Statistical Methods

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

September 10, 2019

Topics

Objectives

Objectives:

- Compute the Type II error probability and power for the one-sample z test for μ.

Digression: Type I and II Errors and Their Probabilities (Cont'd) (8.1, 8.2)

Type II Error Probabilities and the Power of a Test

- We denote the **probability** of making a **Type II error** (when H_a is true) by β:

$$\beta = P(\text{Type II Error}).$$

Digression: Type I and II Errors and Their Probabilities (Cont'd) (8.1, 8.2)

Type II Error Probabilities and the Power of a Test

- We denote the **probability** of making a **Type II error** (when H_a is true) by β:

$$\beta = P(\text{Type II Error}).$$

- The ***power*** of a test is the **probability** that you ***don't*** make a Type II error (when H_a is true):

$$\text{Power} = 1 - \beta\,.$$

Digression: Type I and II Errors and Their Probabilities (Cont'd) (8.1, 8.2)

Type II Error Probabilities and the Power of a Test

- We denote the **probability** of making a **Type II error** (when H_a is true) by β:

$$\beta = P(\text{Type II Error}).$$

- The ***power*** of a test is the **probability** that you ***don't*** make a Type II error (when H_a is true):

$$\text{Power} = 1 - \beta\,.$$

(i.e. the **probability** that you **reject** $\boldsymbol{H_0}$ when $\boldsymbol{H_a}$ is true).

- Takeaway:
  - The more **power** a test has, the more likely it is to detect **departures from H_0.**

- Takeaway:
  - The more **power** a test has, the more likely it is to detect **departures from H_0.**

    Example: Clinical trial to test the effectiveness of a new drug, i.e.

    $$H_0 : \text{The drug has } \textbf{no effect}$$
    $$H_a : \text{The drug has } \textbf{an effect}$$

    The more **power** the hypothesis test has, the more likely it is to detect **an effect** (if there is one).

Type II Error Probability and Power for the One-Sample z Test

- The **power** is (relatively) easy to compute for the ***one-sample z test***, which is used (instead of the t test) when the **population standard deviation σ** is ***known***.

One-Sample z Test Statistic:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Sampling Distribution of the Test Statistic Under H_0:
If Z is the one-sample z test statistic, then when

$$H_0: \ \mu \ = \ \mu_0$$

is true,

$$Z \ \sim \ N(0,1).$$

- The $N(0,1)$ curve gives us:

- The $N(0, 1)$ curve gives us:
  - The ***rejection region*** as the **extreme 100α% of z values** (in the direction(s) specified by H_a).

- The $N(0,1)$ curve gives us:
  - The ***rejection region*** as the **extreme 100α% of z values** (in the direction(s) specified by H_a).
  - The ***p-value*** as the **tail area(s) beyond the observed z value** (in the direction(s) specified by H_a).

Example

Suppose $X_1, X_2, \ldots, X_n$ are a random sample from a N(μ, σ) distribution, where $\boldsymbol{\sigma}$ **is known**, and we want to test

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_a : \mu &> \mu_0 \end{aligned}$$

The **test statistic** is

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

The **decision rule** (rejection region approach) is

Reject H_0 if $\quad \boldsymbol{Z} > \boldsymbol{z_\alpha}$
Fail to reject H_0 if $\quad \boldsymbol{Z} \leq \boldsymbol{z_\alpha}$

The **decision rule** (rejection region approach) is

Reject H_0 if $\quad \boldsymbol{Z} \; > \; \boldsymbol{z_\alpha}$
Fail to reject H_0 if $\quad \boldsymbol{Z} \; \leq \; \boldsymbol{z_\alpha}$

or equivalently

Reject H_0 if $\quad \bar{\boldsymbol{X}} \; > \; \boldsymbol{\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}}$
Fail to reject H_0 if $\quad \bar{\boldsymbol{X}} \; \leq \; \boldsymbol{\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}}$

The **decision rule** (rejection region approach) is

Reject H_0 if $\quad \boldsymbol{Z} > \boldsymbol{z_\alpha}$
Fail to reject H_0 if $\quad \boldsymbol{Z} \le \boldsymbol{z_\alpha}$

or equivalently

Reject H_0 if $\quad \boldsymbol{\bar{X}} > \boldsymbol{\mu_0} + \boldsymbol{z_\alpha \frac{\sigma}{\sqrt{n}}}$
Fail to reject H_0 if $\quad \boldsymbol{\bar{X}} \le \boldsymbol{\mu_0} + \boldsymbol{z_\alpha \frac{\sigma}{\sqrt{n}}}$

where $\boldsymbol{z_\alpha}$ is the $100(1-\alpha)$th percentile of the N$(0,1)$ distribution.

