Design of Experiments

(Lecture IV)

Dr. Adham Ragab

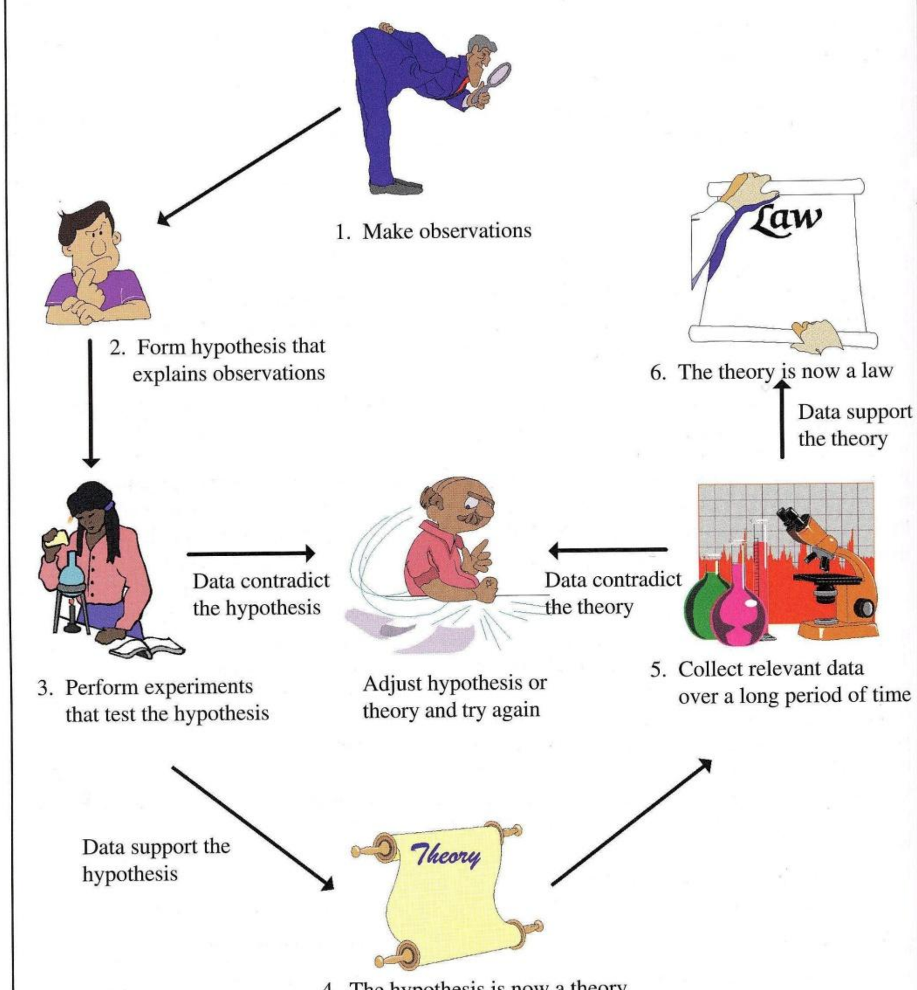
Outline

Objectives

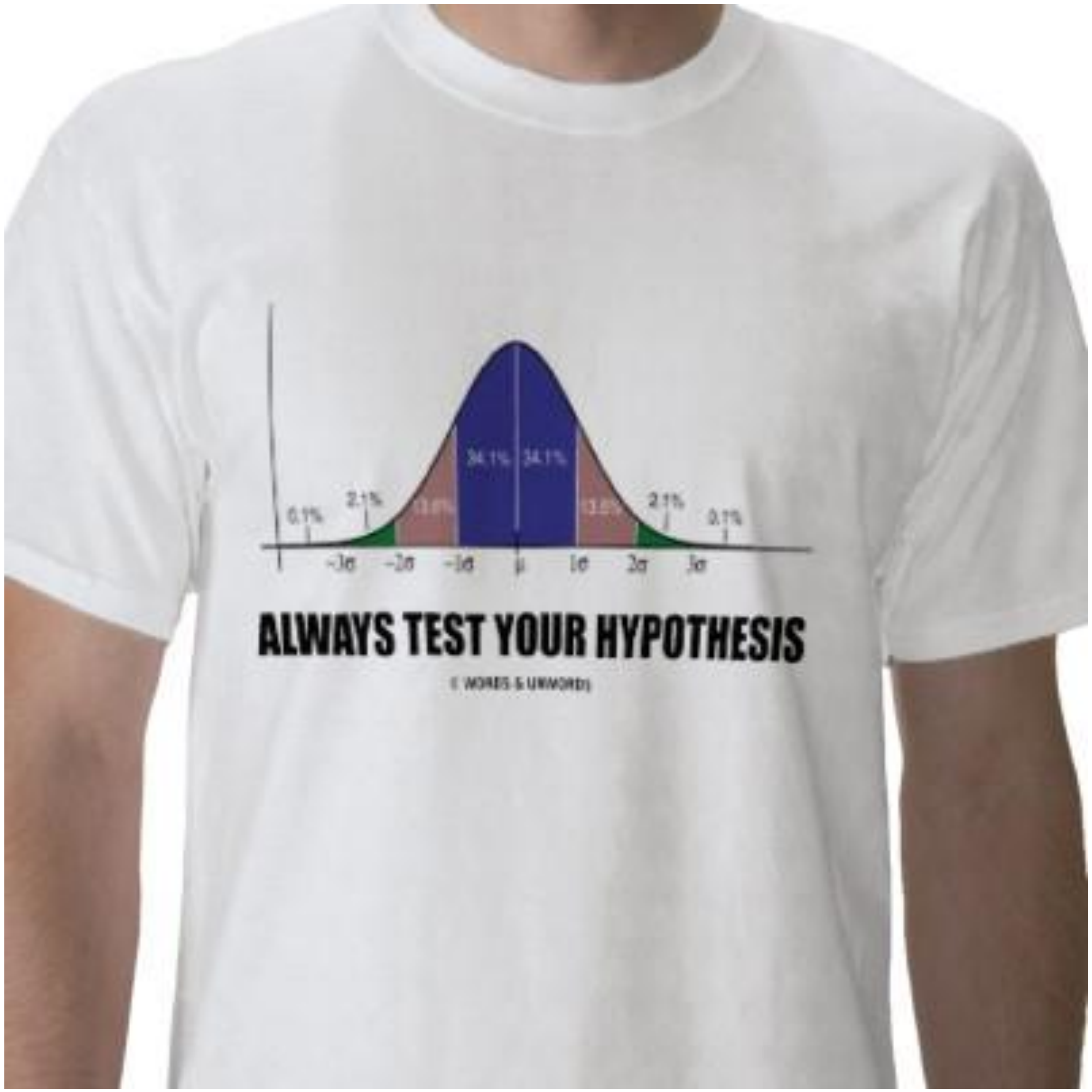
- By the end of this lecture the student should be able to:
  - Conduct tests of hypothesis under different situations
  - Estimating the confidence level.
  - Use Excel in hypothesis testing

What is hypothesis?

FIGURE 2.2
The Scientific Method



What is test of hypothesis?



The Hypothesis Testing Framework

- **Statistical hypothesis testing** is a useful framework for many experimental situations
- Origins of the methodology date from the early 1900s

