Hypothesis Testing

The two main areas of statistics are **estimation** and **hypothesis testing**

Hypothesis Testing

The two main areas of statistics are **estimation** and **hypothesis testing**

The objective in estimation is to determine the unknown value of some parameter assiciated with a statistical distribution.

Hypothesis Testing

The two main areas of statistics are **estimation** and **hypothesis testing**

The objective in estimation is to determine the unknown value of some parameter assiciated with a statistical distribution.

The objective in hypothesis testing is to decide which of two contradictory claims or assertions is likely to be true.

Hypothesis Testing

The two main areas of statistics are **estimation** and **hypothesis testing**

The objective in estimation is to determine the unknown value of some parameter assiciated with a statistical distribution.

The objective in hypothesis testing is to decide which of two contradictory claims or assertions is likely to be true.

In both estimation and hypothesis testing, a sample from the population under consideration is the basis of all procedures.

Terminology

Hypothesis testing has strongly established terminology.

Terminology

Hypothesis testing has strongly established terminology.

The first part of this terminology has to do with the contradictory assertions. These are known as:

- The **null hypothesis** H_0
- The **alternative hypothesis** H_a (or in some texts H_1)

Terminology

Hypothesis testing has strongly established terminology.

The first part of this terminology has to do with the contradictory assertions. These are known as:

- The **null hypothesis** H_0
- The **alternative hypothesis** H_a (or in some texts H_1)

Unfortunately these names are rather nondescript which leads to a certain amount confusion over which of the two claims is the null hypothesis, and which is the alternative.

Terminology

Hypothesis testing has strongly established terminology.

The first part of this terminology has to do with the contradictory assertions. These are known as:

- The **null hypothesis** H_0
- The **alternative hypothesis** H_a (or in some texts H_1)

Unfortunately these names are rather nondescript which leads to a certain amount confusion over which of the two claims is the null hypothesis, and which is the alternative.

The null hypothesis is *favored* in the sense that the burden of proof is on the proponents of the alternative hypothesis.

Terminology

Hypothesis testing has strongly established terminology.

The first part of this terminology has to do with the contradictory assertions. These are known as:

- The **null hypothesis** H_0
- The **alternative hypothesis** H_a (or in some texts H_1)

Unfortunately these names are rather nondescript which leads to a certain amount confusion over which of the two claims is the null hypothesis, and which is the alternative.

The null hypothesis is *favored* in the sense that the burden of proof is on the proponents of the alternative hypothesis.

Only when presented with strong evidence (based on the sample) that it is false will we **reject the null hypothesis**

Terminology

The mechanism for rejecting or failing to reject the null hypothesis is a **test procedure** which consists of:

- A **test statistic** derived from the sample
- A rule for rejecting the null hypothesis.

Terminology

The mechanism for rejecting or failing to reject the null hypothesis is a **test procedure** which consists of:

- A **test statistic** derived from the sample
- A rule for rejecting the null hypothesis.

There are a number of ways to specify the rejection rule. The author mentions two, the "classical" and p-value approaches.

Terminology

The mechanism for rejecting or failing to reject the null hypothesis is a **test procedure** which consists of:

- A **test statistic** derived from the sample
- A rule for rejecting the null hypothesis.

There are a number of ways to specify the rejection rule. The author mentions two, the "classical" and p-value approaches.

We will consider the p-value method because it is easy to understand and we already know how to do the required computations.

Terminology

Additional terminlolgy has to do with the types of *errors* that are possible.

Terminology

Additional terminlolgy has to do with the types of *errors* that are possible.

The null and alternative hypotheses are always constructed to be contradictory: one and only one of the two must be true, the other must be false.

Terminology

Additional terminlolgy has to do with the types of *errors* that are possible.

The null and alternative hypotheses are always constructed to be contradictory: one and only one of the two must be true, the other must be false.

This means there are only two ways to be wrong when we reject or fail to reject H_0:

- Reject H_0 when it is true (**type I error**)
- Fail to reject H_0 when it is false (**type II error**)

Terminology

Additional terminlolgy has to do with the types of *errors* that are possible.

The null and alternative hypotheses are always constructed to be contradictory: one and only one of the two must be true, the other must be false.

