Economics 361
Hypothesis Testing: Part I

Jun Ishii *
Department of Economics
Amherst College

Fall 2023

1 Overview

Not all statistical inference involves pinning down a value for some unknown parameter of the Data Generating Process (DGP). Some statistical inference involves examining the degree to which available data (sample) is consistent with some conjecture about the DGP. More specifically,

- The conjecture is a **hypothesis** about the DGP
- The statistician/econometrician uses some function of the sample, **test statistic**, to summarize the information in the sample concerning the hypothesis.
- A testing procedure is adopted involving some rejection criterion
  - if the value of the test statistic falls within some set of values, **critical region**, the hypothesis is rejected
  - if the value of the test statistic falls outside the critical region, the hypothesis fails to be rejected.

This practice of statistical inference is referred to as formal **hypothesis testing**. The triplet of (hypothesis, test statistic, critical region) operationally defines a hypothesis test. However, the triplet in of themselves do not indicate the properties/desirability of the defined hypothesis test.

In this handout, we extend the Gosset-Fisher insight on the distribution of estimators to the distribution of test statistics and use the Neyman-Pearson criteria for choosing hypothesis tests. This synthesis between Gosset-Fisher and Neyman-Pearson is the current standard for formal hypothesis testing (under the frequentist notion of probability).

*Office: Converse Hall 315 Phone: (413) 542-2901 E-mail: jishii@amherst.edu

2 Hypotheses

There are many types of conjectures one can formulate about the Data Generating Process (DGP). In this course, we focus on conjectures that can be expressed as functions of the DGP parameters.[1] Let θ be the vector of parameters underlying the DGP.

Example: For the case of the OLS Model under the Gauss-Markov Assumptions

$$\theta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{k-1} \\ \sigma^2 \end{pmatrix} \qquad \text{Note: } \theta \text{ is a } (k+1) \times 1 \text{ column vector}$$

Let $h(\cdot)$ be some function. Consider two types of hypotheses based on $h(\theta)$

- **Simple**: $h(\theta) = r$ where r is some constant known value
- **Composite**: $h(\theta) \in S_r$ where S_r is a set of constant known values

Example: Coefficients of the Cobb-Douglas Production Function $Y = AK^{\beta_1} L^{\beta_2}$

- **Simple**: $\beta_1 + \beta_2 = 1$ (Constant Returns to Scale Hypothesis)
- **Composite**: $\beta_1 + \beta_2 > 1$ (Increasing Returns to Scale Hypothesis)

Each hypothesis test consists of two hypotheses:

- **Null** (H_0): the hypothesis being tested
- **Alternative** (H_a): the hypothesis being supported if H_0 is rejected

H_0 and H_a should [1] be mutually exclusive and [2] span the possible states of the world. Together, this means that one and only one of the two hypotheses should be true. If our hypothesis test rejects one hypothesis, we are implicitly supporting the other hypothesis.

Example: Testing for Constant Returns to Scale for Cobb-Douglas Production

- **Null** (H_0): $\beta_1 + \beta_2 = 1$
- **Alternative** (H_a): $\beta_1 + \beta_2 \neq 1$ (assuming increasing/decreasing also possible)

In this course, we will focus only on simple null hypotheses: i.e. those that can be expressed as $g(\theta) = r$. It is possible to conduct hypothesis tests involving composite null hypotheses but they are outside the scope of this course. Moreover, hypothesis tests involving composite null hypotheses are rare in economics and other social sciences.[2] However, we will allow alternative hypotheses to be simple or composite.

We will discuss the role of the alternative hypothesis more when we discuss the choice of critical region for a hypothesis test.

[1] For now, we will deal with a single statement about the DGP. Hypotheses that involve multiple statements – "joint hypotheses" – will be discussed later

[2] This is partly because conducting hypothesis tests involving composite nulls are difficult. See, for example, "An Exact Test for Multiple Inequality and Equality Constrains in the Linear Regression Model," by F. Wolak, *Journal of the American Statistical Association*, September 1987, pp.782-793, which examines tests of nulls like $\beta_1 + \beta_2 \geq 1$

3 Test Statistics

Given a simple null hypothesis expressed as a function of the DGP parameters θ

- $H_0: \; h(\theta) = r$

a natural candidate for the test statistic (TS) is

- $TS \; = \; h(\hat{\theta})$

where $\hat{\theta}$ is an estimate of θ using the available sample.

