STATISTICAL INFERENCES (2cr)
Chapter 10 One- and Two-Sample Tests of Hypotheses

Zhong Guan

Math, IUSB

Outline

1. 10.1 Statistical Hypotheses: General Concepts
2. 10.2 Testing a Statistical Hypothesis
3. 10.3 Using p-value

Example 1.

- Let p be the proportion of printed circuits that fail. Current procedure has $p_0 = 0.06$.
- A new method is proposed. To see if the new method results in an improvement.
- We want to test **statistical hypotheses**

$$H_0 : p = p_0 = 0.06 \quad \text{vs} \quad H_1 : p < p_0 = 0.06.$$

- We decided to select $n = 200$ circuits to test. Let Y be the number of circuits that fail.
- **Decision rule**: accept the improvement hypothesis H_1 if $Y \leqslant k$ (reject H_0), otherwise reject H_1 (accept H_0).
- $k = ?$

Example 1.

- Let p be the proportion of printed circuits that fail. Current procedure has $p_0 = 0.06$.
- A new method is proposed. To see if the new method results in an improvement.
- We want to test **statistical hypotheses**

$$H_0 : p = p_0 = 0.06 \quad \text{vs} \quad H_1 : p < p_0 = 0.06.$$

- We decided to select $n = 200$ circuits to test. Let Y be the number of circuits that fail.
- **Decision rule**: accept the improvement hypothesis H_1 if $Y \leqslant k$ (reject H_0), otherwise reject H_1 (accept H_0).
- $k = ?$

Example 1.

- Let p be the proportion of printed circuits that fail. Current procedure has $p_0 = 0.06$.
- A new method is proposed. To see if the new method results in an improvement.
- We want to test **statistical hypotheses**

$$H_0 : p = p_0 = 0.06 \quad \text{vs} \quad H_1 : p < p_0 = 0.06.$$

- We decided to select $n = 200$ circuits to test. Let Y be the number of circuits that fail.
- **Decision rule**: accept the improvement hypothesis H_1 if $Y \leqslant k$ (reject H_0), otherwise reject H_1 (accept H_0).
- $k = ?$

Example 1.

- Let p be the proportion of printed circuits that fail. Current procedure has $p_0 = 0.06$.
- A new method is proposed. To see if the new method results in an improvement.
- We want to test **statistical hypotheses**

$$H_0 : p = p_0 = 0.06 \quad \text{vs} \quad H_1 : p < p_0 = 0.06.$$

- We decided to select $n = 200$ circuits to test. Let Y be the number of circuits that fail.
- **Decision rule**: accept the improvement hypothesis H_1 if $Y \leqslant k$ (reject H_0), otherwise reject H_1 (accept H_0).
- $k = ?$

Example 1.

- Let p be the proportion of printed circuits that fail. Current procedure has $p_0 = 0.06$.
- A new method is proposed. To see if the new method results in an improvement.
- We want to test **statistical hypotheses**

$$H_0 : p = p_0 = 0.06 \quad \text{vs} \quad H_1 : p < p_0 = 0.06.$$

- We decided to select $n = 200$ circuits to test. Let Y be the number of circuits that fail.
- **Decision rule**: accept the improvement hypothesis H_1 if $Y \leqslant k$ (reject H_0), otherwise reject H_1 (accept H_0).
- $k = ?$

Example 1.

- Let p be the proportion of printed circuits that fail. Current procedure has $p_0 = 0.06$.
- A new method is proposed. To see if the new method results in an improvement.
- We want to test **statistical hypotheses**

$$H_0 : p = p_0 = 0.06 \quad \text{vs} \quad H_1 : p < p_0 = 0.06.$$

- We decided to select $n = 200$ circuits to test. Let Y be the number of circuits that fail.
- **Decision rule**: accept the improvement hypothesis H_1 if $Y \leqslant k$ (reject H_0), otherwise reject H_1 (accept H_0).
- $k =?$

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the test statistic.

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the **test statistic**.

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the test statistic.

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the test statistic.

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the test statistic.

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the test statistic.

Two Types of Errors

- For example, we choose $k = 7$ so that if $Y \leqslant 7$ or $\hat{p} \leqslant 0.035$, then we accept the improvement hypothesis.
- However, whatever k is, the decision made based on a sample could be wrong. There are two types of errors.

	H_0 true	H_0 false
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

- If we accept H_0, then we will keep the current procedure. This is called no change hypothesis. It is also called **null hypothesis**.
- H_1 is called **alternative hypothesis**.
- $y \leqslant k$ is called **rejection** or **critical region** of H_0.
- y or $\hat{p}$ is called the test statistic.

