Hypothesis Testing

Fall 2008

Hypothesis Testing

- Examples:
  - Does the increase of Co2 concentration increase the average temperature?
  - Is the elasticity of housing prices to nitrogen oxide equal to one?
  - Are non-whites (or females) discriminated against in hiring?
- Devising methods for answering such questions, using a sample of data, is known as *hypothesis testing*.

Hypothesis I

- A hypothesis takes the form of a statement of the true value for a coefficient or for an expression involving the coefficient.
  - The hypothesis to be tested is called the *null hypothesis*, H_0.
  - The hypothesis against which the null is tested is called the *alternative hypothesis*, H_A.
- Example: Consider the following regression model:

$$\ln Hprice_i = \beta_0 + \beta_1 \ln Nox_i + \beta_2 rooms_i + \beta_3 stratio_i + \beta_4 \ln dist_i + u_i$$

$$\begin{aligned} H_0 &: \beta_1 = 1, \\ H_A &: \beta_1 \neq 1 \end{aligned}$$

- Rejecting the null hypothesis *does not imply accepting the alternative*.

Hypothesis II

- In hypothesis testing, we can make two kinds of mistakes.

Hypothesis II

- In hypothesis testing, we can make two kinds of mistakes.
- A **Type I Error** is rejecting H_0 when it is true. The probability of a Type I error is called the *significance level*, usually denoted by α.

Hypothesis II

- In hypothesis testing, we can make two kinds of mistakes.
- A **Type I Error** is rejecting H_0 when it is true. The probability of a Type I error is called the *significance level*, usually denoted by α.
- Classical hypothesis testing requires that we initially specify a signficance level for a test, usually $\alpha = 0.10$, 0.05, and 0.01.

Hypothesis II

- In hypothesis testing, we can make two kinds of mistakes.
- A **Type I Error** is rejecting H_0 when it is true. The probability of a Type I error is called the *significance level*, usually denoted by α.
- Classical hypothesis testing requires that we initially specify a signficance level for a test, usually $\alpha = 0.10$, 0.05, and 0.01.
- A **Type II Error** is failing to reject H_0 when it is false. The *power of a test* is just one minus the probability of a Type II error.

Hypothesis II

- In hypothesis testing, we can make two kinds of mistakes.
- A **Type I Error** is rejecting H_0 when it is true. The probability of a Type I error is called the *significance level*, usually denoted by α.
- Classical hypothesis testing requires that we initially specify a signficance level for a test, usually $\alpha = 0.10$, 0.05, and 0.01.
- A **Type II Error** is failing to reject H_0 when it is false. The *power of a test* is just one minus the probability of a Type II error.
- Once we have chosen the significance level, we would like to maximize the power of a test against all relevant alternatives.

Hypothesis II

- In hypothesis testing, we can make two kinds of mistakes.
- A **Type I Error** is rejecting H_0 when it is true. The probability of a Type I error is called the *significance level*, usually denoted by α.
- Classical hypothesis testing requires that we initially specify a signficance level for a test, usually $\alpha = 0.10$, 0.05, and 0.01.
- A **Type II Error** is failing to reject H_0 when it is false. The *power of a test* is just one minus the probability of a Type II error.
- Once we have chosen the significance level, we would like to maximize the power of a test against all relevant alternatives.
- In order to test a null hypothesis against an alternative, we need to choose a test statistic and a critical value.

Testing hypothesis about a single population parameter

- Consider the following multiple regression model:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik} + u_i.$$

- We wish to test the hypothesis that $\beta_j = b$ where b is some known value (e.g., zero) against the alternative that β_j is not equal to b:

$$H_0 : \beta_j = b$$

$$H_A : \beta_j \neq b$$

- To test the null hypothesis, we need to know how the OLS estimator β_j is distributed.