Additionally, null hypotheses are often expressed such that the function equals zero.

- $H_0: \; h(\theta) - r = 0$
- $TS \; = \; h(\hat{\theta}) - r$

For example, consider the OLS estimator of the log-linearized Cobb-Douglas production function

$$\begin{aligned} Y \; = \; AK^{\beta_1}L^{\beta_2} \;\; \Longrightarrow \;\; ln(Y) \; &= \; ln(A) \; + \; \beta_1 \; ln(K) \; + \; \beta_2 \; ln(L) \\ E[ln(Y) \mid ln(K), ln(L)] \; &= \; \underbrace{E[ln(A) \mid ln(K), ln(L)]}_{\beta_0} \; + \; \beta_1 \; ln(K) \; + \; \beta_2 \; ln(L) \\ b^{ols} \; &= \; \begin{pmatrix} b_0^{ols} \\ b_1^{ols} \\ b_2^{ols} \end{pmatrix} \; = \; \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \end{aligned}$$

and the simple null hypothesis of constant returns to scale: $\beta_1 + \beta_2 - 1 = 0$. The natural test statistic candidate would be $\hat{\beta}_1 + \hat{\beta}_2 - 1$ which, in matrix notation, can be expressed as

$$TS \; = \; Rb^{ols} - r \; = \; \underbrace{\begin{pmatrix} 0 & 1 & 1 \end{pmatrix}}_{R} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} - \underbrace{1}_{r} \; = \; \hat{\beta}_1 + \hat{\beta}_2 - 1 \; = \; b_1^{ols} + b_2^{ols} - 1$$

In this course, we will focus primarily on these "natural" test statistics.[3] As the test statistics are functions of the estimators and the estimators random variables, the test statistics themselves are also random variables. We will derive the distribution of the test statistic from the distribution of the underlying estimators.

[3] Most of the test statistics that are taught in standard introductory statistics/econometrics courses – e.g. the t-statistic and F-statistic – are of this kind

3.1 "Trivial" Null with Normal Estimators and Finite Random Sample

When the hypothesis concerns the value of the estimated parameter, itself, then the test statistic is simply the estimator.

Consider the OLS model when the Gauss-Markov and the Normality assumptions hold. Given a finite random sample, we can show that

$$b^{ols} \quad \sim \quad N(\beta, \sigma^2 \underbrace{(X'X)^{-1}}_{Q^{-1}})$$

The above multivariate statement also implies that each element in b^{ols} is also distributed Normal

$$b_j^{ols} \quad \sim \quad N(\beta_j, \sigma^2 Q_{jj}^{-1}) \qquad \text{for } j = 0 \text{ to } k-1$$

where Q_{jj}^{-1} is the j^{th} diagonal element of $Q^{-1} \equiv (X'X)^{-1}$.

Any simple null hypothesis on the value of b_j^{ols}

- $H_0 : \beta - r = 0 \qquad$ (i.e. $\beta = r$)

can be tested using the test statistic

- $TS : b_j^{ols} - r$

We can show using change of variables that

$$\frac{b_j^{ols} - r}{\sqrt{\sigma^2 Q_{jj}^{-1}}} \quad \sim \quad N(\beta - r, 1)$$

If the null hypothesis (H_0) is correct, then $\frac{b_j^{ols} - r}{\sigma^2 \sqrt{Q_{jj}}}$ is distributed standard Normal as $\beta - r = 0$. The above test statistic is known as the **z-statistic**.

If σ^2 is unknown, then the above test-statistic cannot be used. We know b^{ols}, Q_{jj}^{-1}, and r but we do not know σ^2. So we cannot calculate the z-statistic.