Chances to Make Mistakes: Type I Error

- The Type I error of the above rule is that $Y \leqslant 7$ but the truth is $p = 0.06$.
- The probability of this error is

$$P(Y \leqslant 7 | p = 0.06) = \sum_{y=0}^{7} \binom{200}{y} (0.06)^y (0.94)^{200-y}$$

$$= 0.0829 = 8.29\%.$$

Chances to Make Mistakes: Type I Error

- The Type I error of the above rule is that $Y \leqslant 7$ but the truth is $p = 0.06$.
- The probability of this error is

$$P(Y \leqslant 7 | p = 0.06) = \sum_{y=0}^{7} \binom{200}{y} (0.06)^y (0.94)^{200-y}$$

$$= 0.0829 = 8.29\%.$$

Chances to Make Mistakes: Type I Error

- The Type I error of the above rule is that $Y \leqslant 7$ but the truth is $p = 0.06$.
- The probability of this error is

$$P(Y \leqslant 7 | p = 0.06) = \sum_{y=0}^{7} \binom{200}{y} (0.06)^y (0.94)^{200-y}$$

$$= 0.0829 = 8.29\%.$$

Chances to Make Mistakes: Type I Error

- The Type I error of the above rule is that $Y \leqslant 7$ but the truth is $p = 0.06$.
- The probability of this error is

$$P(Y \leqslant 7 | p = 0.06) = \sum_{y=0}^{7} \binom{200}{y} (0.06)^y (0.94)^{200-y}$$

$$= 0.0829 = 8.29\%.$$

Chances to Make Mistakes: Type II Error

- Type II error: $Y > 7$ but the truth is $p < 0.06$.
- The probability of Type II error is, for $p < 0.06$,

$$\beta(p) = P(Y > 7|p) = \sum_{y=8}^{200} \binom{200}{y} p^y (1-p)^{200-y}$$

- For example,

$$\beta(0.03) = 1 - binomialcdf(7, n = 200, p = 0.03) = 0.267$$

$$\beta(0.02) = 1 - binomialcdf(7, n = 200, p = 0.02) = 0.038$$

$$\beta(0.01) = 1 - binomialcdf(7, n = 200, p = 0.01) = 4.64 \times 10^{-5}$$

Chances to Make Mistakes: Type II Error

- Type II error: $Y > 7$ but the truth is $p < 0.06$.
- The probability of Type II error is, for $p < 0.06$,

$$\beta(p) = P(Y > 7|p) = \sum_{y=8}^{200} \binom{200}{y} p^y (1-p)^{200-y}$$

- For example,

$$\beta(0.03) = 1 - binomialcdf(7, n = 200, p = 0.03) = 0.267$$

$$\beta(0.02) = 1 - binomialcdf(7, n = 200, p = 0.02) = 0.038$$

$$\beta(0.01) = 1 - binomialcdf(7, n = 200, p = 0.01) = 4.64 \times 10^{-5}$$

Chances to Make Mistakes: Type II Error

- Type II error: $Y > 7$ but the truth is $p < 0.06$.
- The probability of Type II error is, for $p < 0.06$,

$$\beta(p) = P(Y > 7|p) = \sum_{y=8}^{200} \binom{200}{y} p^y (1-p)^{200-y}$$

- For example,

$$\beta(0.03) = 1 - binomialcdf(7, n = 200, p = 0.03) = 0.267$$

$$\beta(0.02) = 1 - binomialcdf(7, n = 200, p = 0.02) = 0.038$$

$$\beta(0.01) = 1 - binomialcdf(7, n = 200, p = 0.01) = 4.64 \times 10^{-5}$$

Chances to Make Mistakes: Type II Error

- Type II error: $Y > 7$ but the truth is $p < 0.06$.
- The probability of Type II error is, for $p < 0.06$,

$$\beta(p) = P(Y > 7|p) = \sum_{y=8}^{200} \binom{200}{y} p^y (1-p)^{200-y}$$

- For example,

$$\beta(0.03) = 1 - binomialcdf(7, n = 200, p = 0.03) = 0.267$$

$$\beta(0.02) = 1 - binomialcdf(7, n = 200, p = 0.02) = 0.038$$

$$\beta(0.01) = 1 - binomialcdf(7, n = 200, p = 0.01) = 4.64 \times 10^{-5}$$

Chances to Make Mistakes: Type II Error

- Type II error: $Y > 7$ but the truth is $p < 0.06$.
- The probability of Type II error is, for $p < 0.06$,

