(\hat{\alpha})] \right) \\
\frac{1}{n^2} \frac{(\phi(\hat{\alpha}))^2}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \left( n \sum_i x_i^2 - (\sum_i x_i)^2 \right)
\]

\[
= \frac{1}{n} \left( \frac{(\phi(\hat{\alpha}))^2}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \right)^2 \left( \sum_i x_i [y_i - \Phi(\hat{\alpha})]\right)^2 \\
\frac{1}{n^2} \frac{(\phi(\hat{\alpha}))^2}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \left( n \sum_i x_i^2 - (\sum_i x_i)^2 \right)
\]

\[
= \frac{1}{n} \frac{\sum_i x_i [y_i - \Phi(\hat{\alpha})]^2}{\Phi(\hat{\alpha}) \left[ 1 - \Phi(\hat{\alpha}) \right]} \left( \frac{1}{n} \sum_i x_i^2 - \left( \frac{1}{n} \sum_i x_i \right)^2 \right).
\]

6 Confidence Intervals

Confidence intervals are directly analogous to tests. But they are used to make a different (somewhat convoluted statement). Let \( \theta \) a parameter of interest. Using an estimate of \( \theta, \hat{\theta} \), we want to construct a 100 \( (1 - \alpha) \) \% confidence interval for \( \theta \) such that, for any point in the confidence interval, the “probability” of \( \hat{\theta} \) is at least \( 1 - \alpha \).
6.1 Normal Example

Let $X_i \sim iidN(\mu, \sigma^2), i = 1, 2, \ldots, n$, with $\sigma^2$ known. Use $\bar{X}$ as an estimator of $\mu$. Choose $\alpha$, and find $z_{\alpha/2}$. Then

$$\Pr \left[ -z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \right] = 1 - \alpha$$

$$\Pr \left[ -\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq \bar{X} - \mu \leq \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right] = 1 - \alpha$$

$$\Pr \left[ \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \geq \mu - \bar{X} \geq -\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right] = 1 - \alpha$$

$$\Pr \left[ \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \geq \mu \geq \bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right] = 1 - \alpha.$$

$\Rightarrow$ Confidence interval is $\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$.

Compare to test for $\mu$.

What happens once $\bar{X}$ is realized?

6.2 Other Distributions

Use CLT if too difficult to do for small-sample distribution.

6.3 When $\sigma^2$ Unknown

Replace $\sigma^2$ with $s^2$ and use $t$–distribution.

6.4 Confidence Interval for Difference of Means

Let

$$X_i \sim iidN(\mu_x, \sigma_x^2), i = 1, 2, \ldots, n_x$$

$$Y_i \sim iidN(\mu_y, \sigma_y^2), i = 1, 2, \ldots, n_y,$$

and find a confidence interval for $\mu_x - \mu_y$.

Define

$$Z = \bar{X} - \bar{Y}$$

and note that

$$EZ = \mu_x - \mu_y,$$

$$VarZ = \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}$$

$\Rightarrow$

$$CI: \bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$
If \((\sigma_x^2, \sigma_y^2)\) unknown, use

\[
\frac{(n_x - 1) s_x^2 + (n_y - 1) s_y^2}{n_x + n_y - 2} \left[ \frac{1}{n_x} + \frac{1}{n_y} \right]
\]

as an estimate of \(\sigma_x^2\) if comfortable assuming \(\sigma_x^2 = \sigma_y^2\).

6.5 Confidence Interval for Variance

Let \(X_i \sim iidN(\mu, \sigma^2), i = 1, 2, \ldots, n\), with \((\mu, \sigma^2)\) unknown. Then

\[
\Pr \left[ a \leq \frac{(n - 1) s^2}{\sigma^2} \leq b \right] = 1 - \alpha
\]

\[
\Pr \left[ \frac{1}{a} \geq \frac{\sigma^2}{(n - 1) s^2} \geq \frac{1}{b} \right] = 1 - \alpha
\]

\[
\Pr \left[ \frac{(n - 1) s^2}{a} \geq \sigma^2 \geq \frac{(n - 1) s^2}{b} \right] = 1 - \alpha.
\]

Choose \((a, b)\) appropriately.