We can, however, replace the unknown σ^2 with our estimator $s^2 = \frac{e'e}{N-k} = \frac{(Y - Xb^{ols})'(Y - Xb^{ols})}{N-k}$ and arrive at a modified test statistic. If H_0 is true (so $\beta = r$), we can also derive the distribution of this modified test statistic

$$\frac{b_j^{ols} - r}{\sqrt{s^2 Q_{jj}^{-1}}} \quad \sim \quad t_{N-k}$$

This modified test statistic is known as the **t-statistic**

Note the difference between the two test statistic. Under H_0, the z-statistic is distributed standard Normal but the t-statistic the "t" distribution with $N-k$ degrees of freedom.[4] Interestingly enough, as $N \to \infty$ the t-distribution with $N-k$ degrees of freedom converges to the standard Normal distribution; this result is consistent with the following intuition: as the sample size grows to infinity, the estimator s^2 converges to σ^2.

3.2 Linear Null with Normal Estimators and Finite Random Sample

Consider the OLS model under the same set of assumptions as before (Gauss-Markov, Normality, finite random sample). But now, consider a simple null hypothesis concerning the value of a linear function of the parameters

- $H_0 : R\beta - r = 0$

where R is a $(1 \times k)$ vector of known constants and r a known constant.

Example: For $H_0 : 0.5\ \beta_1 - 1.2\ \beta_3 - 2.1 = 0$ and $k = 5$

$$R = \begin{pmatrix} 0 & 0.5 & 0 & -1.2 & 0 \end{pmatrix} \quad \text{and} \quad r = 2.1$$

The natural candidate for the test statistic would be

- $TS = Rb^{ols} - r$

Using change of variables, we can show that the linear function of Normal random variables is, itself, a Normal random variable. therefore

$$Rb^{ols} - r \quad \sim \quad N(R\beta - r, \sigma^2 RQ^{-1}R')$$

ASIDE: If the linear function only involves a few elements of b^{ols}, it is often easier to calculate the mean and variance of $(Rb^{ols} - r)$ directly
Example: $H_0 : 0.5\ \beta_1 - 1.2\ \beta_3 - 2.1 = 0$

$$\begin{aligned}
\text{Var}(0.5\ b_1^{ols} - 1.2\ b_3^{ols} - 2.1|X) &= (0.5)^2\text{Var}(b_1^{ols}|X) + (-1.2)^2\text{Var}(b_3^{ols}|X) \\
&\quad +2(0.5)(-1.2)\text{Cov}(b_1^{ols}, b_3^{ols}|X) \\
&= \underbrace{\sigma^2 \left(0.25\ Q_{11}^{-1} + 1.44\ Q_{33}^{-1} - 1.2\ \underbrace{Q_{13}^{-1}}_{=Q_{31}^{-1}} \right)}_{RQ^{-1}R'} \\
&\quad Q_{ij}^{-1} \text{ is the } i^{th} \text{ row } j^{th} \text{ column element of } Q^{-1} \\
E[0.5\ b_1^{ols} - 1.2\ b_3^{ols} - 2.1|X] &= 0.5\ E[b_1^{ols}|X] - 1.2\ E[b_3^{ols}|X] - 2.1 \\
&= \underbrace{0.5\ \beta_1 - 1.2\ \beta_3}_{R\beta} - \underbrace{2.1}_{r}
\end{aligned}$$

[4] "Under H_0" is statistics shorthand for "Under the assumption that the null hypothesis is true"

Furthermore, we can show that under H_0

$$\underbrace{\frac{Rb^{ols} - r}{\sqrt{\sigma^2 RQ^{-1}R'}}}_{\text{“z-statistic”}} \quad \sim \quad N(0,1) \qquad \text{and} \qquad \underbrace{\frac{Rb^{ols} - r}{\sqrt{s^2 RQ^{-1}R'}}}_{\text{“t-statistic”}} \quad \sim \quad t_{N-k}$$

3.3 Non-linear Null with Normal Estimators and Finite Random Sample

Consider the OLS model under the same set of assumptions as before (Gauss-Markov, Normality, finite random sample). But now, consider a simple null hypothesis concerning the value of a non-linear function of the parameters.

- $H_0 : h(\beta) - r = 0$ where $h(\cdot)$ is some non-linear function

Unfortunately, there is no general statement one can make about the distribution of a non-linear function of Normal random variables. Thus, we cannot make any general statements about the distribution of $h(b^{ols}) - r$.

