Hypothesis Testing

1. Specify hypothesis to be tested

$$H_0: \text{ null hypothesis versus. } H_1: \text{ alternative hypothesis}$$

2. Specify significance level of test

$$\text{level} = \Pr(\text{Reject } H_0 | H_0 \text{ is true})$$

3. Construct test statistic, T, from observed data

4. Use test statistic T to evaluate data evidence regarding H_0

$$\begin{aligned} |T| \text{ is big} \quad &\Rightarrow \quad \text{evidence against } H_0 \\ |T| \text{ is small} \quad &\Rightarrow \quad \text{evidence in favor of } H_0 \end{aligned}$$

Decide to reject H_0 at specified significance level if value of T falls in the rejection region

$$T \in \text{ rejection region } \Rightarrow \text{ reject } H_0$$

Usually the rejection region of T is determined by a critical value, cv, such that

$$\begin{aligned} |T| &> cv \Rightarrow \text{reject } H_0 \\ |T| &\leq cv \Rightarrow \text{ do not reject } H_0 \end{aligned}$$

Decision Making and Hypothesis Tests

	Reality	
Decision	H_0 is true	H_0 is false
Reject H_0	Type I error	No error
Do not reject H_0	No error	Type II error

Significance Level of Test

$$\begin{aligned} \text{level} &= \Pr(\text{Type I error}) \\ & \Pr(\text{Reject } H_0 | H_0 \text{ is true}) \end{aligned}$$

Goal: Constuct test to have a specified small significance level

$$\text{level} = 5\% \text{ or level} = 1\%$$

Power of Test

$$
\begin{aligned}
& 1 - \text{Pr}(\text{Type II error}) \\
= \; & \text{Pr}(\text{Reject } H_0 | H_0 \text{ is false})
\end{aligned}
$$

Goal: Construct test to have high power

Problem: Impossible to simultaneously have level ≈ 0 and power ≈ 1. As level $\longrightarrow 0$ power also $\longrightarrow 0$.

Hypothesis Testing in CER Model

$$
\begin{aligned}
r_{it} &= \mu_i + \epsilon_{it} \quad t = 1, \cdots, T; \quad i = 1, \cdots N \\
\epsilon_{it} &\sim \text{iid } N(0, \sigma_i^2) \\
\text{cov}(\epsilon_{it},\ \epsilon_{jt}) &= \sigma_{ij},\ \text{cor}(\epsilon_{it},\ \epsilon_{jt}) = \rho_{ij} \\
\text{cov}(\epsilon_{it},\ \epsilon_{js}) &= 0 \quad t \neq s,\ \text{for all } i,\ j
\end{aligned}
$$

- Basic significance test

$$H_0 : \mu_i = 0 \text{ vs. } H_1 : \mu_i \neq 0$$

- Test for specific value

$$H_0 : \mu_i = \mu_i^0 \text{ vs. } H_1 : \mu_i \cancel{\neq} \mu_i^0$$

- Test for sign

$$H_0 : \mu_i = 0 \text{ vs. } H_1 : \mu_i > 0 \text{ or } \mu_i < 0$$

- Test for normal distribution

$$\begin{aligned} H_0 &: r_{it} \sim \text{iid } N(\mu_i, \sigma_i^2) \\ H_1 &: r_{it} \sim \text{ not normal} \end{aligned}$$

- Test for no autocorrelation

$$\begin{aligned} H_0 &: \rho_j = \text{corr}(r_{it}, r_{i,t-j}) = 0,\ j > 1 \\ H_1 &: \rho_j = \text{corr}(r_{it}, r_{i,t-j}) \neq 0 \text{ for some } j \end{aligned}$$

- Test of constant parameters

$$\begin{aligned} H_0 &: \mu_i, \sigma_i \text{ and } \rho_i \text{ are constant over entire sample} \\ H_1 &: \mu_i\ \sigma_i \text{ or } \rho_i \text{ changes in some sub-sample} \end{aligned}$$

Definition: Chi-square random variable and distribution

Let $Z_1, \ldots, Z_q$ be iid $N(0, 1)$ random variables. Define

$$X = Z_1^2 + \cdots + Z_q^2$$

Then

$$\begin{aligned} X &\sim \chi^2(q) \\ q &= \text{degrees of freedom (d.f.)} \end{aligned}$$