Normality Assumption

- **Assumption (Normality)**: u_i is independent of $X_1, ..., X_k$ and all other u_j ,and is normally distributed with mean zero and variance σ^2:

$$u_i \sim iid\ N\left(0, \sigma^2\right)$$

Normality Assumption

- **Assumption (Normality)**: u_i is independent of $X_1, ..., X_k$ and all other u_j ,and is normally distributed with mean zero and variance σ^2:

$$u_i \sim iid\ N\left(0, \sigma^2\right)$$

- Now, note that

$$\widehat{\beta}_j = \beta_j + \sum_{i=1}^{N} \omega_i u_i,$$

where

$$\omega_i = \widehat{R}_{ij} / \sum_{s=1}^{N} \widehat{R}_{sj}^2.$$

Normality Assumption

- **Assumption (Normality)**: u_i is independent of $X_1, ..., X_k$ and all other u_j ,and is normally distributed with mean zero and variance σ^2:
$$u_i \sim iid \ N\left(0, \sigma^2\right)$$
- Now, note that
$$\widehat{\beta}_j = \beta_j + \sum_{i=1}^{N} \omega_i u_i,$$
where
$$\omega_i = \widehat{R}_{ij} / \sum_{s=1}^{N} \widehat{R}_{sj}^2.$$
- Then, we can show
$$\widehat{\beta}_j \sim N\left(\beta_j, Var\left(\widehat{\beta}_j\right)\right),$$
where
$$Var\left(\widehat{\beta}_j\right) = \sigma^2 / \sum_{s=1}^{N} \widehat{R}_{sj}^2.$$

Test Statistic

- Naturally, a test statistic can be constructed in the following way: under the null hypothesis ($H_0 : \beta_j = b$),

$$z = \left(\widehat{\beta}_j - b\right) / \sqrt{Var\left(\widehat{\beta}_j\right)} \sim N(0,1)$$

Test Statistic

- Naturally, a test statistic can be constructed in the following way: under the null hypothesis ($H_0 : \beta_j = b$),

$$z = \left(\widehat{\beta}_j - b\right) / \sqrt{Var\left(\widehat{\beta}_j\right)} \sim N(0,1)$$

- The difficulty in using this result is that we do not know $Var\left(\widehat{\beta}_j\right)$ since we do not know σ^2, which needs to be estimated.

Test Statistic

- Naturally, a test statistic can be constructed in the following way: under the null hypothesis ($H_0 : \beta_j = b$),

$$z = \left(\widehat{\beta}_j - b\right) / \sqrt{Var\left(\widehat{\beta}_j\right)} \sim N(0,1)$$

- The difficulty in using this result is that we do not know $Var\left(\widehat{\beta}_j\right)$ since we do not know σ^2, which needs to be estimated.
- Using the unbiased estimator of σ^2, $\widehat{\sigma}^2$, we can constuct the alternative test statistic:

$$z^* = \left(\widehat{\beta}_j - b\right) / \sqrt{\widehat{Var\left(\widehat{\beta}_j\right)}}.$$

Test Statistic

- Naturally, a test statistic can be constructed in the following way: under the null hypothesis ($H_0 : \beta_j = b$),

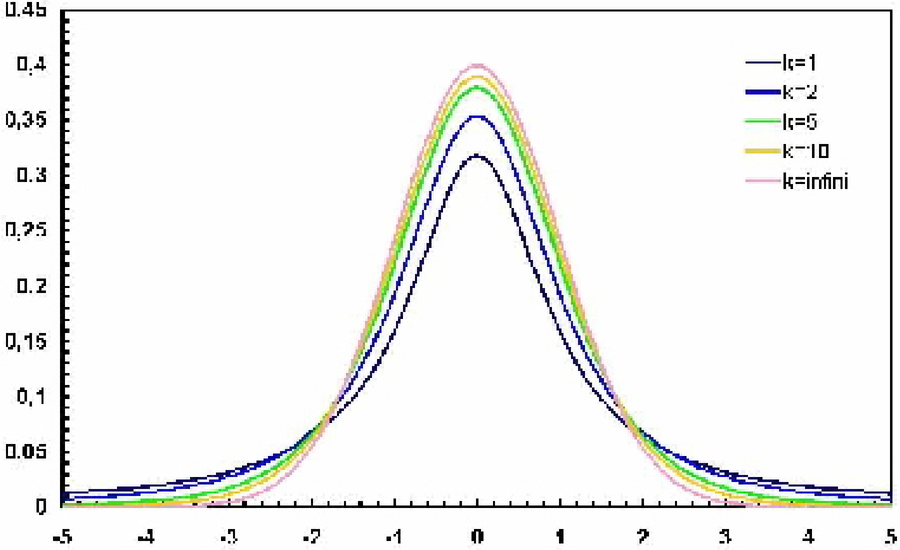
$$z = \left(\widehat{\beta}_j - b\right) / \sqrt{Var\left(\widehat{\beta}_j\right)} \;\sim\; N\left(0, 1\right)$$