We know the distribution of specific non-linear functions of random variables. Of particular interest is the distribution of the square of a standard Normal random variable

$$(Z)^2 \quad \sim \quad \chi^2_1 \qquad \text{where } Z \sim N(0,1)$$

Furthermore, we can show that

$$(Z_1)^2 + (Z_2)^2 + \cdots + (Z_k)^2 \quad = \quad \sum_{j=1}^{J} (Z_j)^2 \quad \sim \quad \chi^2_J$$

if $\{(Z_0)^2, (Z_1)^2, \ldots, (Z_{k-1})^2\}$ are distributed mutually independent.

The above result is what allows us to argue that

$$\frac{1}{\sigma^2} \sum_{i=1}^{N} (Y_i - \sum_{j=0}^{k-1} b_j^{ols} X_{ji})^2 \quad = \quad \frac{(Y - Xb^{ols})'(Y - Xb^{ols})}{\sigma^2} \quad = \quad (N-K)\frac{s^2}{\sigma^2} \quad = \quad \frac{e'e}{\sigma^2} \quad \sim \quad \chi^2_{N-k}$$

The proof is provided in Goldberger Chapter 21.2 and Amemiya Chapter 12.4.1

Test statistics whose distribution under H_0 is chi-squared (χ^2) are known as **chi-squared statistics**. Chi-squared statistics are not often used for tests involving *finite* random samples distributed Normal.

However, they play a major role in **asymptotic tests** – tests based on the asymptotic distribution of a test statistic. Asymptotic tests are used when the sampling distribution is either unknown or of an inconvenient type.[5]

[5] In particular, there are three major classes of asymptotic hypothesis tests that use chi-squared statistics: the Wald test, the Lagrange Multiplier test, and the Likelihood Ratio test. We discuss these later.

4 Critical Region

Given a null hypothesis and a test statistic, we say that a null hypothesis is **rejected** if the value of the test statistic calculated using the sample falls within a **critical region**.

> **Critical Region** is the set of values of the test statistic for which the hypothesis test rejects the null hypothesis (H_0)

There are many ways to choose a critical region, similar to the many ways to choose a test statistic. Here, we adopt the approach credited to Jerzy Neyman and Egon Pearson.

In the Neyman-Pearson approach, we consider two competing hypotheses – the null and the alternative. Two hypotheses must be mutually exclusive and exhausting, the latter meaning that the true state of the world must be consistent with one or the other hypothesis.

> **Example**: Cobb Douglas Production and Economies of Scale: $Y = AK^{\beta_1}L^{\beta_2}$
>
> If all three scale economies (constant, increasing, decreasing) are possible, then
>
> - $H_0 : \beta_1 + \beta_2 = 0$
> - $H_a : \beta_1 + \beta_2 \neq 1$ (Not constant implies either increasing or decreasing)
>
> But if increasing returns to scale is impossible (e.g. technology constraint), then
>
> - $H_0 : \beta_1 + \beta_2 = 0$
> - $H_a : \beta_1 + \beta_2 < 1$ (Not constant implies decreasing only)

There are two decisions that the econometrician may make (reject H_0, fail to reject H_0) and two possible states of the world (H_0 true, H_0 false). This suggests four scenarios which can be illustrated in the following decision matrix

	H_0 True	H_0 False
Reject H_0	**Type I Error**	Correct Decision
Fail to Reject H_0	Correct Decision	**Type II Error**

The econometrician has made the correct decision when she rejects H_0 when H_0 is false and when she fails to reject H_0 when H_0 is true. But the econometrician is in error when she fails to reject H_0 when H_0 is false (**Type II Error**) and when she rejects H_0 when H_0 is true (**Type I Error**).

The Neyman-Pearson criterion for hypothesis testing can now be expressed as

> **Neyman-Pearson Criterion**: select the test statistic and critical region that, for a given upper limit of the probability of Type I Error (**Significance Level**), minimizes the probability of Type II Error.

In essence, the Neyman-Pearson criterion transforms the hypothesis test selection problem into a constrained minimization problem

$$\begin{aligned} \min_{TS,C} \quad & P(\text{Type II Error}) \\ \text{s.t.} \quad & P(\text{Type I Error}) \leq \alpha \end{aligned}$$

(TS, C) represent the test statistic and critical region. α is the "significance level" – the chosen upper limit of the probability of Type I error.