Properties of $\chi^2(q)$ distribution

$$\begin{aligned} X &> 0 \\ E[X] &= q \\ \chi^2(q) &\rightarrow \text{normal as } q \rightarrow \infty \end{aligned}$$

R functions

`rchisq()`: simulate data

`dchisq()`: compute density

`pchisq()`: compute CDF

`qchisq()`: compute quantiles

Definition: Student's t random variable and distribution with q degrees of freedom

$$
\begin{aligned}
& Z \sim N(0,1),\ X \sim \chi^2(q) \\
& Z \text{ and } X \text{ are independent} \\
& T = \frac{Z}{\sqrt{X/q}} \sim t_q \\
q \ &= \ \text{degrees of freedom (d.f.)}
\end{aligned}
$$

Properties of t_q distribution:

$$
\begin{aligned}
E[T] &= 0 \\
\text{skew}(T) &= 0 \\
\text{kurt}(T) &= \frac{3q-6}{q-4},\ q > 4 \\
T &\rightarrow N(0,1) \text{ as } q \rightarrow \infty\ (q \geq 60)
\end{aligned}
$$

R functions

`rt()`: simulate data

`dt()`: compute density

`pt()`: compute CDF

`qt()`: compute quantiles

Basic Significance Test

$$H_0 : \mu_i = 0 \text{ vs. } H_1 : \mu_i \neq 0$$

1. Test statistics: t-statistics

$$t_{\mu_i=0} = \frac{\hat{\mu}_i - 0}{\widehat{\mathsf{SE}}(\hat{\mu}_i)} = \frac{\hat{\mu}_i}{\widehat{\mathsf{SE}}(\hat{\mu}_i)}$$

Intuition:

- If $|t_{\mu_i=0}| \approx 0$ then $\hat{\mu}_i \approx 0$, and $H_0 : \mu_i = 0$ should not be rejected
- If $|t_{\mu_i=0}| > 2$, then $\hat{\mu}_i$ is more than 2 values of $\widehat{\mathsf{SE}}(\hat{\mu}_i)$ away from 0. This is very unlikely if $\mu_i = 0$, so $H_0 : \mu_i = 0$ should be rejected.

Distribution of t-statistic under H_0

Under the assumptions of the CER model, and $H_0 : \mu_i = 0$

$$t_{\mu_i=0} = \frac{\hat{\mu}_i}{\widehat{\mathsf{SE}}(\hat{\mu}_i)} \sim t_{T-1}$$

where

$$\mu_i = \frac{1}{T}\sum_{t=1}^{T} r_{it},\ \widehat{\mathsf{SE}}(\hat{\mu}_i) = \frac{\hat{\sigma}_i}{\sqrt{T}},\ \hat{\sigma}_i = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_{it} - \widehat{\mu_i})^2}$$

$$t_{T-1} = \text{Student's t distribution with}$$
$$T-1 \text{ degrees of freedom (d.f.)}$$

Remarks:

- t_{T-1} is bell-shaped and symmetric about zero (like normal)

- d.f. $=$ sample size - number of estimated parameters. In CER model there is one estimated parameter, μ_i, so df $= T-1$

- Degrees of freedom determines kurtosis (tail thickness)

$$
\begin{aligned}
\text{d.f.} &= T-1 < 10,\ \text{kurt}(t_{T-1}) >> 3 \\
\text{d.f.} &= T-1 \geq 60,\ \text{kurt}(t_{T-1}) \approx 3
\end{aligned}
$$

- For $T \geq 60$, $t_{T-1} \sim N(0,1)$. Therefore, for $T \geq 60$

$$t_{\mu_i=0} = \frac{\hat{\mu}_i}{\widehat{\mathsf{SE}}(\hat{\mu}_i)} \sim N(0,1)$$

2. Set significance level and determine rejection region; e.g.,

$$\text{Pr(Type I error)} = 5\%$$

Rejection region of t-test with 5% level

$$\text{reject } H_0 : \mu_i = 0 \text{ in favor of } H_1; \mu \neq 0 \text{ if}$$
$$t_{\mu_i=0} < cv_{.025} \text{ or } t_{\mu_i=0} > cv_{.975}$$
$$\text{i.e., if } |t_{\mu_i=0}| > cv_{.975}$$

