Hypothesis Testing – Distribution

Jiří Neubauer

Department of Econometrics FVL UO Brno
office 69a, tel. 973 442029
email:Jiri.Neubauer@unob.cz

Hypothesis Testing

A statistical hypothesis is a claim (or statement) about

- a population parameters (μ, σ^2, π, λ, $\dots$),

Hypothesis Testing

A statistical hypothesis is a claim (or statement) about

- a population parameters (μ, σ^2, π, λ, . . .),
- a distribution (normal, Poisson, . . .).

Hypothesis Testing

- **A null hypothesis** H is a claim (or statement) about a population parameter that is assumed to be true until it is declared false ... for example
  - $H : \mu = \mu_0$,

Hypothesis Testing

- **A null hypothesis** H is a claim (or statement) about a population parameter that is assumed to be true until it is declared false ... for example
  - $H: \mu = \mu_0$,
- **An alternative hypothesis** A is a claim about a population parameter that will be true if the null hypothesis is false
  1. $A: \mu \neq \mu_0 \rightarrow$ both-sided test,

Hypothesis Testing

- **A null hypothesis** H is a claim (or statement) about a population parameter that is assumed to be true until it is declared false ... for example
  - $H: \mu = \mu_0$,
- **An alternative hypothesis** A is a claim about a population parameter that will be true if the null hypothesis is false
  1. $A: \mu \neq \mu_0 \rightarrow$ both-sided test,
  2. $A: \mu > \mu_0 \rightarrow$ one-sided test,

Hypothesis Testing

- **A null hypothesis** H is a claim (or statement) about a population parameter that is assumed to be true until it is declared false . . . for example
  - $H: \mu = \mu_0$,
- **An alternative hypothesis** A is a claim about a population parameter that will be true if the null hypothesis is false
  1. $A: \mu \neq \mu_0 \rightarrow$ both-sided test,
  2. $A: \mu > \mu_0 \rightarrow$ one-sided test,
  3. $A: \mu < \mu_0 \rightarrow$ one-sided test.

Hypothesis Testing

reality	H is true		H is false	
decision about H		prob.		prob.
not reject	correct decision	$1 - \alpha$	type II error	β
reject	type I error	α	correct decision	$1 - \beta$

Hypothesis Testing

- If we reject the null hypothesis which is true, we call this **type I error**. The probability of this error is $\alpha \Rightarrow$ **a significance level**. A number $1 - \alpha$ is probability that we do not reject the true hypothesis H.

Hypothesis Testing

- If we reject the null hypothesis which is true, we call this **type I error**. The probability of this error is $\alpha \Rightarrow$ **a significance level**. A number $1 - \alpha$ is probability that we do not reject the true hypothesis H.
- If we accept the null hypothesis although is false, we call this **type II error**. The probability of this error is β. A number $1 - \beta \Rightarrow$ **a power of the test** is the probability that we reject the null hypothesis H, if it is false.

Hypothesis Testing

To test the null hypothesis we use a function of a random sample $T = T(x_1, x_2, \ldots, x_n)$, so called test statistic, which has under the null hypothesis H known distribution (usually t, u, χ^2, F).
We divide the all possible values of the test statistic into

Hypothesis Testing

To test the null hypothesis we use a function of a random sample $T = T(x_1, x_2, \ldots, x_n)$, so called test statistic, which has under the null hypothesis H known distribution (usually t, u, χ^2, F).
We divide the all possible values of the test statistic into

- $W_{1-\alpha}$ - **nonrejection region of** H – the set of values connected with the hypothesis H,

Hypothesis Testing

To test the null hypothesis we use a function of a random sample $T = T(x_1, x_2, \ldots, x_n)$, so called test statistic, which has under the null hypothesis H known distribution (usually t, u, χ^2, F).
We divide the all possible values of the test statistic into

- $W_{1-\alpha}$ - **nonrejection region of H** – the set of values connected with the hypothesis H,
- W_{α} - **rejection region of H** – the set of values connected with the hypothesis A.

Steps to Perform a Test of Hypothesis

1. State the null and alternative hypothesis H and A.

Steps to Perform a Test of Hypothesis

1. State the null and alternative hypothesis H and A.
2. Select a significance level α (usually 0.05 a 0.01).

Steps to Perform a Test of Hypothesis

1. State the null and alternative hypothesis H and A.
2. Select a significance level α (usually 0.05 a 0.01).
3. Choose the test statistic.

Steps to Perform a Test of Hypothesis

1. State the null and alternative hypothesis H and A.
2. Select a significance level α (usually 0.05 a 0.01).
3. Choose the test statistic.
4. Determine the rejection region W_α.

Steps to Perform a Test of Hypothesis

1. State the null and alternative hypothesis H and A.
2. Select a significance level α (usually 0.05 a 0.01).
3. Choose the test statistic.
4. Determine the rejection region W_α.
5. Calculate the value of the test statistic.