$$\beta(p) = P(Y > 7|p) = \sum_{y=8}^{200} \binom{200}{y} p^y (1-p)^{200-y}$$

- For example,

$$\beta(0.03) = 1 - binomialcdf(7, n = 200, p = 0.03) = 0.267$$

$$\beta(0.02) = 1 - binomialcdf(7, n = 200, p = 0.02) = 0.038$$

$$\beta(0.01) = 1 - binomialcdf(7, n = 200, p = 0.01) = 4.64 \times 10^{-5}$$

Significance Level and Power of the Test

- Generally, we want to control the Type I error first: for a given **significance level** α, choose k so that

$$P(\text{Type I error}) \leqslant \alpha$$

- The **power** of the test is,

$$\text{Power}(p) = 1 - \beta(p).$$

- In the above example, for $p < 0.06$,

$$\text{Power}(p) = P(Y \leqslant 7|p) = 1 - P(Y > 7|p).$$

$$\text{Power}(0.03) = 0.833$$
$$\text{Power}(0.02) = 0.962,$$
$$\text{Power}(0.01) \approx 1.$$

Significance Level and Power of the Test

- Generally, we want to control the Type I error first: for a given **significance level** α, choose k so that

$$P(\text{Type I error}) \leqslant \alpha$$

- The **power** of the test is,

$$\text{Power}(p) = 1 - \beta(p).$$

- In the above example, for $p < 0.06$,

$$\text{Power}(p) = P(Y \leqslant 7|p) = 1 - P(Y > 7|p).$$

$$\text{Power}(0.03) = 0.833$$
$$\text{Power}(0.02) = 0.962,$$
$$\text{Power}(0.01) \approx 1.$$

Significance Level and Power of the Test

- Generally, we want to control the Type I error first: for a given **significance level** α, choose k so that

$$P(\text{Type I error}) \leqslant \alpha$$

- The **power** of the test is,

$$\text{Power}(p) = 1 - \beta(p).$$

- In the above example, for $p < 0.06$,

$$\text{Power}(p) = P(Y \leqslant 7|p) = 1 - P(Y > 7|p).$$

$$\text{Power}(0.03) = 0.833$$
$$\text{Power}(0.02) = 0.962,$$
$$\text{Power}(0.01) \approx 1.$$

Find Critical Value k

If H_0 is true, approximately,

$$Z = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} \sim N(0,1).$$

Thus

$$\alpha = P(Y \leqslant k | p = p_0) \approx \Phi\left(\frac{k - np_0}{\sqrt{np_0(1-p_0)}}\right)$$

So

$$\frac{k - np_0}{\sqrt{np_0(1-p_0)}} = -z_\alpha.$$

and

$$k = np_0 - z_\alpha \sqrt{np_0(1-p_0)}.$$

Rejection Region of H_0

- The **rejection region** of H_0 is $R = \{Y : Y \leqslant k\}$
- or $R = \{Z = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} \leqslant -z_\alpha\}$.
- Here k and $-z_\alpha$ are called the critical values of the **test statistics** Y and Z respectively.
- In the above example, $n = 200$, $p_0 = 0.06$. Choose $\alpha = 0.05$. $z_\alpha = z_{0.05} = 1.645$.

$$k = np_0 - z_\alpha\sqrt{np_0(1 - p_0)} = 6.42$$

- The rejection region of H_0 is

$$R = \{Y : Y \leqslant 6.4167\}$$

or

$$R = \{Z = \frac{Y - 12}{3.358571} \leqslant -1.645\}.$$

Rejection Region of H_0

- The **rejection region** of H_0 is $R = \{Y : Y \leqslant k\}$
- or $R = \{Z = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} \leqslant -z_\alpha\}$.
- Here k and $-z_\alpha$ are called the critical values of the **test statistics** Y and Z respectively.
- In the above example, $n = 200$, $p_0 = 0.06$. Choose $\alpha = 0.05$. $z_\alpha = z_{0.05} = 1.645$.

$$k = np_0 - z_\alpha \sqrt{np_0(1 - p_0)} = 6.42$$

- The rejection region of H_0 is

$$R = \{Y : Y \leqslant 6.4167\}$$

or

$$R = \{Z = \frac{Y - 12}{3.358571} \leqslant -1.645\}.$$

Rejection Region of H_0

- The **rejection region** of H_0 is $R = \{Y : Y \leqslant k\}$
- or $R = \{Z = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} \leqslant -z_\alpha\}$.
- Here k and $-z_\alpha$ are called the critical values of the **test statistics** Y and Z respectively.
- In the above example, $n = 200$, $p_0 = 0.06$. Choose $\alpha = 0.05$. $z_\alpha = z_{0.05} = 1.645$.

