# Null and Alternative Hypotheses (Continued)

**Null hypothesis:**

- Hypothesis that we formulate with the hope of rejecting, denoted by $H_0$.
- A null hypothesis concerning a population parameter will always be stated to specify an exact value of the parameter.

**Alternative hypothesis:**

- The rejection of $H_0$ leads to the acceptance of an alternative hypothesis, denoted by $H_1$.
- It allows for the possibility of several values.



✓ For a valid set of hypotheses (i.e., $H_0$ versus $H_1$), they need to be disjoint. For example "$H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$" is a valid set; "$H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$" is another valid set.

✓ By convention (and theoretically supported), we usually write the null hypothesis in the "equal" form, i.e., $H_0 : \theta = \theta_0$, with $\theta_0$ being a given value. So the hypotheses have three possible forms:

★ $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$; in this case $H_0 : \theta_0 = \theta_0$ in fact means $\theta \leq \theta_0$;

★ $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$; in this case $H_0 : \theta_0 = \theta_0$ in fact means $\theta \geq \theta_0$;

★ $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

So we need to check the form of $H_1$ to ensure the meaning of $H_0 : \theta = \theta_0$ in practice.

## 7.1.2 Types of Error

- Two types of errors in the hypothesis testing:

| | State of Nature | |
|---|---|---|
| **Decision** | $H_0$ is true | $H_0$ is false |
| Reject $H_0$ | **Type I error** Pr(Reject $H_0$ given that $H_0$ is true) = α | Correct decision Pr(Reject $H_0$ given that $H_0$ is false) = 1 – β |
| Do not reject $H_0$ | Correct decision Pr(Do not reject $H_0$ given that $H_0$ is true) = 1 – α | **Type II error** Pr(Do not reject $H_0$ given that $H_0$ is false) = β |



This page gives us a fundamental explanation of the outcomes of the hypothesis testing.

✓ Type I error is usually treated more seriously; we need to control it in the first place. We set an $\alpha$, called the significant level, and require the decision rule satisfy:

$$Pr\left(\text{Reject } H_0 \middle| H_0 \text{ is true }\right) \leq \alpha. \tag{1}$$

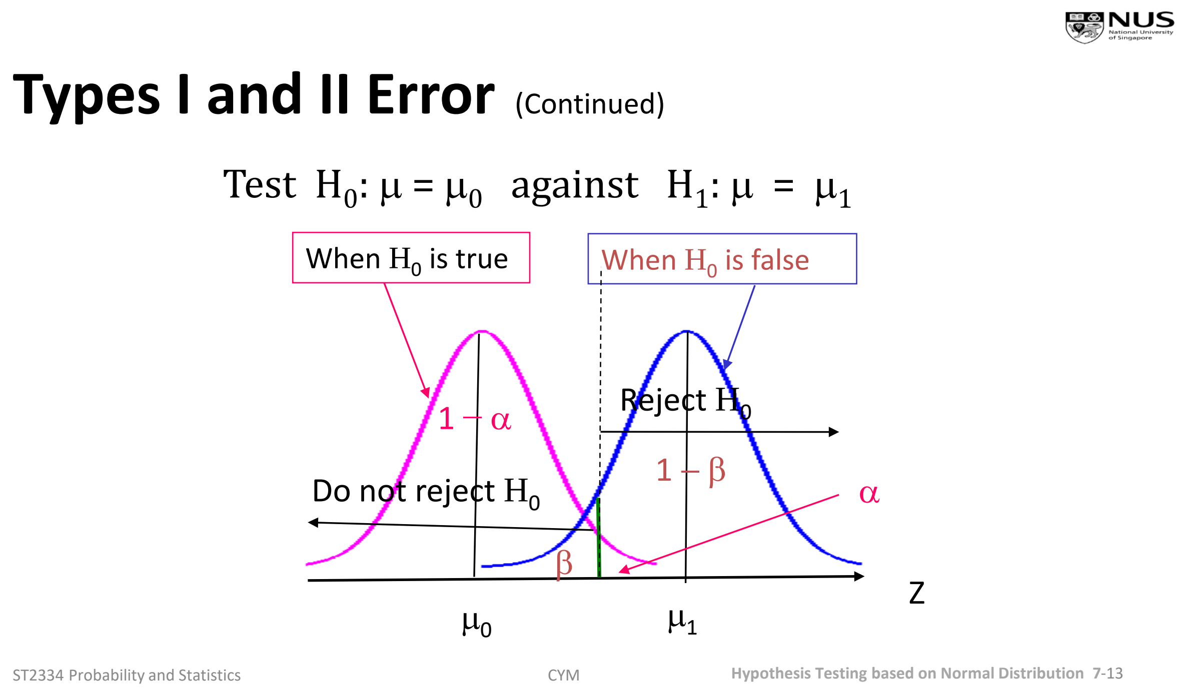
Note that "given $H_0$ is true" in the statement essentially tells that "given $\theta = \theta_0$ is true". This is important when we are to derive the distribution of the test statistic, and it is why we write $H_0 : \theta = \theta_0$.