We reject H_0 when $TS \in C$. So we can re-write the above as

$$\begin{aligned} \min_{TS,C} \quad & P(TS \notin C \mid H_0 \text{ False}) \\ \text{s.t.} \quad & P(TS \in C \mid H_0 \text{ True}) \leq \alpha \end{aligned}$$

$P(TS \notin C \mid H_0 \text{ False})$ is the probability that the value of the test statistic falls outside the critical region given that H_0 is false ("under H_a"). Similarly, $P(TS \in C \mid H_0 \text{ True})$ is the probability that the value of the test statistic falls within the critical region given that H_0 is true ("under H_0").

We can further re-write the above by using the **power function**, PF. Consider the special case where we examine test statistics based on the estimators of the DGP parameters: $TS(\hat{\theta})$.

$$PF(\theta) \quad = \quad P(TS(\hat{\theta}) \in C \mid \theta)$$

Here, the power function $PF(\theta_0)$ is the probability that we reject the null hypothesis using $TS(\hat{\theta})$ given the real DGP parameter values θ. Note that the power function is a function of the underlying DGP parameter values.

Let θ_0 be the value of θ that satisfies the null hypothesis $H_0 : h(\theta) - r = 0$. For an invertible $h(\cdot)$, $\theta_0 = h^{-1}(r)$. Then

$$P(TS(\hat{\theta}) \in C \mid H_0 \text{ True}) \quad = \quad \underbrace{P(TS(\hat{\theta}) \in C \mid \theta = \theta_0)}_{PF(\theta_0)}$$

Trying to re-express $P(TS(\hat{\theta}) \notin C \mid H_0 \text{ False})$ in terms of the power function illustrates a complication. While simple hypotheses (like the H_0 considered above) usually map to a single value of θ, composite hypotheses do not. But the conditional and unconditional probability of $TS \in C$ (and the complement $TS \notin C$) depends on the precise value of θ. So $P(TS(\hat{\theta}) \notin C \mid H_0 \text{ False})$ is not well defined when the alternative hypothesis is composite.[6]

Therefore, we replace the idea of minimizing $P(TS(\hat{\theta}) \notin C \mid H_0 \text{ False})$ with the idea that the chosen (TS, C) has a lower $P(TS(\hat{\theta}) \notin C \mid \theta = \theta_a)$ for any θ_a that satisfies the alternative hypothesis. This modified concept is known as **"uniformly most powerful" (UMP)**. Not all hypotheses have an associated UMP test. But, most of the popular test statistics are of the UMP variety.

[6]One might be tempted to "integrate" across the different θ_a that satisfy the alternative hypothesis; this requires some "distribution" of θ itself. This is disallowed in classical/frequentist statistics (but is allowed in Bayesian).

Note that $P(TS(\hat{\theta}) \notin C \mid \theta = \theta_a) \ = \ 1 - PF(\theta_a)$. We can re-write the Neyman-Pearson Criterion in terms of the PF and UMP:

Neyman-Pearson UMP Test: the test, (TS^*, C^*), of H_0 and H_a such that

- $PF(\theta_0) \leq \alpha$ for θ_0 satisfying H_0
- $PF(\theta_a)$ is greater at all θ_a that satisfies H_a for (TS^*, C^*) than any other (TS, C) for which $PF(\theta_0) \leq \alpha$

In short, we want to choose (TS, C) that has a low value of the power function ("low power") when $\theta = \theta_0$ but a high value ("high power") for all other values of θ. Remember, we want to reject the null hypothesis as much as possible when $\theta \neq \theta_0$. In general, the probability of rejecting H_0 when H_0 is false is known as the **power** of the test.

Technically, the significance level is the upper limit of the probability of the Type I error. But in practice, this upper limit is usually binding. Reducing the probability of Type II error requires us to incur greater probability of Type I error. Therefore, for most applications, $\alpha = PF(\theta_0)$.[7]

4.1 Normally Distributed Test Statistics

Consider the following test statistic for $H_0 : h(\theta) - r = 0$ that is distributed Normal for any θ

$$h(\hat{\theta}) - r \quad \sim \quad N(h(\theta) - r, V_h) \qquad \text{where } V_h \text{ is the variance of } h(\hat{\theta}) - r$$

We can construct the z-statistic

$$\frac{h(\hat{\theta}) - r}{\sqrt{V_h}} \quad \sim \quad N(h(\theta) - r, 1)$$

Under H_0, the z-statistic is distributed standard Normal. Under H_a, the z-statistic is distributed Normal but with a non-zero mean.