- The difficulty in using this result is that we do not know $Var\left(\widehat{\beta}_j\right)$ since we do not know σ^2, which needs to be estimated.
- Using the unbiased estimator of σ^2, $\widehat{\sigma}^2$, we can constuct the alternative test statistic:

$$z^* = \left(\widehat{\beta}_j - b\right) / \sqrt{\widehat{Var\left(\widehat{\beta}_j\right)}}.$$

- This test statistic is no longer Normally distributed, but follows the t distribution with $N - (k+1)$ degrees of freedom.

The Student's t Distribution



Classical Approach of testing the Hypothesis

- First, we choose the size of the test (significance level). The conventional size is 5%, $\alpha = 0.05$.

Classical Approach of testing the Hypothesis

- First, we choose the size of the test (significance level). The conventional size is 5%, $\alpha = 0.05$.
- Then, we find the two critical values (since it is a two-tailed test), $t_{\frac{\alpha}{2},N-(k+1)}$ and $t_{1-\frac{\alpha}{2},N-(k+1)}$, using the table of t distribution.

Classical Approach of testing the Hypothesis

- First, we choose the size of the test (significance level). The conventional size is 5%, $\alpha = 0.05$.
- Then, we find the two critical values (since it is a two-tailed test), $t_{\frac{\alpha}{2},N-(k+1)}$ and $t_{1-\frac{\alpha}{2},N-(k+1)}$, using the table of t distribution.
  - We accept the null hypothesis if the test statistic is between the two critical values corresponding to our chosen size.

Classical Approach of testing the Hypothesis

- First, we choose the size of the test (significance level). The conventional size is 5%, $\alpha = 0.05$.
- Then, we find the two critical values (since it is a two-tailed test), $t_{\frac{\alpha}{2}, N-(k+1)}$ and $t_{1-\frac{\alpha}{2}, N-(k+1)}$, using the table of t distribution.
  - We accept the null hypothesis if the test statistic is between the two critical values corresponding to our chosen size.
  - Otherwise we reject the null hypothesis.

Classical Approach of testing the Hypothesis

- First, we choose the size of the test (significance level). The conventional size is 5%, $\alpha = 0.05$.
- Then, we find the two critical values (since it is a two-tailed test), $t_{\frac{\alpha}{2}, N-(k+1)}$ and $t_{1-\frac{\alpha}{2}, N-(k+1)}$, using the table of t distribution.
  - We accept the null hypothesis if the test statistic is between the two critical values corresponding to our chosen size.
  - Otherwise we reject the null hypothesis.
- The logic of hypothesis testing is that if the null hypothesis is true, then the statistic will lie within the two critical values with $100 \times (1 - \alpha)\,\%$ of the time of random samples.