$$k = np_0 - z_\alpha\sqrt{np_0(1-p_0)} = 6.42$$

- The rejection region of H_0 is

$$R = \{Y : Y \leqslant 6.4167\}$$

or

$$R = \{Z = \frac{Y-12}{3.358571} \leqslant -1.645\}.$$

Rejection Region of H_0

- The **rejection region** of H_0 is $R = \{Y : Y \leqslant k\}$
- or $R = \left\{Z = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} \leqslant -z_\alpha\right\}$.
- Here k and $-z_\alpha$ are called the critical values of the **test statistics** Y and Z respectively.
- In the above example, $n = 200$, $p_0 = 0.06$. Choose $\alpha = 0.05$. $z_\alpha = z_{0.05} = 1.645$.

$$k = np_0 - z_\alpha\sqrt{np_0(1-p_0)} = 6.42$$

- The rejection region of H_0 is

$$R = \{Y : Y \leqslant 6.4167\}$$

or

$$R = \left\{Z = \frac{Y-12}{3.358571} \leqslant -1.645\right\}.$$

Rejection Region of H_0

- The **rejection region** of H_0 is $R = \{Y : Y \leqslant k\}$
- or $R = \{Z = \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} \leqslant -z_\alpha\}$.
- Here k and $-z_\alpha$ are called the critical values of the **test statistics** Y and Z respectively.
- In the above example, $n = 200$, $p_0 = 0.06$. Choose $\alpha = 0.05$. $z_\alpha = z_{0.05} = 1.645$.

$$k = np_0 - z_\alpha \sqrt{np_0(1 - p_0)} = 6.42$$

- The rejection region of H_0 is

$$R = \{Y : Y \leqslant 6.4167\}$$

or

$$R = \{Z = \frac{Y - 12}{3.358571} \leqslant -1.645\}.$$

A Continuous Random Variable Example

Example 2: Let X be the Brinell hardness measurement of ductile iron substantially annealed. Assume that the distribution of X is $N(\mu, 10^2)$. We shall test the null hypothesis $H_0 : \mu = 170$ against the alternative hypothesis $H_1 : \mu > 170$ based on a sample of size n.

(a) Define the test statistic;

(b) Define a critical region with a significance level $\alpha = 0.05$.

(c) A random sample of $n = 25$ observations of X yielded the following measurements.

170 167 174 179 179 156 163 156 187 156 183 179 174 179
170 156 187 179 183 174 187 167 159 170 179

Calculate the value of the test statistic.

(d) State your conclusion.

A Continuous Random Variable Example

Example 2: Let X be the Brinell hardness measurement of ductile iron substantially annealed. Assume that the distribution of X is $N(\mu, 10^2)$. We shall test the null hypothesis $H_0 : \mu = 170$ against the alternative hypothesis $H_1 : \mu > 170$ based on a sample of size n.

(a) Define the test statistic;

(b) Define a critical region with a significance level $\alpha = 0.05$.

(c) A random sample of $n = 25$ observations of X yielded the following measurements.

170 167 174 179 179 156 163 156 187 156 183 179 174 179
170 156 187 179 183 174 187 167 159 170 179

Calculate the value of the test statistic.

(d) State your conclusion.

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \,|\, H_0 \text{ is true}) = P(\bar{X} > c \,|\, \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \,|\, H_0 \text{ is true}) = P(\bar{X} > c \,|\, \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \,|\, H_0 \text{ is true}) = P(\bar{X} > c \,|\, \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

Solution of Example 2

(a) Since $\bar{X}$ is an estimator of μ, we use $\bar{X}$ as the test statistic.

(b) The rejection region of $H_0 : \mu = 170$ in favor of $H_1 : \mu > 170$:

$$\bar{X} > c$$

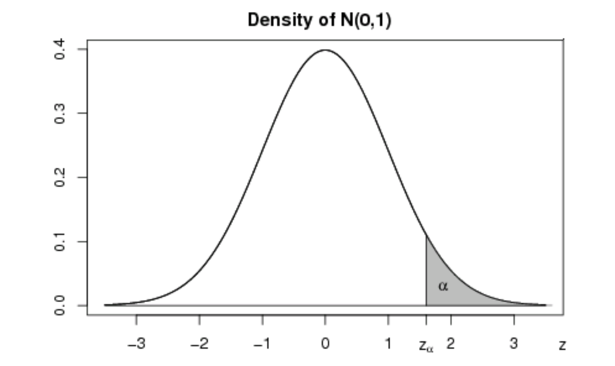
for some pre-determined critical value c according to α.

$$0.05 = \alpha = Prob(\text{Type I}) = ?$$

$$P(\text{Type I}) = P(\text{Rejec } H_0 \mid H_0 \text{ is true}) = P(\bar{X} > c \mid \mu = 170)$$

$$0.05 = \alpha = P\left(Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > \frac{c - 170}{10/\sqrt{n}}\right)$$