So to evaluate the probability of a Type I Error associated with this z-statistic, we use the standard Normal distribution and to evaluate the probability of a Type II Error, we use the Normal distribution with mean $r - h(\theta)$ and variance 1.

There are three main types of composite alternative hypotheses used in economics applications

- $H_a : h(\theta) - r > 0$ (rejecting H_0 implies support for greater values)
- $H_a : h(\theta) - r < 0$ (rejecting H_0 implies support for lesser values)
- $H_a : h(\theta) - r \neq 0$ (rejecting H_0 implies support for both greater and lesser)

The first two are associated with **"one-sided"** hypothesis tests and the third with **"two-sided"** hypothesis tests – for reasons that are self-evident. For the null hypothesis $H_0 : h(\theta) - r = 0$ and the above alternative hypotheses, we can derive the UMP test based on the above z-statistic

[7] This is analogous to the budget constraint in consumer theory. We maximize utility conditional on the budget constraint. Technically, the budget constraint indicates that we cannot consume *up to* our income. In practice, the budget constraint is interpreted as consumer expenditure equaling income. Higher consumption, higher utility.

- For $H_a : h(\theta) - r > 0$, the UMP test is the above z-statistic and a critical region of $\{\text{z-statistic} > c_\alpha\}$
- For $H_a : h(\theta) - r < 0$, the UMP test is the above z-statistic and a critical region of $\{\text{z-statistic} < -c_\alpha\}$
- For $H_a : h(\theta) - r \neq 0$, the UMP test is the above z-statistic and a critical region of $\{\text{z-statistic} < -c_{0.5\alpha}\} \cup \{\text{z-statistic} > c_{0.5\alpha}\}$

where C_α is the value such that $P(\text{z-statistic} > C_\alpha \mid H_0 \text{ True}) = \alpha$. Similarly, $C_{0.5\alpha}$ is the value such that $P(\text{z-statistic} > C_{0.5\alpha} \mid H_0 \text{ True}) = \frac{1}{2}\alpha$.

The above z-statistic is distributed standard Normal when H_0 is true. This implies that the distribution of the z-statistic, under H_0, is symmetric around 0. Therefore

$$\begin{aligned} P(\text{z-statistic} > C_\alpha \mid H_0 \text{ True}) &= P(\text{z-statistic} < -C_\alpha \mid H_0 \text{ True}) \\ P(\text{z-statistic} > C_{0.5\alpha} \mid H_0 \text{ True}) &= P(\text{z-statistic} < -C_{0.5\alpha} \mid H_0 \text{ True}) \end{aligned}$$

We used this result when we expressed the UMP critical regions above. The above result also indicates that we can solve for the threshold values of the critical region $(C_\alpha, C_{0.5\alpha})$ using the cumulative distribution function (cdf) of the standard Normal.

$$\begin{aligned} P(z < 1.28 \mid H_0) = F(1.28) = 0.9 \quad &\Longrightarrow \quad C_{0.1} = 1.28 \\ P(z < 1.65 \mid H_0) = F(1.65) = 0.95 \quad &\Longrightarrow \quad C_{0.05} = 1.65 \\ P(z < 1.96 \mid H_0) = F(1.96) = 0.975 \quad &\Longrightarrow \quad C_{0.025} = 1.96 \\ P(z < 2.31 \mid H_0) = F(2.31) = 0.99 \quad &\Longrightarrow \quad C_{0.01} = 2.31 \end{aligned}$$

These values can be obtained from tables at the back of many econometrics textbooks (e.g. Goldberger) or from commands within econometrics software (e.g. STATA).