t Distribution

Degrees of freedom	.005 (one tail) .01 (two tails)	.01 (one tail) .02 (two tails)	.025 (one tail) .05 (two tails)	.05 (one tail) .10 (two tails)	.10 (one tail) .20 (two tails)	.25 (one tail) .50 (two tails)
	α					
1	63.657	31.821	12.706	6.314	3.078	1.000
2	9.925	6.965	4.303	2.920	1.886	.816
3	5.841	4.541	3.182	2.353	1.638	765
4	4.604	3.747	2.776	2.132	1.533	741
5	4.032	3.365	2.571	2.015	1.476	727
6	3.707	3.143	2.447	1.943	1.440	.718
7	3.500	2.998	2.365	1.895	1.415	.711
8	3.355	2.896	2.306	1.860	1.397	.706
9	3.250	2.821	2.262	1.833	1.383	.703
10	3.169	2.764	2.228	1.812	1.372	.700
11	3.106	2.718	2.201	1.796	1.363	.697
12	3.054	2.681	2.179	1.782	1.356	.696
13	3.012	2.650	2.160	1.771	1.350	.694
14	2.977	2.625	2.145	1.761	1.345	.692
15	2.947	2.602	2.132	1.753	1.341	.691
16	2.921	2.584	2.120	1.746	1.337	.690
17	2.898	2.567	2.110	1.740	1.333	.689
18	2.878	2.552	2.101	1.734	1.330	.688
19	2.861	2.540	2.093	1.729	1.328	.688
20	2.845	2.528	2.086	1.725	1.325	.687
21	2.831	2.518	2.080	1.721	1.323	.686
22	2.819	2.508	2.074	1.717	1.321	.686
23	2.807	2.500	2.069	1.714	1.320	.685
24	2.797	2.492	2.064	1.711	1.318	.685
25	2.787	2.485	2.060	1.708	1.316	.684
26	2.779	2.479	2.056	1.706	1.315	.684
27	2.771	2.473	2.052	1.703	1.314	.684
28	2.763	2.467	2.048	1.701	1.313	.683
29	2.756	2.462	2.045	1.699	1.311	.683
Large (z)	2.575	2.327	1.960	1.645	1.282	.675

p-Value and Confidence Interval

- Since there is no "correct" significance level, it may be more informative to report *the* **smallest** *significance level at which the null would be rejected*.
- This level is known as the **p-value** for the test.
- We can also construct an interval estimate for the population parameter β_j, called *confidence interval* (CI), such that the chance that the true β_j lies within that interval is $1-\alpha$.
- That is,

$$\Pr\left(t_{\frac{\alpha}{2},N-(k+1)} < z^* < t_{1-\frac{\alpha}{2},N-(k+1)}\right) = 1-\alpha.$$

- With some manipulation,

$$\Pr\left(\widehat{\beta}_j - s.e.\left(\widehat{\beta}_j\right) \times t < \beta_j < \widehat{\beta}_j + s.e.\left(\widehat{\beta}_j\right) \times t\right) = 1-\alpha,$$

  where $t = t_{1-\frac{\alpha}{2},N-(k+1)}$.
- The term in the bracket is the confidence interval for β_j.

Example: Housing Prices

- The Stata program, by default, reports the *p*-values, *t* statistics and confidence intervals under the null that each population parameter is zero.

Var	Coeff.	s.e.	t value	p-value	Conf. Int.	
lnox	-0.95	0.12	-8.17	0.000	-1.18	-0.72
ldist	-0.13	0.04	-3.12	0.002	-0.22	-0.05
rooms	0.25	0.02	13.74	0.000	0.22	0.29
stratio	-0.05	0.01	-8.89	0.000	-0.06	-0.04
const	11.08	0.32	34.84	0.000	10.46	11.71

Example: Housing Prices

- We want to test the following hypothesis: the elasticity of housing prices to the amount of nitrogen oxide is equal to 1:

$$H_0 : \beta_1 = -1, \; H_A : \beta_1 \neq -1.$$

- We have 506 observations and so 501 degrees of freedom. At 95% confidence interval, $t_{0.025,501} = 1.96$. Then,

$$\begin{aligned} & \Pr\left(-0.95 - 0.12 \times 1.96 < \beta_1 < -0.95 + 0.12 \times 1.96\right) \\ = \; & \Pr\left(-1.1852 < \beta_1 < -0.7148\right) = 0.95. \end{aligned}$$

- The true value β_1 has 95% chance of being in $[-1.1852, -0.7148]$.
- Alternatively,

$$z^* = \frac{-0.95 - (-1)}{0.12} = 0.417.$$

- Since the critical value is 1.96 at the 5% significance level. Thus, we cannot reject the null hypothesis.