The Hypothesis Testing Framework

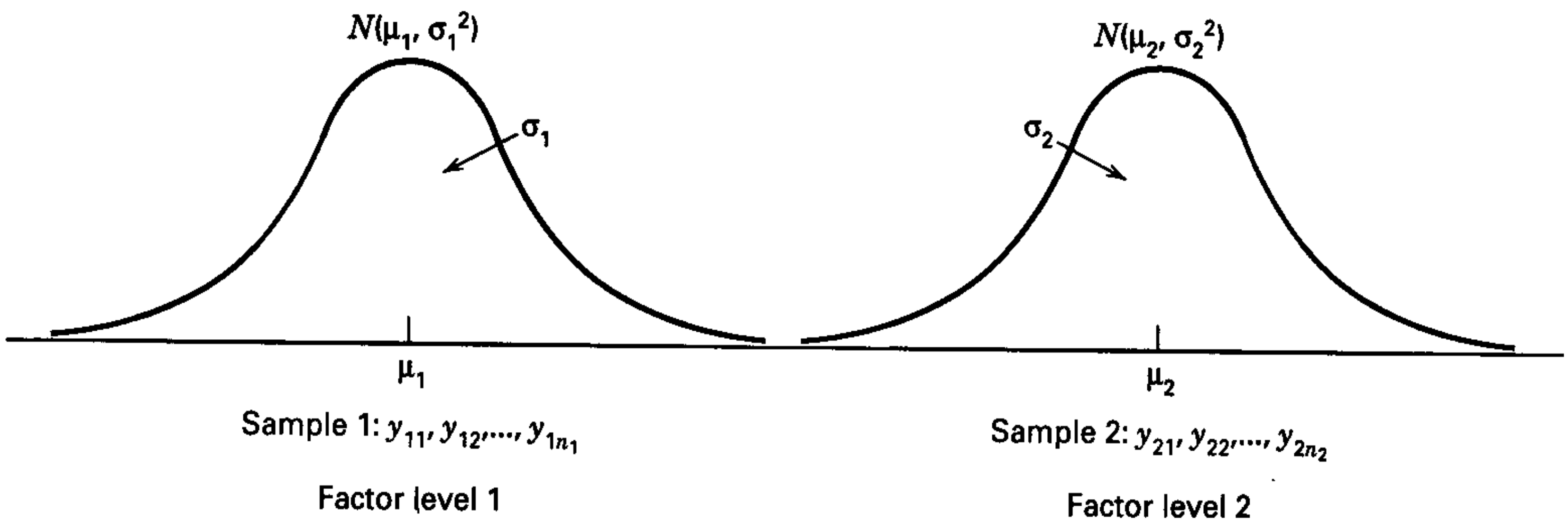


Figure 2-9 The sampling situation for the two-sample *t*-test.

- Sampling from a **normal** distribution
- Statistical hypotheses:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Estimation of Parameters

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ estimates the population mean } \mu$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2 \text{ estimates the variance } \sigma^2$$

The following examples are copied from the University of North Carolina web page.

Example 1

An insurance company is reviewing its current policy rates. When originally setting the rates they believed that the average claim amount was $1,800. They are concerned that the true mean is actually higher than this, which would cause them to lose a lot of money.
They randomly selected 40 claims, and calculated a sample mean of $1,950.
Assuming that the standard deviation of claims is $500, and set confidence level to 95%, test to see if the insurance company should be concerned.