The **decision rule** (rejection region approach) is

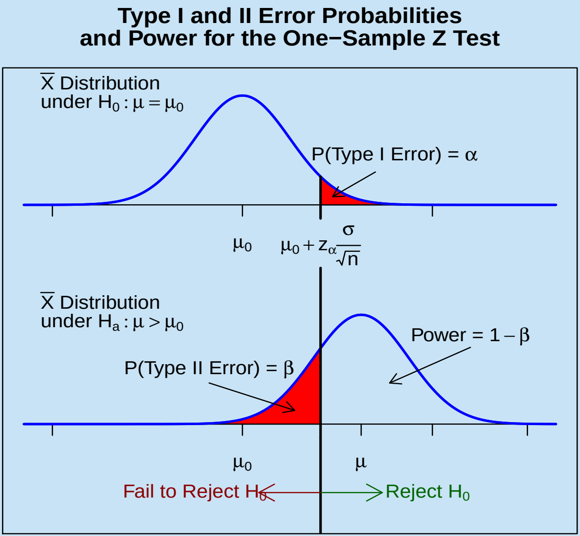
Reject H_0 if $\quad \boldsymbol{Z} > \boldsymbol{z_\alpha}$
Fail to reject H_0 if $\quad \boldsymbol{Z} \leq \boldsymbol{z_\alpha}$

or equivalently

Reject H_0 if $\quad \bar{\boldsymbol{X}} > \boldsymbol{\mu_0} + \boldsymbol{z_\alpha \frac{\sigma}{\sqrt{n}}}$
Fail to reject H_0 if $\quad \bar{\boldsymbol{X}} \leq \boldsymbol{\mu_0} + \boldsymbol{z_\alpha \frac{\sigma}{\sqrt{n}}}$

where z_α is the $100(1-\alpha)$th percentile of the N$(0,1)$ distribution.

The **Type II error probability** β is the *shaded* area under the bottom curve on the next slide. The **power** is the *unshaded* area.


Type I and II Error Probabilities
and Power for the One−Sample Z Test
$\overline{X}$ Distribution
under $H_0 : \mu = \mu_0$
P(Type I Error) = α
μ_0
$\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$
$\overline{X}$ Distribution
under $H_a : \mu > \mu_0$
Power = $1 - \beta$
P(Type II Error) = β
μ_0
μ
Fail to Reject H_0
Reject H_0

To compute β and the **power**, first recall that

$$\bar{X} \sim \text{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right),$$

and so

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

When $H_a : \mu > \mu_0$ is true,

$$\boldsymbol{\beta} \;=\; P(\text{Type II error}) \;=\; P(\textit{Not}\text{ rejecting } H_0)$$

When $H_a : \mu > \mu_0$ is true,

$$\begin{aligned} \boldsymbol{\beta} = P(\text{Type II error}) &= P(\textit{Not}\ \text{rejecting}\ H_0) \\ &= P\left(\bar{X} \leq \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \end{aligned}$$

When $H_a : \mu > \mu_0$ is true,

$$\begin{aligned}\boldsymbol{\beta} = P(\text{Type II error}) &= P(\textit{Not}\text{ rejecting } H_0) \\ &= P\left(\bar{X} \leq \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \\ &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)\end{aligned}$$

When $H_a : \mu > \mu_0$ is true,

$$\begin{aligned} \boldsymbol{\beta} = P(\text{Type II error}) &= P(\textit{Not}\text{ rejecting } H_0) \\ &= P\left(\bar{X} \le \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \le z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \\ &= P\left(Z \le z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \end{aligned}$$

When $H_a : \mu > \mu_0$ is true,

$$\begin{aligned}
\boldsymbol{\beta} &= P(\text{Type II error}) = P(\textit{Not}\text{ rejecting } H_0) \\
&= P\left(\bar{X} \leq \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha + \frac{\mu_0-\mu}{\sigma/\sqrt{n}}\right) \\
&= P\left(Z \leq z_\alpha + \frac{\mu_0-\mu}{\sigma/\sqrt{n}}\right) \\
&= \Phi\left(z_\alpha + \frac{\mu_0-\mu}{\sigma/\sqrt{n}}\right) \qquad (1)
\end{aligned}$$

where $\boldsymbol{\Phi(z) = P(Z \leq z)}$ denotes the **cdf** of the $\boldsymbol{N(0, \ , 1)}$ distribution.