This means there are only two ways to be wrong when we reject or fail to reject H_0:

- Reject H_0 when it is true (**type I error**)
- Fail to reject H_0 when it is false (**type II error**)

- The probability of a **type I error** is denoted by α
- The probability of a **type II error** is denoted by β

Terminology

We will use a slight modification of the p-value method presented in the text, which assumes we are going to transform to a standard normal before calculating the p-value.

Terminology

We will use a slight modification of the p-value method presented in the text, which assumes we are going to transform to a standard normal before calculating the p-value.

We will take advantage of widely available software programs to directly calculate p-values without transforming to standard normal.

Terminology

We will use a slight modification of the p-value method presented in the text, which assumes we are going to transform to a standard normal before calculating the p-value.

We will take advantage of widely available software programs to directly calculate p-values without transforming to standard normal.

Recall that if we take a sample of size n from a $N(\mu, \sigma)$ population, that is, a normal population with mean μ and standard deviation σ, the **sample mean** has a $N(\mu, \sigma/sqrtn)$ distribution.

Terminology

We will use a slight modification of the p-value method presented in the text, which assumes we are going to transform to a standard normal before calculating the p-value.

We will take advantage of widely available software programs to directly calculate p-values without transforming to standard normal.

Recall that if we take a sample of size n from a $N(\mu, \sigma)$ population, that is, a normal population with mean μ and standard deviation σ, the **sample mean** has a $N(\mu, \sigma/sqrtn)$ distribution.

This means we can treat the sample mean as a single observation from a normal population with mean μ and standard deviation $\sigma/\sqrt{n}$.

p-value method for left-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu < \mu_0$ is the following:

p-value method for left-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu < \mu_0$ is the following:

Let $\overline{x}$ be the mean of a sample of size n from a normal population with standard deviation σ.

Reject H_0 if $P(X < \overline{x}) < \alpha$ where α is the alpha level (probability of Type I error).

p-value method for left-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu < \mu_0$ is the following:

Let $\overline{x}$ be the mean of a sample of size n from a normal population with standard deviation σ.

Reject H_0 if $P(X < \overline{x}) < \alpha$ where α is the alpha level (probability of Type I error).

Computationally, this can be stated as:

Compute one of the following:

- `=NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

p-value method for left-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu < \mu_0$ is the following:

Let $\overline{x}$ be the mean of a sample of size n from a normal population with standard deviation σ.

Reject H_0 if $P(X < \overline{x}) < \alpha$ where α is the alpha level (probability of Type I error).

Computationally, this can be stated as:

Compute one of the following:

- `=NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

Reject H_0 if the result is less than (or equal to) α.

p-value method for right-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$ is the following:

p-value method for right-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$ is the following:

Let $\overline{x}$ be the mean of a sample of size n from a normal population with standard deviation σ.

Reject H_0 if $P(X > \overline{x}) < \alpha$ where α is the alpha level (probability of Type I error).

p-value method for right-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$ is the following:

Let $\overline{x}$ be the mean of a sample of size n from a normal population with standard deviation σ.

Reject H_0 if $P(X > \overline{x}) < \alpha$ where α is the alpha level (probability of Type I error).

Computationally, this can be stated as:

Compute one of the following:

- `=1-NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `1-pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

p-value method for right-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$ is the following:

Let $\overline{x}$ be the mean of a sample of size n from a normal population with standard deviation σ.

Reject H_0 if $P(X > \overline{x}) < \alpha$ where α is the alpha level (probability of Type I error).

Computationally, this can be stated as:

Compute one of the following:

- `=1-NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `1-pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

Reject H_0 if the result is less than (or equal to) α.

p-value method for two-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu \neq \mu_0$ is the following:

p-value method for two-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu \neq \mu_0$ is the following:

Case 1: $\overline{x} \leq \mu_0$

Compute one of the following:

- `=NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

Reject H_0 if the result is less than (or equal to) $\alpha/2$.

p-value method for two-tailed test

The p-value method for testing $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu \neq \mu_0$ is the following:

Case 1: $\overline{x} \leq \mu_0$

Compute one of the following:

- `=NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

Reject H_0 if the result is less than (or equal to) $\alpha/2$.

Case 2: $\overline{x} > \mu_0$

Compute one of the following:

- `=1-NORMDIST(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (spreadsheet)
- `1-pnorm(`$\overline{x}$`,`μ_0`,`$\sigma/\sqrt{n}$`)` (R)

Reject H_0 if the result is less than (or equal to) $\alpha/2$.