3. Decision rule:

$$\text{Reject } H_0 : \mu_i = 0 \text{ at 5\% level if}$$
$$|t_{\mu_i=0}| > cv_{.975} = q_{.975}^{t_{T-1}}$$

Useful Rule of Thumb:

If $T \geq 60$ then $cv_{.975} \approx 2$ and the decision rule is

$$\text{Reject } H_0 : \mu_i = 0 \text{ at } 5\% \text{ level if}$$
$$|t_{\mu_i=0}| > 2$$

4. P-Value of test

$$\begin{aligned}
&\text{significance level at which test is just rejected} \\
&= \Pr(|t_{T-1}| > t_{\mu_i=0}) \\
&= \Pr(t_{T-1} < -t_{\mu_i=0}) + \Pr(t_{T-1} > t_{\mu_i=0}) \\
&= 2 \cdot \Pr(t_{T-1} > |t_{\mu_i=0}|) \\
&= 2 \times (1 - \Pr(t_{T-1} \leq |t_{\mu_i=0}|))
\end{aligned}$$

Decision rule based on P-Value

$$\text{Reject } H_0 : \mu_i = 0 \text{ at } 5\% \text{ level if}$$
$$\text{P-Value } < 5\%$$

For $T \geq 60$

$$\text{P-value} = 2 \times \Pr(z > |t_{\mu_i = 0}|), \; z \sim N(0, 1)$$

Test for Specific Coefficient Value

$$H_0 : \mu_i = \mu_i^0 \text{ vs. } H_1 : \mu_i \neq \mu_i^0$$

1. Test statistic

$$t_{\mu_i = \mu_i^0} = \frac{\hat{\mu}_i - \mu_i^0}{\widehat{\mathsf{SE}}(\hat{\mu}_i)}$$

Intuition:

- If $t_{\mu_i = \mu_i^0} \approx \mu_i^0$ then $\hat{\mu}_i \approx \mu_i^0$, and $H_0 : \mu_i = \mu_i^0$ should not be rejected

- If $|t_{\mu_i = \mu_i^0}| > 2$, say, then $\hat{\mu}_i$ is more than 2 values of $\widehat{\mathsf{SE}}(\hat{\mu}_i)$ away from μ_i^0. This is very unlikely if $\mu_i = \mu_i^0$, so $H_0 : \mu_i = \mu_i^0$ should be rejected.

2. Set significance level and determine critical value

$$\Pr(\text{Type I error}) = 5\%$$

3. Decision rule:

$$\text{Reject } H_0 : \mu_i = \mu_i^0 \text{ at } 5\% \text{ level if}$$
$$|t_{\mu_i=\mu_i^0}| > cv_{.975} = q_{.975}^{t_{T-1}}$$

Useful Rule of Thumb:

If $T \geq 60$ then $cv_{.975} \approx 2$ and the decision rule is

$$\text{Reject } H_0 : \mu_i = \mu_i^0 \text{ at } 5\% \text{ level if}$$
$$|t_{\mu_i=\mu_i^0}| > 2$$

Relationship Between Hypothesis Tests and Confidence Intervals

$$H_0 : \mu_i = \mu_i^0 \text{ vs. } H_1 : \mu_i \neq \mu_i^0$$
$$\text{level } = 5\%$$
$$cv_{.975} = q_{.975}^{t_{T-1}} \approx 2 \text{ for } T > 60$$
$$t_{\mu_i = \mu_i^0} = \frac{\hat{\mu}_i - \mu_i^0}{\widehat{\mathsf{SE}}(\hat{\mu}_i)}$$
$$\text{Reject at 5\% level if } |t_{\mu_i = \mu_i^0}| > 2$$

Approximate 95% confidence interval for μ_i

$$\begin{aligned} \hat{\mu}_i &= \pm 2 \cdot \widehat{\mathsf{SE}}(\hat{\mu}_i) \\ &= [\hat{\mu}_i - 2 \cdot \widehat{\mathsf{SE}}(\hat{\mu}_i),\ \hat{\mu}_i + 2 \cdot \widehat{\mathsf{SE}}(\hat{\mu}_i)] \end{aligned}$$

Decision: Reject $H_0 : \mu_i = \mu_i^0$ at 5% level if μ_i^0 does not lie in 95% confidence interval.