Solution

- Step 1: Set the null and alternative hypotheses

$$H_0 : \ \mu \leq 1800$$

$$H_1 : \ \mu > 1800$$

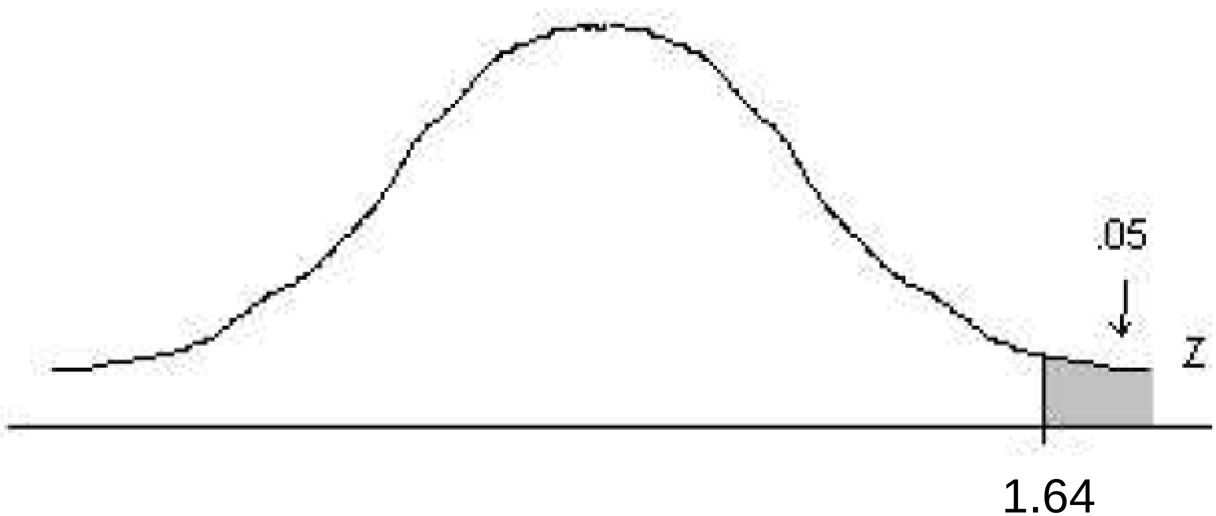
- Step 2: Calculate the test statistic

$$Z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{1950 - 1800}{500 / \sqrt{40}} = 1.897$$

Solution (cont.)

- Step 3: Set Rejection Region
  Looking at the the picture below, we need to put all of α in the right tail. Thus,

$$R: Z > 1.64$$



Solution (cont.)

Step 4: Conclude

We can see that 1*.897 > 1.64, thus our test statistic is in the rejection region. Therefore we have* to reject the null hypothesis. We can tell the insurance company that they should be concerned about their current policies.

Example 2

- Trying to encourage people to stop driving to campus, the university claims that on average it takes people 30 minutes to find a parking space on campus. I don't think it takes so long to find a spot.
- In fact I have a sample of the last five times I drove to campus, and I calculated $x = 20$. *Assuming that the time* it takes to find a parking spot is normal, and that $\sigma = 6$ *minutes, then perform a hypothesis test with level* $\alpha = 0.10$ *to see if my claim is correct.*

Solution (cont.)

- *Step 1: Set the null and alternative hypotheses*

  $H0 : \mu \geq 30$

  $H1 : \mu < 30$

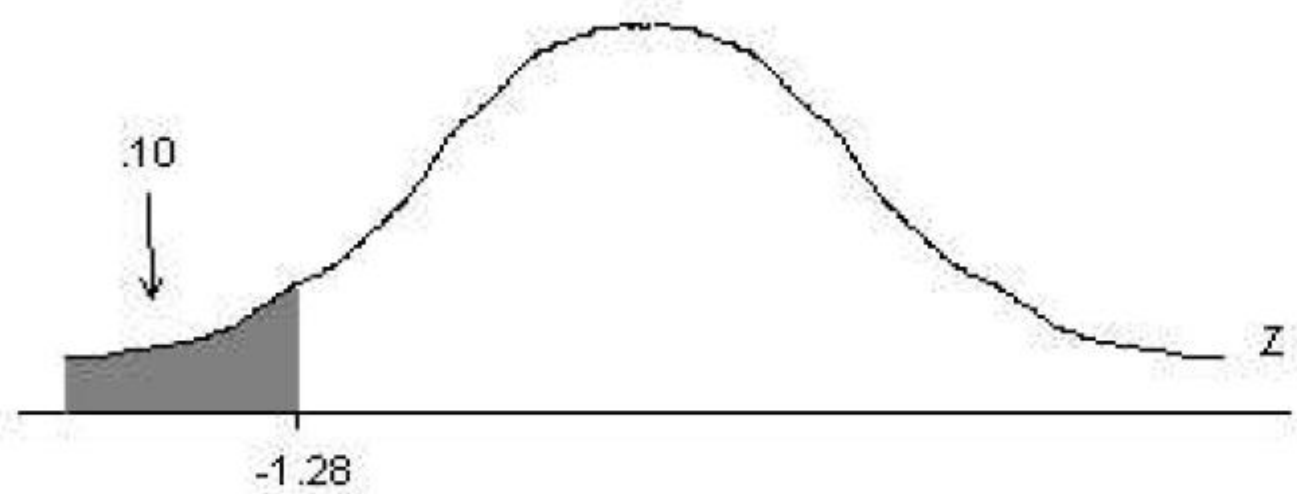
- Step 2: Calculate the test statistic

$$Z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{20 - 30}{6/\sqrt{5}} = -3.727$$

Solution (cont.)