More on Testing

- Do we need the assumption of normality of the error term to carry out inference (hypothesis testing)?
- Under normality our test is exact in a sense that the test statistic exactly follows the t distribution.
- Without the normality, we can still carry out hypothesis testing, relying on asymptotic approximations when we have large enough samples.
- To do this, we need to use the Central Limit Theorem: under regular conditions,

$$z^* = \frac{\left(\widehat{\beta}_j - b\right)}{\sqrt{\widehat{Var\left(\widehat{\beta}_j\right)}}} \sim^a N(0,1) \quad \text{as } N \to \infty$$

Testing Multiple Restrictions

- Now suppose we wish to test multiple hypotheses about the underlying parameters.
- A test of multiple restrictions on the parameters is called a *joint hypotheses test*.
- In this case we will use the so called "F-test".
- Examples:
  - (Exclusion restrictions) $H_0 : \beta_1 = 0, \beta_2 = 0$ and $\beta_3 = 0$; $H_A : H_0$ is not ture.
  - $H_0 : \beta_1 = 0, \beta_2 = \beta_3$; $H_A : H_0$ is not ture.

The Unrestricted and Restricted Regression Models

- The **Unrestricted Model** is the model without any of the restrictions imposed from the null hypothesis.
- The **Restricted Model** is the model on which the restrictions have been imposed.
- Example 1: $H_0 : \beta_1 = 0, \beta_2 = 0$ and $\beta_3 = 0$

$$\begin{cases} Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + u_i & \text{Unrestricted} \\ Y_i = \beta_0 + \beta_4 X_{i4} + u_i & \text{Restricted} \end{cases}$$

- Example 2: $H_0 : \beta_1 = 0, \beta_2 = \beta_3$

$$\begin{cases} Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + u_i & \text{Unrestricted} \\ Y_i = \beta_0 + \beta_2 (X_{i2} + X_{i3}) + \beta_4 X_{i4} + u_i & \text{Restricted} \end{cases}$$

Heuristic Illustration of the Test

- Inference will be based on comparing the fit of the restricted and unrestricted regression.
- Note that the unrestricted regression will always fit *at least as as well as* the restricted one (why?).
- So the question will be how much improvement in the fit we get by relaxing the restrictions relative to the loss of precision that follows.
- The distribution of the test statistic will give us a measure of this so that we can construct a statistical decision rule.

Definitions

- Define the *Unrestricted Sum of Squared Residual (USSR)* as the residual sum of squares obtained from estimating the unrestricted model.
- Define the *Restricted Sum of Squared Residual (RSSR)* as the residual sum of squares obtained from estimating the restricted model.
- Note that $RRSS \geq URSS$ (why?).
- Denote q as the number of restrictions mposed.

The F-Statistic

- The statistic for testing multiple restrictions we discussed is

$$F = \frac{(RSSR - USSR)/q}{USSR/(N-k-1)} = \frac{\left(R_{UR}^2 - R_R^2\right)/q}{\left(1 - R_{UR}^2\right)/(N-k-1)} \sim F(q, N-k-1)$$

The F-Statistic

- The statistic for testing multiple restrictions we discussed is

$$\begin{aligned} F &= \frac{(RSSR - USSR)/q}{USSR/(N-k-1)} \\ &= \frac{\left(R^2_{UR} - R^2_R\right)/q}{\left(1 - R^2_{UR}\right)/(N-k-1)} \sim F(q, N-k-1) \end{aligned}$$

- Under the normality assumption of errors, the F statistic exactly follows the F distribution with degrees of freedom $(q, N-k-1)$.

The F-Statistic

- The statistic for testing multiple restrictions we discussed is

$$
\begin{aligned}
F &= \frac{(RSSR - USSR)/q}{USSR/(N-k-1)} \\
&= \frac{\left(R_{UR}^2 - R_R^2\right)/q}{\left(1 - R_{UR}^2\right)/(N-k-1)} \sim F(q, N-k-1)
\end{aligned}
$$

- Under the normality assumption of errors, the F statistic exactly follows the F distribution with degrees of freedom $(q, N-k-1)$.
- The test statistic is always non-negative. If the null is true, we would expect this to be "small".

