ECON 61001: *Hypothesis Testing*

Alastair R. Hall

The University of Manchester

- asset returns: $R - R_f = \beta_0(R_m - R_f) + error$

  - $\beta_0 < 0 \Rightarrow$ stock is inversely related to market index;
  - $\beta_0 = 1 \Rightarrow$ stock moves in line with market index.

- returns to education (dropping exp to simplify and reparameterizing):

$$ln(w) = \beta_{0,1} + \beta_{0,2} * ed + \beta_{0,3} * D + \beta_{0,4} * (D * ed) + error$$

  - $D = 1$ is female and zero else
  - $\beta_{0,3} = 0$, $\beta_{0,4} = 0 \Rightarrow$ no difference between men and women.

- aggregate production function:

$$lnQ = \beta_{0,1} + \beta_{0,2} * ln(L) + \beta_{0,3} * ln(K) + error$$

- $\beta_{0,2} + \beta_{0,3} \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 1 \Rightarrow \left\{ \begin{array}{c} \text{diminishing} \\ \text{constant} \\ \text{increasing} \end{array} \right\}$ returns to scale

- A statistical hypothesis is a conjecture about the distribution of one or more random variables.

- The classical theory of hypothesis testing provides a framework for deciding whether a particular hypothesis is correct.

- Within this framework, there are only two possible decisions: the hypothesis is true or it is not. A decision procedure for such a problem is called a test.

Assume: our hypothesis involves θ, the parameter vector indexing distribution of V, and Θ denotes the parameter space with $\Theta \subset \mathbb{R}^p$.

Divide Θ into two mutually exclusive and exhaustive parts:

$$\begin{aligned} \Theta_0 &= \{\theta : \text{ such that the hypothesis is true}\}, \\ \Theta_1 &= \{\theta : \text{ such that the hypothesis is false}\} \end{aligned}$$

Using this partition, we can state the object as being to test the null hypothesis

$$H_0 : \theta \in \Theta_0$$

against the alternative hypothesis,

$$H_1 : \theta \in \Theta_1.$$

To facilitate the choice between H_0 and H_1, collect sample of T observations.

Base inference on some function of this sample, known as a test statistic; denoted by S_T.

In a test procedure, divide sample space of S_T into two mutually exclusive and exhaustive regions, R_0 and R_1, such that

$$S_T \in R_0 \quad \Rightarrow \quad H_0 \text{ is accepted or rather not rejected}$$

$$S_T \in R_1 \quad \Rightarrow \quad H_0 \text{ is rejected in favour of } H_1$$

where

- R_0 is known as the acceptance region.
- R_1 is known as the rejection region or the *critical region*.

Choice of R_0 and R_1 based on possible outcomes of test.

Decision may be correct, but can also make error:

- a Type I error, in which H_0 is rejected when it is true;
- a Type II error, in which H_0 is not rejected when its false.

Ideally we would make both errors as small as possible but there is a tension between them.

So limit P(Type I error) to be no larger than: 0.1, 0.05 or 0.01.

To do this, we need to know the distribution of the test statistic under H_0.

Size and significance level of a test

Let $P_\theta(\,\cdot\,)$ denote the probability of the event in parentheses if the parameter vector takes the value θ.

Define $\alpha(\theta) = P_\theta(R_1)$ that is, $\alpha(\theta)$ describes the probability of a type I error for values of θ that satisfy H_0.

The quantity $\sup_{\theta \in \Theta_0} \alpha(\theta)$ is known as the size of the test, and equals the maximal probability of a type I error.

To implement test, specify α such that $\alpha(\theta) \leq \alpha$ for all $\theta \in \Theta_0$.

- α is an upper bound on the probability of a type one error.
- $100\alpha\%$ is known as the significance level of the test.
- In general, the size and α coincide (but they need not.)

Suppose that:

- $v_t \sim IIN(\theta, \sigma^2)$, $t = 1, 2, \ldots T$; assume σ^2 known.
- wish to test $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$.

- can show $\bar{v}_T \sim N(\theta, \sigma^2/T)$, where $\bar{v}_T = T^{-1}\sum_{t=1}^{T} v_t$.

$\rightarrow$ test statistic is:

$$\tau_T = \frac{\bar{v}_T}{\sqrt{\sigma^2/T}}$$

Under H_0, $\tau_T \sim N(0, 1)$.

Reject H_0 if τ_T is sufficiently far away from zero.

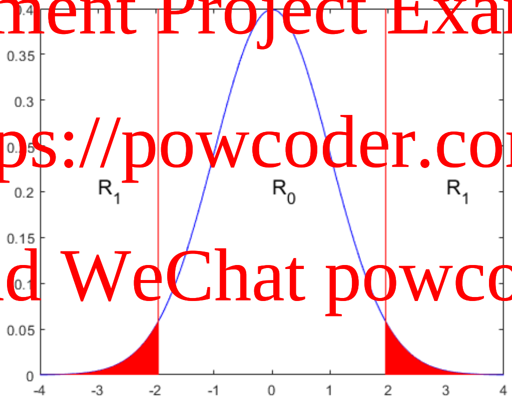
So the decision rule takes form: reject H_0 if $|\tau_T| > c$ for some constant c.

Choose c to control P(Type I error).

- $\Theta_0 = \{0\}$ and so $sup_{\theta \in \Theta_0} \alpha(\theta) = P(|\tau_T| > c \,\big|\, \theta = 0)$.
- $\Rightarrow c = z_{1-\alpha/2}$

$\Rightarrow$ (for example): reject $H_0 : \theta = 0$ in favour of $H_1 : \theta \neq 0$ at the 5% significance level if $|\tau_T| > 1.96$.

Example 2.8 in Lecture Notes: Acceptance and rejection regions



Observed significance level *aka* the p-value

The observed significance level or p-value is the significance level for which the test statistic lies on the boundary of the acceptance and rejection regions.

In our example: p-value $= \delta$ such that $|\tau_T| = z_{1-\delta/2}$.

Interpretation of p-value: we reject H_0 at all significance levels $100\alpha\%$ for which $\alpha > p - value$.

Further discussion of the topics in this podcast can be found in Section 2.3.1 of the Lecture Notes.

Same source discusses application of framework to hypotheses of the form:

- $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$
- $H_0 : \theta \geq 0$ versus $H_1 : \theta < 0$.