Example: Two-sided Test for Constant Returns in Cobb-Douglas (known σ^2)
$H_0 : \beta_1 + \beta_2 - 1 = 0$ versus $H_a : \beta_1 + \beta_2 - 1 \neq 0$

1. Obtain b^{ols} using the sample $\{\ ln(Y_i), ln(K_i), ln(L_i)\ \}_{i=1...N}$
2. Calculate $Rb^{ols} - r$ and $\text{Var}(Rb^{ols} - r \mid ln(K), ln(L)) = \sigma^2 RQ^{-1}R'$
   Note: $R = (\ 0\ \ 1\ \ 1\ )$ and $r = 1$ and $X = (\ \iota\ \ ln(K)\ \ ln(L)\ )$ (Nx3 matrix) and $Q = X'X$ $(3 \times 3$ matrix)
3. Calculate the z-statistic: $z = \frac{Rb^{ols} - r}{\sqrt{\sigma^2 RQ^{-1}R'}}$
4. Choose the significance level α (usually some value less than 0.1)
5. Find the critical region associated with that significance level (use cdf of N(0,1))
   $\{z < -C_{0.5\alpha}\} \cup \{z > C_{0.5\alpha}\}$
6. If z-statistic is less than $-C_{0.5\alpha}$ or greater than $C_{0.5\alpha}$, reject H_0.
   Otherwise, fail to reject H_0

Note: the last two steps are what differ for a one-sided test

4.2 Non-Normally Distributed Test Statistics

For test statistics that are not distributed Normally, deriving the UMP test can be trickier.

If the distribution of the test statistic is sufficiently similar to the Normal distribution, the UMP test is analogous. Specifically, if the distribution of the test-statistic is uni-modal and symmetric around the zero mean then the steps involved in performing the UMP test is near-identical.

- A distribution that is uni-modal around the zero mean resembles a single hill, with the apex at the zero mean: the value of the pdf (pmf) is monotonically increasing from $-\infty$ to 0 and monotonically decreasing from 0 to $+\infty$.
- A distribution that is symmetric around the zero mean is one where $f(x) = f(-x)$ for any possible x – i.e. the pdf value is the same for any real value and its negative counterpart.

One distribution we have encountered that is both uni-modal and symmetric around the zero mean is the **t-distribution**.[8]

Consequently, the UMP test involving the t-statistic is similar to that of the z-statistic. The only difference is that the threshold values of the critical region, C_α or $C_{0.5\alpha}$, are obtained from the cdf of the t-distribution and not the cdf of the standard Normal.

Example: One-sided Test for Constant Returns in Cobb-Douglas (unknown σ^2)
$H_0 : \beta_1 + \beta_2 - 1 = 0$ versus $H_a : \beta_1 + \beta_2 - 1 > 0$ (constant or increasing returns only)

1. Obtain b^{ols} using the sample $\{ \ ln(Y_i), ln(K_i), ln(L_i) \ \}_{i=1...N}$
2. Calculate $Rb^{ols} - r$ and $\text{Var}(Rb^{ols} - r \mid ln(K), ln(L)) = \sigma^2 RQ^{-1}R'$
   Note: $R = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}$ and $r = 1$ and $X = \begin{pmatrix} \iota & ln(K) & ln(L) \end{pmatrix}$ (Nx3 matrix) and $Q = X'X$ $(3 \times 3$ matrix)
3. Calculate $s^2 = \frac{1}{N-3} \sum_{i=1}^{N} \left(\ ln(Y_i) - b_0^{ols} - b_1^{ols} ln(K_i) - b_2^{ols} ln(L_i) \ \right)^2$
4. Calculate the t-statistic: $t = \frac{Rb^{ols} - r}{\sqrt{s^2 RQ^{-1}R'}}$
5. Choose the significance level α (usually some value less than 0.1)
6. Find the critical region associated with that significance level (use cdf of t-distribution)
   $\{t > C_\alpha\}$
7. If t-statistic is greater than C_α, reject H_0. Otherwise, fail to reject H_0

Note: the last two steps are what differ for a two-sided test (or the other one-sided test, where allowed alternative is decreasing returns)

We discuss the steps for the UMP test involving other non-Normally distributed test statistics later. the steps for some of these test statistics – such as those distributed chi-squared – are similar to those above but depart in some key manners.

[8] Food for thought: one can show that a distribution that is both uni-modal and symmetric around the same value can only be so around the mean of the distribution.