- Step 3: Set Rejection Region
  Looking at the the picture below, we need to put all of α in the left tail. Thus,

$$R : Z < -1.28$$



Solution (cont.)

- Step 4: Conclude
- We can see that *$-3.727 < -1.28$, thus our test statistic is in the rejection region. Therefore we reject* the null hypothesis in favor of the alternative.
- We can conclude that the mean is significantly less than 30, thus I have proven that the mean time to find a parking space is less than 30.

Example 3

- A sample of 40 sales receipts from a grocery store has x = *$137 and* σ = *$30.2. Use these values to test* whether or not the mean of sales at the grocery store are different from $150. $\alpha = 0.01$.

Solution

- Step 1: Set the null and alternative hypotheses

$$H0 : \mu = 150$$

$$H1 : \mu \neq 150$$

- *Step 2 : Calculate the test statistic*

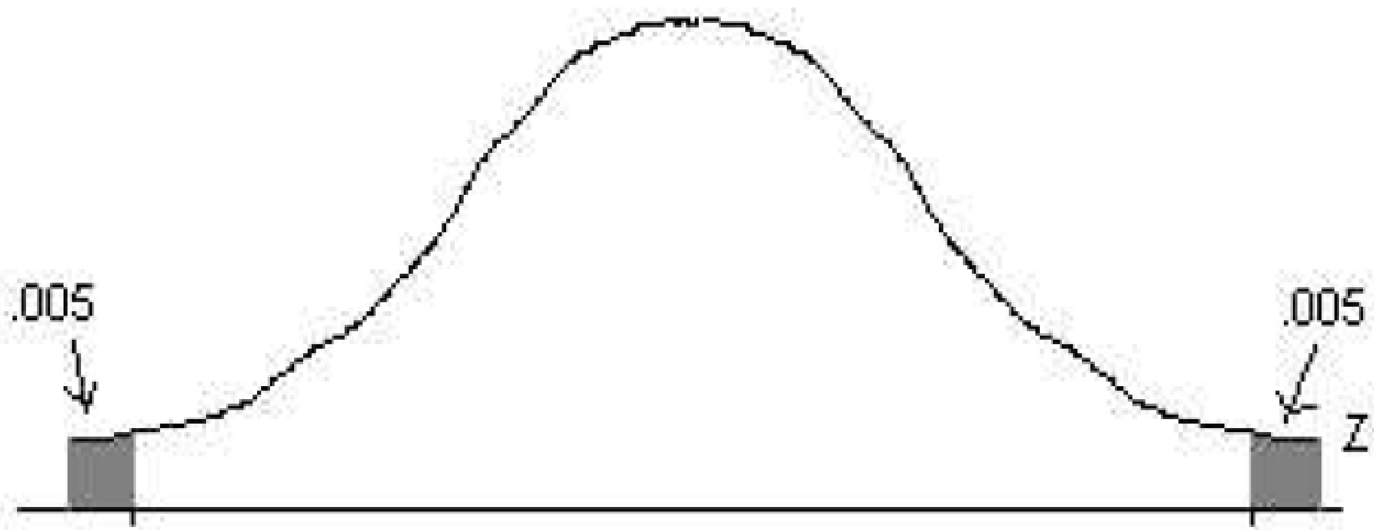
$$Z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{137 - 150}{30.2/\sqrt{40}} = -2.722$$

Solution (cont.)

- Step 3: Set rejection region.
- Looking at the picture below, we need to put half of α *in the left tail, and the other half of α in* the right tail. Thus,

$$R : ABS(Z) > 2.58$$

Solution (cont.)



Step 4: Conclude

We can see that *-2.722 < -2:58, thus our test statistic is in the rejection region. Therefore* we reject the null hypothesis in favor of the alternative. We can conclude that the mean is significantly different from $150, thus I have proven that the mean sales at the grocery store is not $150.