Steps to Perform a Test of Hypothesis

6. Make a decision:

Steps to Perform a Test of Hypothesis

6. Make a decision:
    - If the value of the test statistic falls in the rejection region, we reject the null hypothesis H and say that we accept the alternative hypothesis A with the probability $1 - \alpha$.

Steps to Perform a Test of Hypothesis

6. Make a decision:
   - If the value of the test statistic falls in the rejection region, we reject the null hypothesis H and say that we accept the alternative hypothesis A with the probability $1 - \alpha$.
   - If the value of the test statistic falls in the nonrejection region, we do not reject the null hypothesis H.

Chi-Square Goodness of Fit Test

We divide values of a random sample $x_1, x_2, \ldots, x_n$ into k disjunct classes, where $n_j, j = 1, 2, \ldots, k$, is frequency of the class j and π_j is a probability that the random variable X has value from the class j, calculated on condition that X has an assumed distribution.

The main idea of the test is to compare relative frequencies n_j/n with theoretical probabilities π_j.

Chi-Square Goodness of Fit Test

We state the null and alternative hypothesis:
H : the random X has an assumed distribution $\rightarrow$ A : the random X has not an assumed distribution.
The test statistic is

$$\chi^2 = \sum_{j=1}^{k} \frac{(n_j - n\pi_j)^2}{n\pi_j},$$

which has under the null hypothesis H for large n (asymptotically) a Pearson χ^2-distribution with $\nu = k - c - 1$ degrees of freedom, where c is a number of estimated parameters of the assumed distribution.
A rejection region is

$$W_\alpha = \left\{\chi^2, \chi^2 \geq \chi^2_{1-\alpha}(\nu)\right\},$$

where $\chi^2_{1-\alpha}(\nu)$ is a quantile of the Pearson χ^2-distribution.

Chi-Square Goodness of Fit Test

Recommendation:

$$n\pi_j > 5, \quad j = 1, 2, \ldots, k.$$

If this condition is not satisfied, it is necessary to join the classes.

Tests of Skewness and Kurtosis

The normal distribution has $\alpha_3 = 0$ a $\alpha_4 = 0$. We can use these properties to test normality. We calculate a sample skewness and kurtosis (they are estimates of α_3 and α_4)

$$\hat{\alpha}_3 = a_3 = \frac{1}{ns_n^3}\sum_{i=1}^{n}(x_i - \overline{x})^3, \quad \hat{\alpha}_4 = a_4 = \frac{1}{ns_n^4}\sum_{i=1}^{n}(x_i - \overline{x})^4 - 3.$$

We state hypothesis:
$H_1 : \alpha_3 = 0 \rightarrow A_1 : \alpha_3 \neq 0$
$H_2 : \alpha_4 = 0 \rightarrow A_2 : \alpha_4 \neq 0$

Tests of Skewness and Kurtosis

1. $H_1 : \alpha_3 = 0 \rightarrow A_1 : \alpha_3 \neq 0$
   The test statistic is

$$u_3 = \frac{a_3}{\sqrt{D(a_3)}}, \quad \text{where} \quad D(a_3) = \frac{6(n-2)}{(n+1)(n+3)},$$

   which has under the null hypothesis H_1 asymptotically a normal distribution $N(0,1)$. A rejection region is

$$W_\alpha = \left\{u_3, |u_3| \geq u_{1-\frac{\alpha}{2}}\right\},$$

   where $u_{1-\frac{\alpha}{2}}$ is a quantile of $N(0,1)$.

Tests of Skewness and Kurtosis

2. $H_2 : \alpha_4 = 0 \to A_2 : \alpha_4 \neq 0$
   The test statistic is

$$u_4 = \frac{a_4 + \frac{6}{n+1}}{\sqrt{D(a_4)}}, \quad \text{where} \quad D(a_4) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)},$$

   which has under the null hypothesis H_2 asymptotically a normal distribution $N(0,1)$. A rejection region is

$$W_\alpha = \left\{u_4, |u_4| \geq u_{1-\frac{\alpha}{2}}\right\},$$

   where $u_{1-\frac{\alpha}{2}}$ is a quantile of $N(0,1)$.

Compound Tests of Skewness and Kurtosis

We state hypothesis:
H : a random variable X has a normal distribution $\rightarrow$ A : a random variable X has not a normal distribution.
Test statistic is

$$C = u_3^2 + u_4^2,$$

which has under the null hypothesis H approximately χ^2 distribution with two degrees of freedom.
u_3 and u_4 are test statistics defined above. A rejection region is

$$W_\alpha = \left\{C, C \geq \chi^2_{1-\alpha}(2)\right\},$$

where $\chi^2_{1-\alpha}(2)$ is a quantile of the Pearson χ^2-distribution.