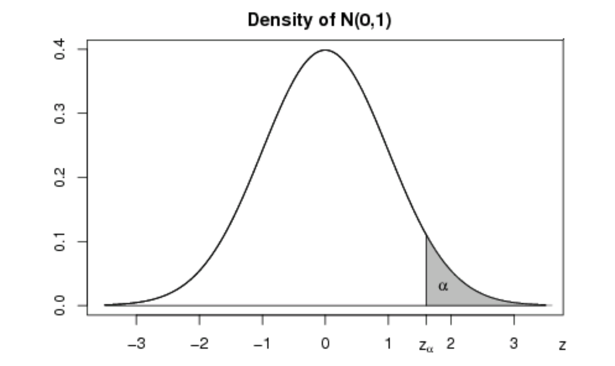
Solution of Example 2



So

$$\frac{c-170}{10/\sqrt{n}} = z_\alpha \quad \rightarrow \quad c = 170 + \frac{10 z_\alpha}{\sqrt{n}}$$

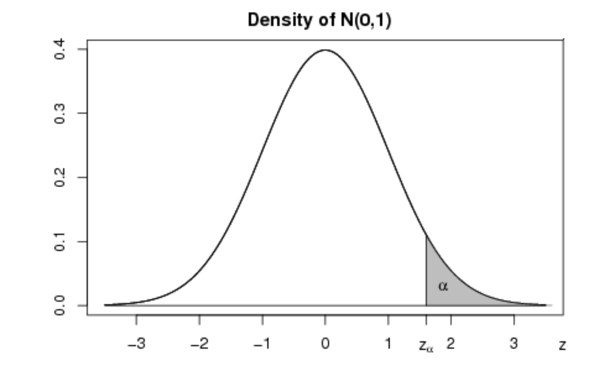
Solution of Example 2



So

$$\frac{c-170}{10/\sqrt{n}} = z_\alpha \quad \rightarrow \quad c = 170 + \frac{10 z_\alpha}{\sqrt{n}}$$

Solution of Example 2



So

$$\frac{c-170}{10/\sqrt{n}} = z_\alpha \quad \rightarrow \quad c = 170 + \frac{10 z_\alpha}{\sqrt{n}}$$

Solution of Example 2

The rejection region is

$$\bar{X} > 170 + \frac{10 z_\alpha}{\sqrt{n}}, \quad \text{or} \quad Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > z_\alpha$$

If $\alpha = 0.05$ then $z_\alpha = 1.645$.

(c) $n = 25$, $\bar{x} = 172.52$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$z = \frac{\bar{x} - 170}{10/\sqrt{n}} = \frac{172.52 - 170}{10/\sqrt{25}} = 1.26$$

(d) Conclusion: Since $z = 1.26 < z_\alpha = 1.645$, don't reject H_0.

Solution of Example 2

The rejection region is

$$\bar{X} > 170 + \frac{10 z_\alpha}{\sqrt{n}}, \quad \text{or} \quad Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > z_\alpha$$

If $\alpha = 0.05$ then $z_\alpha = 1.645$.

(c) $n = 25$, $\bar{x} = 172.52$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$z = \frac{\bar{x} - 170}{10/\sqrt{n}} = \frac{172.52 - 170}{10/\sqrt{25}} = 1.26$$

(d) Conclusion: Since $z = 1.26 < z_\alpha = 1.645$, don't reject H_0.

Solution of Example 2

The rejection region is

$$\bar{X} > 170 + \frac{10z_\alpha}{\sqrt{n}}, \quad \text{or} \quad Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > z_\alpha$$

If $\alpha = 0.05$ then $z_\alpha = 1.645$.

(c) $n = 25$, $\bar{x} = 172.52$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$z = \frac{\bar{x} - 170}{10/\sqrt{n}} = \frac{172.52 - 170}{10/\sqrt{25}} = 1.26$$

(d) Conclusion: Since $z = 1.26 < z_\alpha = 1.645$, don't reject H_0.

Solution of Example 2

The rejection region is

$$\bar{X} > 170 + \frac{10 z_\alpha}{\sqrt{n}}, \qquad \text{or} \quad Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > z_\alpha$$

If $\alpha = 0.05$ then $z_\alpha = 1.645$.

(c) $n = 25$, $\bar{x} = 172.52$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$z = \frac{\bar{x} - 170}{10/\sqrt{n}} = \frac{172.52 - 170}{10/\sqrt{25}} = 1.26$$

(d) Conclusion: Since $z = 1.26 < z_\alpha = 1.645$, don't reject H_0.