Of course the **power** of the test is

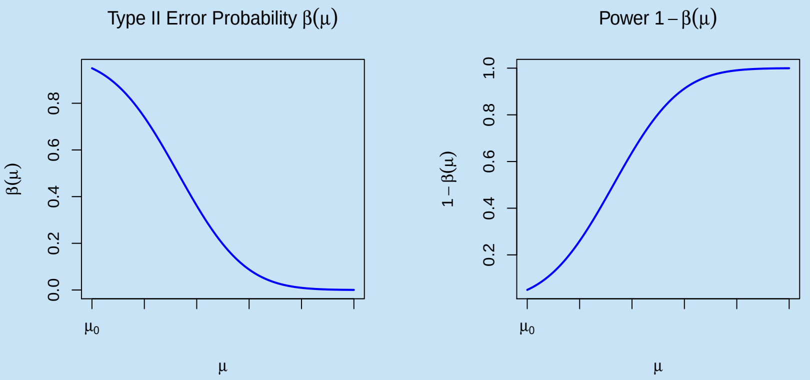
$$\text{Power} \;=\; 1 - \beta\,.$$

Of course the **power** of the test is

$$\text{Power} \;=\; 1-\beta\,.$$

β depends on the value of μ, so it's more appropriate to denote the **Type II error probability** as $\beta(\mu)$ and the **power** as

$$\text{Power} \;=\; 1-\beta(\mu)\,.$$


Type II Error Probability β(μ)
β(μ)
0.0
0.2
0.4
0.6
0.8
μ0
μ
Power 1 – β(μ)
1 – β(μ)
0.2
0.4
0.6
0.8
1.0
μ0
μ

- For lower-tailed and two-tailed z tests, expressions for $\beta(\mu)$ are found in a similar manner.

Alternative Hypothesis	**Type II Error Probability** $\beta(\mu)$	
$H_a : \mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)$	(1)
$H_a : \mu < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)$	(2)
$H_a : \mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right)$	(3)

Exercise

Suppose we have a sample of size $n = 16$ and we want to test

$$\begin{aligned} H_0 : \mu &= 10 \\ H_a : \mu &> 10 \end{aligned}$$

using a level of significance $\alpha = 0.05$. Suppose also that we know that $\sigma = 4$, so the z **test** is appropriate.

Exercise

Suppose we have a sample of size $n = 16$ and we want to test

$$H_0 : \mu = 10$$
$$H_a : \mu > 10$$

using a level of significance $\alpha = 0.05$. Suppose also that we know that $\sigma = 4$, so the z **test** is appropriate.

a) Find $\beta(12)$, the **probability** of a **Type II error** when $\mu = 12$.

Exercise

Suppose we have a sample of size $n = 16$ and we want to test

$$H_0 : \mu = 10$$
$$H_a : \mu > 10$$

using a level of significance $\alpha = 0.05$. Suppose also that we know that $\sigma = 4$, so the z **test** is appropriate.

a) Find $\beta(12)$, the **probability** of a **Type II error** when $\mu = 12$.

Hint: $z_{0.05} = 1.64$, and you should get $\beta(12) = 0.3594$.

Exercise

Suppose we have a sample of size $n = 16$ and we want to test

$$H_0 : \mu = 10$$
$$H_a : \mu > 10$$

using a level of significance $\alpha = 0.05$. Suppose also that we know that $\sigma = 4$, so the z **test** is appropriate.

a) Find $\beta(12)$, the **probability** of a **Type II error** when $\mu = 12$.

Hint: $z_{0.05} = 1.64$, and you should get $\beta(12) = 0.3594$.

b) Find the **power** of the test when $\mu = 12$.

Exercise

Suppose we have a sample of size $n = 16$ and we want to test

$$H_0 : \mu = 10$$
$$H_a : \mu > 10$$

using a level of significance $\alpha = 0.05$. Suppose also that we know that $\sigma = 4$, so the z **test** is appropriate.

a) Find $\beta(12)$, the **probability** of a **Type II error** when $\mu = 12$.

Hint: $z_{0.05} = 1.64$, and you should get $\beta(12) = 0.3594$.

b) Find the **power** of the test when $\mu = 12$.