Test for Sign

$$H_0 : \mu_i = 0 \text{ vs. } H_1 : \mu_i > 0$$

1. Test statistic

$$t_{\mu_i=0} = \frac{\hat{\mu}_i}{\widehat{\text{SE}}(\hat{\mu}_i)}$$

Intuition:

- If $t_{\mu_i=\mu_i^0} \approx 0$ then $\hat{\mu}_i \approx 0$, and $H_0 : \mu_i = 0$ should not be rejected
- If $t_{\mu_i=\mu_i^0} >> 0$, then this is very unlikely if $\mu_i = 0$, so $H_0 : \mu_i = 0$ vs. $H_1 : \mu_i > 0$ should be rejected.

2. Set significance level and determine critical value

$$\text{Pr(Type I error)} = 5\%$$

Critical value cv is determined using

$$\begin{aligned} \Pr(t_{T-1} &> cv_{.05}) = 0.05 \\ &\Rightarrow cv_{.05} = q_{.95}^{t_{T-1}} \end{aligned}$$

where $q_{.95}^{t_{T-1}} = 95\%$ quantile of Student-t distribution with $T-1$ degrees of freedom.

3. Decision rule:

$$\text{Reject } H_0 : \mu_i = 0 \text{ vs. } H_1 : \mu_i > 0 \text{ at } 5\% \text{ level if}$$
$$t_{\mu_i=0} > q_{.95}^{t_{T-1}}$$

Useful Rule of Thumb:

If $T \geq 60$ then $q_{.95}^{t_{T-1}} \approx q_{.95}^{z} = 1.645$ and the decision rule is

$$\text{Reject } H_0 : \mu_i = 0 \text{ vs. } H_1 : \mu_i > 0 \text{ at 5\% level if}$$
$$t_{\mu_i=0} > 1.645$$

4. P-Value of test

$$\begin{aligned}&\text{significance level at which test is just rejected}\\ &= \Pr(t_{T-1} > t_{\mu_i=0})\\ &= \Pr(Z > t_{\mu_i=0}) \text{ for } T \geq 60\end{aligned}$$

Paired Two-Sample Test for Equality of Means

Example: Suppose it is thought that $E[r_i] = E[r_j]$. That is, the hypotheses of interest are

$$\begin{aligned} H_0 &: \mu_i = \mu_j \\ H_1 &: \mu_i \neq \mu_j \end{aligned}$$

An equivalent set of hypotheses are

$$\begin{aligned} H_0 &: \mu_i - \mu_j = 0 \\ H_1 &: \mu_i - \mu_j \neq 0 \end{aligned}$$

Q: How can we test these hypotheses?

Compute paired return differences

$$d_t = r_{it} - r_{jt},\ t = 1, \ldots, T$$

Under $H_0 : \mu_i - \mu_j = 0$ it follows that $\mu_d = E[d_t] = 0$. Therefore, the appropriate test statistic is

$$\begin{aligned} t_{\mu_d=0} &= \frac{\hat{\mu}_d}{\widehat{\mathsf{SE}}(\hat{\mu}_d)} \\ \hat{\mu}_d &= \frac{1}{T}\sum_{t=1}^{T} d_t \\ \widehat{\mathsf{SE}}(\hat{\mu}_d) &= \frac{\hat{\sigma}_d}{\sqrt{T}} \end{aligned}$$

Result: Under the CER model and $H_0 : \mu_i - \mu_j = 0$,

$$t_{\mu_d=0} \sim t_{T-1}$$

Decision rule: Reject $H_0 : \mu_i - \mu_j = 0$ at 5% level if $|t_{\mu_d=0}| > q_{.975}^{T-1}$ or if p-value is less than .05.