Solution of Example 2

The rejection region is

$$\bar{X} > 170 + \frac{10 z_\alpha}{\sqrt{n}}, \quad \text{or} \quad Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > z_\alpha$$

If $\alpha = 0.05$ then $z_\alpha = 1.645$.

(c) $n = 25$, $\bar{x} = 172.52$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$z = \frac{\bar{x} - 170}{10/\sqrt{n}} = \frac{172.52 - 170}{10/\sqrt{25}} = 1.26$$

(d) Conclusion: Since $z = 1.26 < z_\alpha = 1.645$, don't reject H_0.

Solution of Example 2

The rejection region is

$$\bar{X} > 170 + \frac{10z_\alpha}{\sqrt{n}}, \quad \text{or} \quad Z = \frac{\bar{X} - 170}{10/\sqrt{n}} > z_\alpha$$

If $\alpha = 0.05$ then $z_\alpha = 1.645$.

(c) $n = 25$, $\bar{x} = 172.52$, $\alpha = 0.05$, $z_\alpha = 1.645$.

$$z = \frac{\bar{x} - 170}{10/\sqrt{n}} = \frac{172.52 - 170}{10/\sqrt{25}} = 1.26$$

(d) Conclusion: Since $z = 1.26 < z_\alpha = 1.645$, don't reject H_0.

The p-value Approach

- One important statistic which is usually reported in statistical analysis is the **p-value**:
- the probability of rejecting true null H_0 using the observed value of test statistic as the critical value.
- For example, for the above hypotheses $H_0 : p = p_0 \quad \text{vs} \quad H_1 : p < p_0$, if the observed value of Y is y, then the observed value of Z is

$$z = \frac{y - np_0}{\sqrt{np_0(1-p_0)}}$$

- the p-value is

$$p_{value} = P(Z \leqslant z | p = p_0) = P(Y \leqslant y | p = p_0) \approx \Phi\left(\frac{y - np_0}{\sqrt{np_0(1-p_0)}}\right).$$

The p-value Approach

- One important statistic which is usually reported in statistical analysis is the **p-value**:
- the probability of rejecting true null H_0 using the observed value of test statistic as the critical value.
- For example, for the above hypotheses $H_0 : p = p_0 \quad \text{vs} \quad H_1 : p < p_0$, if the observed value of Y is y, then the observed value of Z is

$$z = \frac{y - np_0}{\sqrt{np_0(1-p_0)}}$$

- the p-value is

$$p_{value} = P(Z \leqslant z | p = p_0) = P(Y \leqslant y | p = p_0) \approx \Phi\left(\frac{y - np_0}{\sqrt{np_0(1-p_0)}}\right).$$

The p-value Approach

- One important statistic which is usually reported in statistical analysis is the **p-value**:
- the probability of rejecting true null H_0 using the observed value of test statistic as the critical value.
- For example, for the above hypotheses $H_0 : p = p_0 \quad \text{vs} \quad H_1 : p < p_0$, if the observed value of Y is y, then the observed value of Z is

$$z = \frac{y - np_0}{\sqrt{np_0(1-p_0)}}$$

- the p-value is

$$p_{value} = P(Z \leqslant z | p = p_0) = P(Y \leqslant y | p = p_0)$$
$$\approx \Phi\left(\frac{y - np_0}{\sqrt{np_0(1-p_0)}}\right).$$

The p-value Approach

- One important statistic which is usually reported in statistical analysis is the ***p*-value**:
- the probability of rejecting true null H_0 using the observed value of test statistic as the critical value.
- For example, for the above hypotheses $H_0 : p = p_0 \quad \text{vs} \quad H_1 : p < p_0$, if the observed value of Y is y, then the observed value of Z is

$$z = \frac{y - np_0}{\sqrt{np_0(1 - p_0)}}$$

- the p-value is

$$p_{value} = P(Z \leqslant z | p = p_0) = P(Y \leqslant y | p = p_0) \approx \Phi\left(\frac{y - np_0}{\sqrt{np_0(1 - p_0)}}\right).$$

Decision Based on p-value

- If p-value $< \alpha$, then we reject H_0 at the significance level α. Otherwise accept H_0.
- If in Example 1, the observed value of Y is $y = 4$, then the p-value is

$$p_{value} \approx \Phi\left(\frac{4-12}{3.358571}\right) = 0.0086.$$

- So we can reject H_0 at the significance level $\alpha = 0.05$. We can also reject H_0 at the significance level $\alpha = 0.01$.