Hint: You should get $\beta(12) = 0.6406$.

c) Which of the following would you expect to be true?

$$\beta(12) < \beta(13) < \beta(14)$$

or

$$\beta(12) > \beta(13) > \beta(14)$$

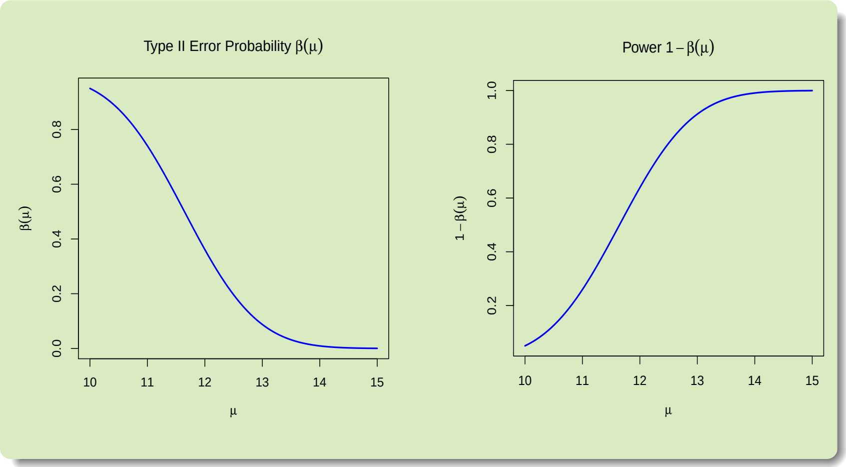
c) Which of the following would you expect to be true?

$$\beta(12) < \beta(13) < \beta(14)$$

or

$$\beta(12) > \beta(13) > \beta(14)$$

Hint: Refer to the graphs of $\beta(\mu)$ and $1 - \beta(\mu)$ on the next slide.


Type II Error Probability β(μ)
β(μ)
0.0
0.2
0.4
0.6
0.8
10
11
12
13
14
15
μ
Power 1 – β(μ)
1 – β(μ)
0.2
0.4
0.6
0.8
1.0
10
11
12
13
14
15
μ

- For t **tests** (and **other tests**), expressions for $\beta(\mu)$ are more complicated.

- For t **tests** (and **other tests**), expressions for $\beta(\mu)$ are more complicated.

  (See the textbook.)

- For t **tests** (and **other tests**), expressions for $\beta(\mu)$ are more complicated.

  (See the textbook.)

  The **Type II error probability** and **power** can be obtained using software.

The Sample Size Needed to Attain a Desired Type II Error Probability for the One-Sample z Test

- For **fixed** μ, as n **increases**, $\beta(\mu)$ **decreases**, and therefore the **power increases** (see (1), (2), and (3)).

The Sample Size Needed to Attain a Desired Type II Error Probability for the One-Sample z Test

- For **fixed** μ, as n **increases**, $\beta(\mu)$ **decreases**, and therefore the **power increases** (see (1), (2), and (3)).

  A **larger sample size** results in a **more powerful** test.

The Sample Size Needed to Attain a Desired Type II Error Probability for the One-Sample z Test

- For **fixed** μ, as n **increases**, $\beta(\mu)$ **decreases**, and therefore the **power increases** (see (1), (2), and (3)).

  A **larger sample size** results in a **more powerful** test.

- We can determine **how big** n **needs to be** in order to attain a **desired** $\beta(\mu)$ (for any specified value of μ).

The Sample Size Needed to Attain a Desired Type II Error Probability for the One-Sample z Test

- For **fixed** μ, as n **increases**, $\beta(\mu)$ **decreases**, and therefore the **power increases** (see (1), (2), and (3)).

  A **larger sample size** results in a **more powerful** test.

- We can determine **how big** n **needs to be** in order to attain a **desired** $\beta(\mu)$ (for any specified value of μ).

  (See the textbook or the **optional** sides ahead for details).

(Optional Slide)

Example

Consider again a test of

$$\begin{aligned} H_0 : \mu &= 10 \\ H_a : \mu &> 10 \end{aligned}$$

using a level of significance $\alpha = 0.05$, and that $\boldsymbol{\sigma = 4}$.

(Optional Slide)

Example

Consider again a test of

$$\begin{aligned} H_0 : \mu &= 10 \\ H_a : \mu &> 10 \end{aligned}$$

using a level of significance $\alpha = 0.05$, and that $\boldsymbol{\sigma} = \mathbf{4}$.

Now suppose we want n to be **large enough** so that the **Type II error probability** when $\mu = 12$ is **0.20**, which will give a **power** of **0.80**.