✓ When type I error is controlled by $\alpha$, i.e., (1) is satisfied, we try to minimize type II error. So, we typically require

$$Pr\left(\text{Reject } H_0 \middle| H_0 \text{ is true }\right) = \alpha,$$

so that the type I error is controlled but maximized to reduce the type II error.

✓ When establishing a testing rule, type I and type II errors trade off each other. Page 7-13 gives a clear explanation of the idea:




Think of how type I and type II errors may change if we push the green line, which defines the rejection/acceptance region of the decision rule, to the left or the right.

✓ When a test is performed, one can make only one type of error, but not both:

★ If $H_0$ is rejected, type I error is possible.

★ If $H_0$ is not rejected, type II error is possible.

Test statistic plays a key role in performing the hypothesis testing. Test statistic must be a function of the sample, e.g., $X_1, X_2, \ldots, X_n$, and does not rely on any unknown parameter.

Similarly to the construction of the confidence interval, we can summarize the procedure for constructing the test statistic for mean related hypothesis tests as follows. Denote by $\theta$ the parameter of interest. We consider the hypotheses:

$$H_0 : \theta = \theta_0 \qquad \text{versus} \qquad H_1 : \ldots$$

where $\theta_0$ is a given value, e.g., 0, which is assumed to be the true value of $\theta$ in developing the distribution of the test statistic. $H_1$ could be $\theta < \theta_0$, $\theta > \theta_0$, or $\theta \neq \theta_0$.

✓ Step 1: Look for an estimator $\widehat{\theta}$ for $\theta$, e.g., $\bar{X}$ for $\mu$.

✓ Step 2: Derive the formula for $V(\widehat{\theta})$.

✓ Step 3: The test statistic is constructed to be

$$T = \frac{\widehat{\theta} - \theta_0}{\sqrt{V}}. \tag{2}$$

★ If $V(\widehat{\theta})$ does not depend on any unknown parameter, e.g., when $\sigma^2$ is known, $V(\bar{X}) = \sigma^2/n$, we set $V = V(\widehat{\theta})$. The statistic $T$ given in (2) (approximately) follows the $N(0, 1)$ distribution, when the data are normal or the sample size is sufficiently large.

Sections 7.2.1 and 7.3.1 belong to this situation.

★ If $V(\widehat{\theta})$ contains some other unknown parameters, e.g., $\sigma^2$, we replace the parameter with its estimator, e.g., $S^2$ can be used to replace $\sigma^2$, and result in $\widehat{V}(\widehat{\theta})$. We set $V = \widehat{V}(\widehat{\theta})$. The distribution of $T$ given in (2) has two possibilities:

(1) The sample size is sufficiently large; then $T \sim N(0, 1)$ approximately.

Section 7.3.2 belongs to this situation. The test problem in Sections 7.2.2 and 7.3.4 can also be solved by this situation if their corresponding sample sizes are sufficiently large, and if so we don't need to require the distribution of the observations are normal.

(2) If the sample size is small, but the observations are normally distributed, then $T \sim t(df)$, where $df$ is the degrees of the freedom of the parameter estimated in $V(\widehat{\theta})$.

The testing problems discussed in Sections 7.2.2, 7.3.3, and 7.3.4 belong to this situation.

The above strategy is not applicable to construct test statistic for the variance related tests.

Page 7-14 & 15

## 7.1.3 Acceptance and Rejection Regions

- To test a hypothesis about a population parameter, we first select a suitable test statistic for the parameter under the hypothesis.
- Once the significance level, $\alpha$, is given, a decision rule can be found such that it divides the set of all possible values of the test statistic into two regions,
- one being the **rejection region** (or critical region) and the other the **acceptance region**.