The F-Statistic

- The statistic for testing multiple restrictions we discussed is

$$F = \frac{(RSSR - USSR)/q}{USSR/(N-k-1)} = \frac{\left(R_{UR}^2 - R_R^2\right)/q}{\left(1 - R_{UR}^2\right)/(N-k-1)} \sim F(q, N-k-1)$$

- Under the normality assumption of errors, the F statistic exactly follows the F distribution with degrees of freedom $(q, N-k-1)$.
- The test statistic is always non-negative. If the null is true, we would expect this to be "small".
- The smaller the F-statistic is, the less the loss of fit due to the restrictions is.

The F-Statistic

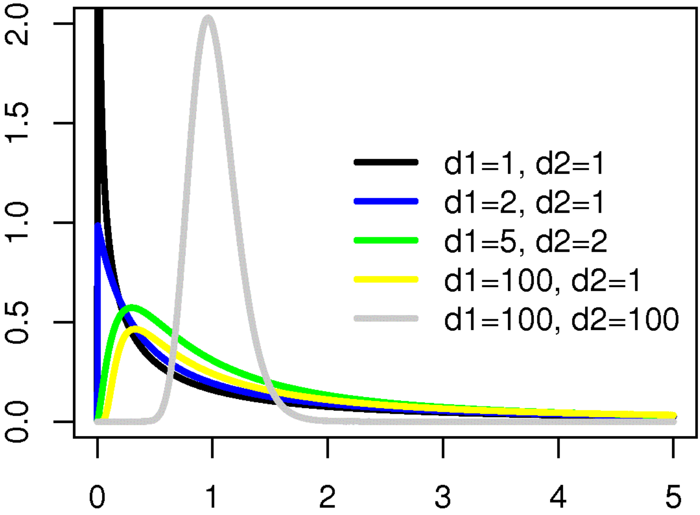
- The statistic for testing multiple restrictions we discussed is

$$
\begin{aligned}
F &= \frac{(RSSR - USSR)/q}{USSR/(N-k-1)} \\
&= \frac{\left(R_{UR}^2 - R_R^2\right)/q}{\left(1 - R_{UR}^2\right)/(N-k-1)} \sim F(q, N-k-1)
\end{aligned}
$$

- Under the normality assumption of errors, the F statistic exactly follows the F distribution with degrees of freedom $(q, N-k-1)$.
- The test statistic is always non-negative. If the null is true, we would expect this to be "small".
- The smaller the F-statistic is, the less the loss of fit due to the restrictions is.
- Without the normality assumption, we can show, using the Central Limit Theorem, that

$$qF \sim^a \mathcal{X}_q^2 \text{ as } N \to \infty$$

The F Distribution



Example: Housing Prices

- Consider the following regression model:

$$\ln Hprice_i = \beta_0 + \beta_1 \ln Nox_i + \beta_2 rooms_i + \beta_3 stratio_i + \beta_4 \ln dist_i + u_i$$

$$H_0 : \beta_1 = -1, \beta_3 = \beta_4 = 0; H_A : H_0 \text{ is not true.}$$

- The restricted regression model is

$$\ln Hprice_i + \ln Nox_i = \beta_0 + \beta_2 rooms_i + u_i$$

- The F-statistic is

$$\begin{aligned} F &= \frac{\left(R^2_{UR} - R^2_R\right)/q}{\left(1 - R^2_{UR}\right)/\left(N - k - 1\right)} = \frac{\left(0.584 - 0.316\right)/3}{\left(1 - 0.584\right)/501} \\ &= 107.59 \end{aligned}$$

- Given the degrees of freedom $(3, 501)$ and 5% significance level, the critical value is 2.60. Thus, we reject the null hypothesis.

Summary

- We use the OLS estimators in the simple and multiple linear regression models.
- Key assumptions:
    - The error term is uncorrelated with independent variables.
    - The variance of error term is constant (homoskedsticity).
    - The covariance of error term is zero (no autocorrelation).
- Departures from this simple framework:
    - Heteroskedasticity;
    - Autocorrelation;
    - Simultaneity and Endogeneity;
    - Non linear models.