Test for Normal Distribution

$$\begin{aligned} H_0 &: r_t \sim \text{iid } N(\mu, \sigma^2) \\ H_1 &: r_t \sim \text{ not normal} \end{aligned}$$

1. Test statistic (Jarque-Bera statistic)

$$\text{JB} = \frac{T}{6}\left(\widehat{\text{skew}}^2 + \frac{(\widehat{\text{kurt}} - 3)^2}{4}\right)$$

Intuition

- If $r_t \sim$ iid $N(\mu, \sigma^2)$ then $\widehat{\mathsf{skew}}(r_t) \approx 0$ and $\widehat{\mathsf{kurt}}(r_t) \approx 3$ so that JB ≈ 0.

- If r_t is not normally distributed then $\widehat{\mathsf{skew}}(r_t) \neq 0$ and/or $\widehat{\mathsf{kurt}}(r_t) \neq 3$ so that JB $>> 0$

Distribution of JB under H_0

If $H_0 : r_t \sim$ iid $N(\mu, \sigma^2)$ is true then

$$\text{JB} \sim \chi^2(2)$$

where $\chi^2(2)$ denotes a chi-square distribution with 2 degrees of freedom (d.f.).

2. Set significance level and determine critical value

$$\Pr(\text{Type I error}) = 5\%$$

Critical value cv is determined using

$$\begin{aligned} \Pr(\chi^2(2) &> cv) = 0.05 \\ &\Rightarrow cv = q_{.95}^{\chi^2(2)} \approx 6 \end{aligned}$$

where $q_{.95}^{\chi^2(2)} \approx 6 \approx 95\%$ quantile of chi-square distribution with 2 degrees of freedom.

3. Decision rule:

$$\text{Reject } H_0 : r_t \sim \text{iid } N(\mu, \sigma^2)$$
$$\text{at } 5\% \text{ level if JB} > 6$$

4. P-Value of test

$$\text{significance level at which test is just rejected}$$
$$= \mathsf{Pr}(\chi^2(2) > \mathsf{JB})$$

Test for No Autocorrelation

Recall, the jth lag autocorrelation for r_t is

$$\begin{aligned} \rho_j &= \text{cor}(r_t, r_{t-j}) \\ &= \frac{\text{cov}(r_t, r_{t-j})}{\text{var}(r_t)} \end{aligned}$$

Hypotheses to be tested

$$\begin{aligned} H_0 &: \rho_j = 0, \text{ for all } j = 1, \ldots, q \\ H_1 &: \rho_j \neq 0 \text{ for some } j \end{aligned}$$

1. Estimate ρ_j using sample autocorrelation

$$\hat{\rho}_j = \frac{\frac{1}{T}\sum_{t=j+1}^{T}(r_t - \hat{\mu})(r_{t-j} - \hat{\mu})}{\frac{1}{T}\sum_{t=1}^{T}(r_t - \hat{\mu})^2}$$

Result: Under $H_0 : \rho_j = 0$ for all $j = 1, \dots, q$, if T is large then

$$\hat{\rho}_j \sim N\left(0, \frac{1}{T}\right) \text{ for all } j \geq 1$$

$$\mathrm{SE}(\hat{\rho}_j) = \frac{1}{\sqrt{T}}$$

2. Test Statistic

$$t_{\rho_{j=0}} = \frac{\hat{\rho}_j}{\mathrm{SE}(\hat{\rho}_j)} = \sqrt{T}\hat{\rho}_j$$

and 95% confidence interval

$$\hat{\rho}_j \pm 2 \cdot \frac{1}{\sqrt{T}}$$

3. Decision rule

$$\text{Reject } H_0 : \rho_j = 0 \text{ at 5\% level}$$
$$\text{if } |t_{\rho_{j=0}}| = \left|\sqrt{T}\hat{\rho}_j\right| > 2$$

That is, reject if

$$\hat{\rho}_j > \frac{2}{\sqrt{T}} \text{ or } \hat{\rho}_j < \frac{-2}{\sqrt{T}}$$

Diagnostics for Constant Parameters

$$r_t \sim \text{iid } N(\mu, \sigma^2),\ t = 1, \ldots, T$$
$$H_0 : \mu \text{ is constant over time}$$
$$H_1 : \mu \text{ changes over time}$$