Decision Based on p-value

- If p-value $< \alpha$, then we reject H_0 at the significance level α. Otherwise accept H_0.
- If in Example 1, the observed value of Y is $y = 4$, then the p-value is

$$p_{value} \approx \Phi\left(\frac{4-12}{3.358571}\right) = 0.0086.$$

- So we can reject H_0 at the significance level $\alpha = 0.05$. We can also reject H_0 at the significance level $\alpha = 0.01$.

Decision Based on p-value

- If p-value $< \alpha$, then we reject H_0 at the significance level α. Otherwise accept H_0.
- If in Example 1, the observed value of Y is $y = 4$, then the p-value is

$$p_{value} \approx \Phi\left(\frac{4-12}{3.358571}\right) = 0.0086.$$

- So we can reject H_0 at the significance level $\alpha = 0.05$. We can also reject H_0 at the significance level $\alpha = 0.01$.

Decision Based on p-value

- If p-value $< \alpha$, then we reject H_0 at the significance level α. Otherwise accept H_0.
- If in Example 1, the observed value of Y is $y = 4$, then the p-value is

$$p_{value} \approx \Phi\left(\frac{4-12}{3.358571}\right) = 0.0086.$$

- So we can reject H_0 at the significance level $\alpha = 0.05$. We can also reject H_0 at the significance level $\alpha = 0.01$.

P-value–the formal definition

Definition

If W is a test statistic, the *p-value*, or *attained significance level*, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

Let RR_α be the rejection region of significance level α:

$$P(W \in RR_\alpha | H_0) = \alpha.$$

If the data resulted in an observed value w of W, then the

$$p\text{-value} = \min\{\alpha : w \in RR_\alpha\}.$$

Rejection Region: p-value $< \alpha$.

P-value–the formal definition

Definition

If W is a test statistic, the *p-value*, or *attained significance level*, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

Let RR_α be the rejection region of significance level α:

$$P(W \in RR_\alpha | H_0) = \alpha.$$

If the data resulted in an observed value w of W, then the

$$p\text{-value} = \min\{\alpha : w \in RR_\alpha\}.$$

Rejection Region: p-value $< \alpha$.

P-value–the formal definition

Definition

If W is a test statistic, the *p-value*, or *attained significance level*, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

Let RR_α be the rejection region of significance level α:

$$P(W \in RR_\alpha | H_0) = \alpha.$$

If the data resulted in an observed value w of W, then the

$$p\text{-value} = \min\{\alpha : w \in RR_\alpha\}.$$

Rejection Region: p-value $< \alpha$.

Example

A health club claims that its members lose an average of 10 pounds or more within the 1st month after joining the club. A consumer agency wanted to test this claim.

A sample 36 members was selected and an average weight lost of 9.2 pounds was obtained. Assume the population standard deviation is known to be 2.4 pounds. What is the p-value of this test?

Solution:

Step 1. Hypotheses:

$$H_0 : \mu \geqslant 10 \quad H_A : \mu < 10$$

Step 2. Since $\sigma = 2.4$ known, $n = 36 > 30$, use normal distribution and z test statistic $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ and rejection region $RR_\alpha = \{Z < -z_\alpha\}$.

Example

A health club claims that its members lose an average of 10 pounds or more within the 1st month after joining the club. A consumer agency wanted to test this claim.
A sample 36 members was selected and an average weight lost of 9.2 pounds was obtained. Assume the population standard deviation is known to be 2.4 pounds. What is the p-value of this test?

Solution:

Step 1. Hypotheses:

$$H_0 : \mu \geqslant 10 \quad H_A : \mu < 10$$

Step 2. Since $\sigma = 2.4$ known, $n = 36 > 30$, use normal distribution and z test statistic $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ and rejection region $RR_\alpha = \{Z < -z_\alpha\}$.

Step 3. $\bar{x} = 9.2$, the observed value of Z is
$z = (\bar{x} - \mu_0)/(\sigma/\sqrt{n}) = -2.00$.
The p-value is

$$p\text{-value} = \min\{\alpha : -2.00 < -z_\alpha\} = P(Z < -2.00) = 0.0228$$

If $\alpha = 0.01$, then since p-value > 0.01, we do not reject H_0 at the significance level 0.01.
But p-value < 0.05, we reject H_0 at the significance level 0.05.

Step 3. $\bar{x} = 9.2$, the observed value of Z is
$z = (\bar{x} - \mu_0)/(\sigma/\sqrt{n}) = -2.00$.
The p-value is

$$p\text{-value} = \min\{\alpha : -2.00 < -z_\alpha\} = P(Z < -2.00) = 0.0228$$

If $\alpha = 0.01$, then since p-value > 0.01, we do not reject H_0 at the significance level 0.01.
But p-value < 0.05, we reject H_0 at the significance level 0.05.