ST2334 Probability and Statistics CYM Hypothesis Testing based on Normal Distribution 7-14

## Acceptance and Rejection Regions (Continued)

- Once a sample is taken, the value of the test statistic is obtained.
- If the test statistic assumes a value in the rejection region, the null hypothesis is rejected; otherwise it is not rejected.
- The value that separates the rejection and acceptance regions is called the **critical value**.

ST2334 Probability and Statistics CYM Hypothesis Testing based on Normal Distribution 7-15

✓ Acceptance and rejection regions are defined based on the test statistic so that the type I error is controlled and type II error is minimized.

✓ They are disjoint, and they together contain all the possible values of the test statistic.

✓ When constructing the the acceptance and rejection regions, one should jointly view the hypotheses (mainly the alternative) and how the corresponding test statistic will perform when $H_0$ or $H_1$ is true.

For example,

★ Suppose we are to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$, and use $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ as the test statistic, where $\sigma^2$ is assumed to be known.

★ When $H_1$ is true, $Z$ should be "stay away" from 0 as $\bar{X}$ should be stay away from $\mu_0$ when $\mu \neq \mu_0$. So the rejection region must be $Z < z_1$ and $Z > z_2$, with $z_1 < 0$ and $z_2 > 0$. Because $N(0,1)$ is symmetric about 0, therefore, we set $z_1 = -z_2 \equiv -z$ with $z > 0$.

★ To get type I error controlled by $\alpha$, it is equivalent to

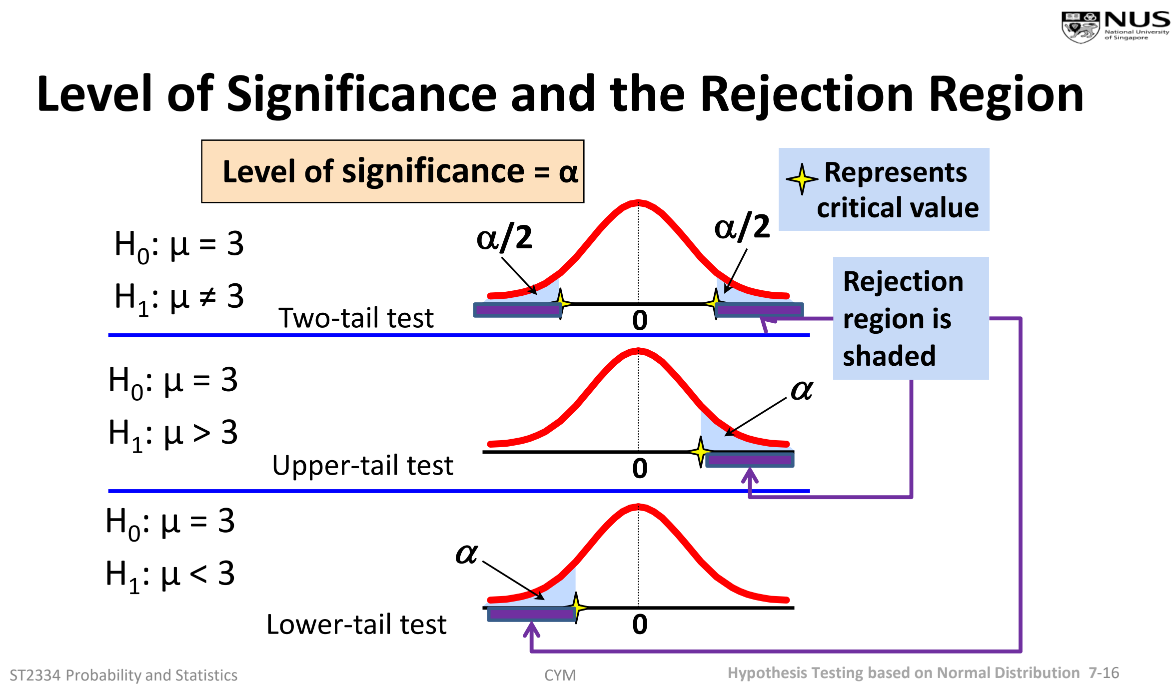
$$\alpha \geq Pr(Z < -z \quad or \quad Z > z) = 2Pr(Z > z).$$

Therefore we need $z = z_{\alpha/2}$.