Common test and diagnostic

- Two-Sample t-test for Structural Change
- Rolling estimates of μ

Two-Sample t-test for Structural Change

Idea: Split sample into two pieces

$$\begin{aligned} \text{sample 1} &: \quad t = 1, \ldots, T_1 \\ \text{sample 2} &: \quad t = T_1 + 1, \ldots, T; T_2 = T - T_1 \end{aligned}$$

The CER model assumptions imply that these two samples are independent. Define

$$\begin{aligned} \mu_1 &= E[r_t] \text{ on sample 1} \\ \mu_2 &= E[r_t] \text{ on sample 2} \end{aligned}$$

Then, we can test for structural change using the hypotheses

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \text{ (no structural change)} \\ H_1 &: \mu_1 \neq \mu_2 \text{ (structural change)} \end{aligned}$$

Remark

The test statistic can be computed assuming that var(r_t) is the same on samples 1 and 2, and it can be computed assuming that it is different on the two samples.

Test Statistic (assuming constant variance)

$$
\begin{aligned}
t_{\mu_1=\mu_2} &= \frac{\hat{\mu}_1 - \hat{\mu}_2}{\widehat{\mathsf{SE}}(\hat{\mu}_1 - \hat{\mu}_2)} \\
\widehat{\mathsf{SE}}(\hat{\mu}_1 - \hat{\mu}_2) &= \hat{\sigma}_{pool}\sqrt{\frac{1}{T_1} + \frac{1}{T_2}} \\
\hat{\sigma}^2_{pool} &= \left\{\frac{(T_1 - 1)\hat{\sigma}_1^2 + (T_2 - 1)\hat{\sigma}_2^2}{T_1 + T_2 - 1}\right\}
\end{aligned}
$$

Result: Under the CER model and $H_0 : \mu_1 = \mu_2$

$$
t_{\mu_1=\mu_2} \sim t_{T-2}
$$

Decision Rule: Reject $H_0 : \mu_i - \mu_j = 0$ at 5% level if $|t_{\mu_d=0}| > q^{T-2}_{.975}$ or if p-value is less than .05.

Rolling Means

Idea: compute estimate of μ over rolling windows of length $n < T$

$$
\begin{aligned}
\hat{\mu}_t(n) &= \frac{1}{n}\sum_{i=0}^{n-1} r_{t-i} \\
&= \frac{1}{n}(r_t + r_{t-1} + \cdots + r_{t-n+1})
\end{aligned}
$$

R function (package zoo)

```
rollapply
```

If $H_0 : \mu$ is constant is true, then $\hat{\mu}_t(n)$ should stay fairly constant over different windows.

If $H_0 : \mu$ is constant is false, then $\hat{\mu}_t(n)$ should fluctuate across different windows

Rolling Variances and Standard Deviations

Idea: Compute estimates of σ^2 and σ over rolling windows of length $n < T$

$$\begin{aligned} \hat{\sigma}_t^2(n) &= \frac{1}{n-1} \sum_{i=0}^{n-1} (r_{t-i} - \hat{\mu}_t(n))^2 \\ \hat{\sigma}_t(n) &= \sqrt{\hat{\sigma}_t^2(n)} \end{aligned}$$

If $H_0 : \sigma$ is constant is true, then $\hat{\sigma}_t(n)$ should stay fairly constant over different windows.

If $H_0 : \sigma$ is constant is false, then $\hat{\sigma}_t(n)$ should fluctuate across different windows

Rolling Covariances and Correlations

Idea: Compute estimates of σ_{jk} and ρ_{jk} over rolling windows of length $n < T$

$$
\begin{aligned}
\hat{\sigma}_{jk,t}(n) &= \frac{1}{n-1}\sum_{i=0}^{n-1}(r_{jt-i} - \hat{\mu}_j(n))(r_{kt-i} - \hat{\mu}_k(n)) \\
\hat{\rho}_{jk,t}(n) &= \frac{\hat{\sigma}_{jk,t}(n)}{\hat{\sigma}_{jt}(n)\hat{\sigma}_{kt}(n)}
\end{aligned}
$$

If $H_0 : \rho_{ij}$ is constant is true, then $\hat{\rho}_{jk,t}(n)$ should stay fairly constant over different windows.

If $H_0 : \rho_{ij}$ is constant is false, then $\hat{\rho}_{jk,t}(n)$ should fluctuate across different windows