Step 3. $\bar{x} = 9.2$, the observed value of Z is
$z = (\bar{x} - \mu_0)/(\sigma/\sqrt{n}) = -2.00$.
The p-value is

$$p\text{-value} = \min\{\alpha : -2.00 < -z_\alpha\} = P(Z < -2.00) = 0.0228$$

If $\alpha = 0.01$, then since p-value > 0.01, we do not reject H_0 at the significance level 0.01.
But p-value < 0.05, we reject H_0 at the significance level 0.05.

Step 3. $\bar{x} = 9.2$, the observed value of Z is
$z = (\bar{x} - \mu_0)/(\sigma/\sqrt{n}) = -2.00$.
The p-value is

$$p\text{-value} = \min\{\alpha : -2.00 < -z_\alpha\} = P(Z < -2.00) = 0.0228$$

If $\alpha = 0.01$, then since p-value > 0.01, we do not reject H_0 at the significance level 0.01.
But p-value < 0.05, we reject H_0 at the significance level 0.05.

Finding p-value

Let the observed value of the test statistic be w. If the rejection region is $\{W < k\}$, then the p-value is

$$p - \text{value} = P(W < w | H_0).$$

Given a significance level α, if $p -$ value is smaller than α, we reject the null hypothesis H_0. This is equivalent to the critical value approach.

Click here for tables of the commonly used tests.

Finding p-value

Let the observed value of the test statistic be w. If the rejection region is $\{W < k\}$, then the p-value is

$$p - \text{value} = P(W < w | H_0).$$

Given a significance level α, if $p -$ value is smaller than α, we reject the null hypothesis H_0. This is equivalent to the critical value approach.

Click here for tables of the commonly used tests.

Finding p-value

Let the observed value of the test statistic be w. If the rejection region is $\{W < k\}$, then the p-value is

$$p - \text{value} = P(W < w|H_0).$$

Given a significance level α, if $p -$ value is smaller than α, we reject the null hypothesis H_0. This is equivalent to the critical value approach.
Click here for tables of the commonly used tests.

An Example of One-Sided Test

Example 1. Let x be the Brinell hardness measurement of ductile iron subcritically annealed. Assume that x has a normal distribution with mean μ. Test $H_0 : \mu = 170$ against the alternative hypothesis that the mean hardness is greater than 170 based on the following 25 observations(Use significance level $\alpha = 0.05$.)

```
170, 167, 174, 179, 179, 156, 163, 156, 187
156, 183, 179, 174, 179, 170, 156, 187,
179, 183, 174, 187, 167, 159, 170, 179
```

Hypotheses:

$$H_0 : \mu = 170 \quad H_1 : \mu > 170$$

An Example of One-Sided Test

Example 1. Let x be the Brinell hardness measurement of ductile iron subcritically annealed. Assume that x has a normal distribution with mean μ. Test $H_0 : \mu = 170$ against the alternative hypothesis that the mean hardness is greater than 170 based on the following 25 observations(Use significance level $\alpha = 0.05$.)

```
170, 167, 174, 179, 179, 156, 163, 156, 187
156, 183, 179, 174, 179, 170, 156, 187,
179, 183, 174, 187, 167, 159, 170, 179
```

Hypotheses:

$$H_0 : \mu = 170 \quad H_1 : \mu > 170$$

An Example of Two-Sided Test

Example 2. A company that manufactures brackets for an auto maker selected 15 brackets from the production line and performs a torque test. The goal is for mean torque to equal 125. Let the toque have a normal distribution. The 15 observations are

```
128, 149, 136, 114, 126, 142, 124,
122, 118, 122, 129, 118, 122, 129, 136
```

Hypotheses: $H_0 : \mu = 125 \quad H_1 : \mu \neq 125$

An Example of Two-Sided Test

Example 2. A company that manufactures brackets for an auto maker selected 15 brackets from the production line and performs a torque test. The goal is for mean torque to equal 125. Let the toque have a normal distribution. The 15 observations are

```
128, 149, 136, 114, 126, 142, 124,
122, 118, 122, 129, 118, 122, 129, 136
```

Hypotheses: $H_0 : \mu = 125 \quad H_1 : \mu \neq 125$

Another Example of Two-Sided Test

Example 3. A Casino want to test whether the probability p that a red number comes up on a Nevada roulette wheel equals 18/38 or not.

Hypotheses: $H_0 : p = 18/38 \quad H_1 : p \neq 18/38.$

Another Example of Two-Sided Test

Example 3. A Casino want to test whether the probability p that a red number comes up on a Nevada roulette wheel equals 18/38 or not.

Hypotheses: $H_0 : p = 18/38 \quad H_1 : p \neq 18/38.$