(Optional Slide)

We need n to be **large enough** that

$$\beta(12) \; = \; 0.20.$$

(Optional Slide)

We need n to be **large enough** that

$$\beta(12) \;=\; 0.20.$$

From (1), this means we need

$$0.20 \;=\; \Phi\left(z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \;=\; \Phi\left(1.64 + \frac{10 - 12}{4/\sqrt{n}}\right).$$

(Optional Slide)

We need n to be **large enough** that

$$\beta(12) \;=\; 0.20.$$

From (1), this means we need

$$0.20 \;=\; \Phi\left(z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \;=\; \Phi\left(1.64 + \frac{10-12}{4/\sqrt{n}}\right).$$

But from a $N(0,1)$ table,

$$0.20 \;=\; \Phi(-0.84)$$

(i.e. $z_{0.20} = 0.84$).

(Optional Slide)

We need n to be **large enough** that

$$\beta(12) \;=\; 0.20.$$

From (1), this means we need

$$0.20 \;=\; \Phi\left(z_\alpha + \frac{\mu_0 - \mu}{\sigma/\sqrt{n}}\right) \;=\; \Phi\left(1.64 + \frac{10-12}{4/\sqrt{n}}\right).$$

But from a $N(0,1)$ table,

$$0.20 \;=\; \Phi(-0.84)$$

(i.e. $z_{0.20} = 0.84$). This implies

$$1.64 + \frac{10-12}{4/\sqrt{n}} \;=\; -0.84.$$

(Optional Slide)

Solving for n gives

$$n = \left(\frac{4(1.64 + 0.84)}{10 - 12}\right)^2 = 24.60, \tag{5}$$

(Optional Slide)

Solving for n gives

$$n = \left(\frac{4(1.64 + 0.84)}{10 - 12} \right)^2 = 24.60, \tag{5}$$

so we'd need a **sample size** of **at least** $\mathbf{n = 25}$.

(Optional Slide)

- Generalizing from (5):

Sample Size for a Desired Type II Error Probability: The sample size n for which a level α test has a desired Type II error probability β at the specified value μ is

$$n = \begin{cases} \left(\frac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu}\right)^2 & \text{for a one-tailed (upper or lower) test} \\ \left(\frac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu}\right)^2 & \text{for a two-tailed test} \end{cases} \tag{6}$$

(Optional Slide)

Exercise

Consider again a test of

$$\begin{aligned} H_0 : \mu &= 10 \\ H_a : \mu &> 10 \end{aligned}$$

using $\alpha = 0.05$, and suppose again that $\boldsymbol{\sigma} = 4$.

(Optional Slide)

Exercise

Consider again a test of

$$H_0 : \mu = 10$$
$$H_a : \mu > 10$$

using $\alpha = 0.05$, and suppose again that $\boldsymbol{\sigma} = 4$.

a) Use (6) to determine how big n needs to be for the **Type II error probability** to be **0.10** when $\mu = 11$.

(Optional Slide)

Exercise

Consider again a test of

$$H_0 : \mu = 10$$
$$H_a : \mu > 10$$

using $\alpha = 0.05$, and suppose again that $\boldsymbol{\sigma = 4}$.

a) Use (6) to determine how big n needs to be for the **Type II error probability** to be **0.10** when $\boldsymbol{\mu = 11}$.

Hint: $z_{0.10} = \mathbf{1.28}$ and $z_{0.05} = \mathbf{1.64}$.

(Optional Slide)

b) Use (6) to determine how big n needs to be for the **Type II error probability** to be **0.10** when $\mu = 12$.

(Optional Slide)

b) Use (6) to determine how big n needs to be for the **Type II error probability** to be **0.10** when $\mu = 12$.

c) If we want the **Type II error probability** to be **0.10** when $\mu = 13$, will n need to be **larger** or **smaller** than the one in Part *b*?

(Optional Slide)

b) Use (6) to determine how big n needs to be for the **Type II error probability** to be **0.10** when $\mu = 12$.

c) If we want the **Type II error probability** to be **0.10** when $\mu = 13$, will n need to be **larger** or **smaller** than the one in Part *b*?

d) If we want the **Type II error probability** to be **0.20** when $\mu = 12$, will n need to be **larger** or **smaller** than the one in Part *b*?

(Optional Slide)

- For t **tests** (and **other tests**), expressions for the **sample size** n needed to attain a desired **Type II error probability** or **power** are more complicated.

(Optional Slide)

- For t **tests** (and **other tests**), expressions for the **sample size** n needed to attain a desired **Type II error probability** or **power** are more complicated.

  (See the textbook.)

(Optional Slide)

- For t **tests** (and **other tests**), expressions for the **sample size** n needed to attain a desired **Type II error probability** or **power** are more complicated.

  (See the textbook.)

  The **sample size** n can be obtained using software.