★ The rejection region should be $|z_{\text{obs}}| > z_{\alpha/2}$, where $z_{\text{obs}}$ denotes the computed value of $Z$ when observations are obtained.

With the similar strategy, we can obtain the rejection region if "$H_1 : \mu > \mu_0$": $z_{\text{obs}} > z_\alpha$; and if "$H_1 : \mu < \mu_0$": $z_{\text{obs}} < -z_\alpha$.

The idea has been well summarized as a figure in page 7-16 of the lecture slide.




Think of how the rejection/acceptance region might be changed if we use $Z_1 = \frac{\mu_0 - \bar{X}}{\sigma^2/\sqrt{n}} \sim N(0,1)$ as the test statistic instead.

Page 7-36

## *p*-Value Approach to Testing

- *p*-value: Probability of obtaining a test statistic more extreme ( ≤ or ≥ ) than the observed sample value given $H_0$ is true
  - Also called **observed level of significance**

ST2334 Probability and Statistics CYM Hypothesis Testing based on Normal Distribution 7-36

✓ p-value is the probability that we may obtain a test statistic value that is as extreme or more extreme than the observed value of the test statistic, when $H_0$ is true.

✓ So the smaller the p-value, the better that we reject $H_0$, as less likely we shall make such an extreme observation given $H_0$ is true.

✓ Given a significant level $\alpha$, we reject $H_0$ when p-value$< \alpha$; and do not reject $H_0$ otherwise.

✓ Using p-value or the rejection region approach to perform the test must result in exactly the same conclusion (in terms of reject or do not reject $H_0$).

★ p-value$< \alpha$ if and only if the test statistic is in the rejection region.

★ p-value$\geq \alpha$ if and only if the test statistics is in the acceptance region.

This idea is well illustrated by page 7-38 of the lecture slides:

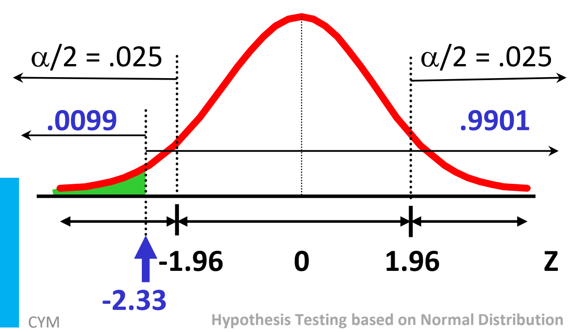
# Example 1 (Continued)

- How likely is it to see a sample mean of 34.5 (or something further from the mean, in either direction) if the true mean is 35? ( $\sigma = 1.5$ and $n = 49$)

$\bar{X}$ = 34.5 is translated to a Z score of Z = -2.33

Pr($Z$ < -2.33) = 0.0099
Pr($Z$ > -2.33) = 0.9901

***p*-value**
**= 2 min{Pr($Z$ < -2.33), Pr(Z >-2.33) }**
**= 2(0.0099) = .0198**



Page 7-34

## Relationship between two-sided test and confidence interval

- The two-sided test procedure just described is equivalent to finding a $(1-\alpha)100\%$ confidence interval for $\mu$
- $H_0$ is accepted if the confidence interval covers $\mu_0$.
- If the C.I. does not cover $\mu_0$, we reject $\mu = \mu_0$ in favour of the alternative $H_1$: $\mu \neq \mu_0$ since

$$\Pr\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$\Leftrightarrow \quad \Pr\left(\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$



Confidence interval is (approximately; see the example in the next slide) equivalent to the hypothesis testing if and only if the latter is two sided; this is applicable not only for the mean related inference but also for the variance related inference. The theoretical foundation is given in the formulation at the bottom of the slide above. Use $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ as an example:

- ✓ If the confidence interval contains $\mu_0$, we do not reject $H_0$.
- ✓ If the confidence interval does not contain $\mu_0$, we reject $H_0$.

Furthermore, note that

$$\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

is equivalent to $|Z| > z_{\alpha/2}$ with $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$.

One example: solve the following problem.

✓ The observation $X$ follows a binomial distribution: $X \sim Bin(200, p)$.

✓ Establish a testing procedure for the following hypotheses:

$$H_0 : p = 0.5 \qquad H_1 : p \neq 0.5.$$

✓ If we observe $X = 90$, perform the above test.