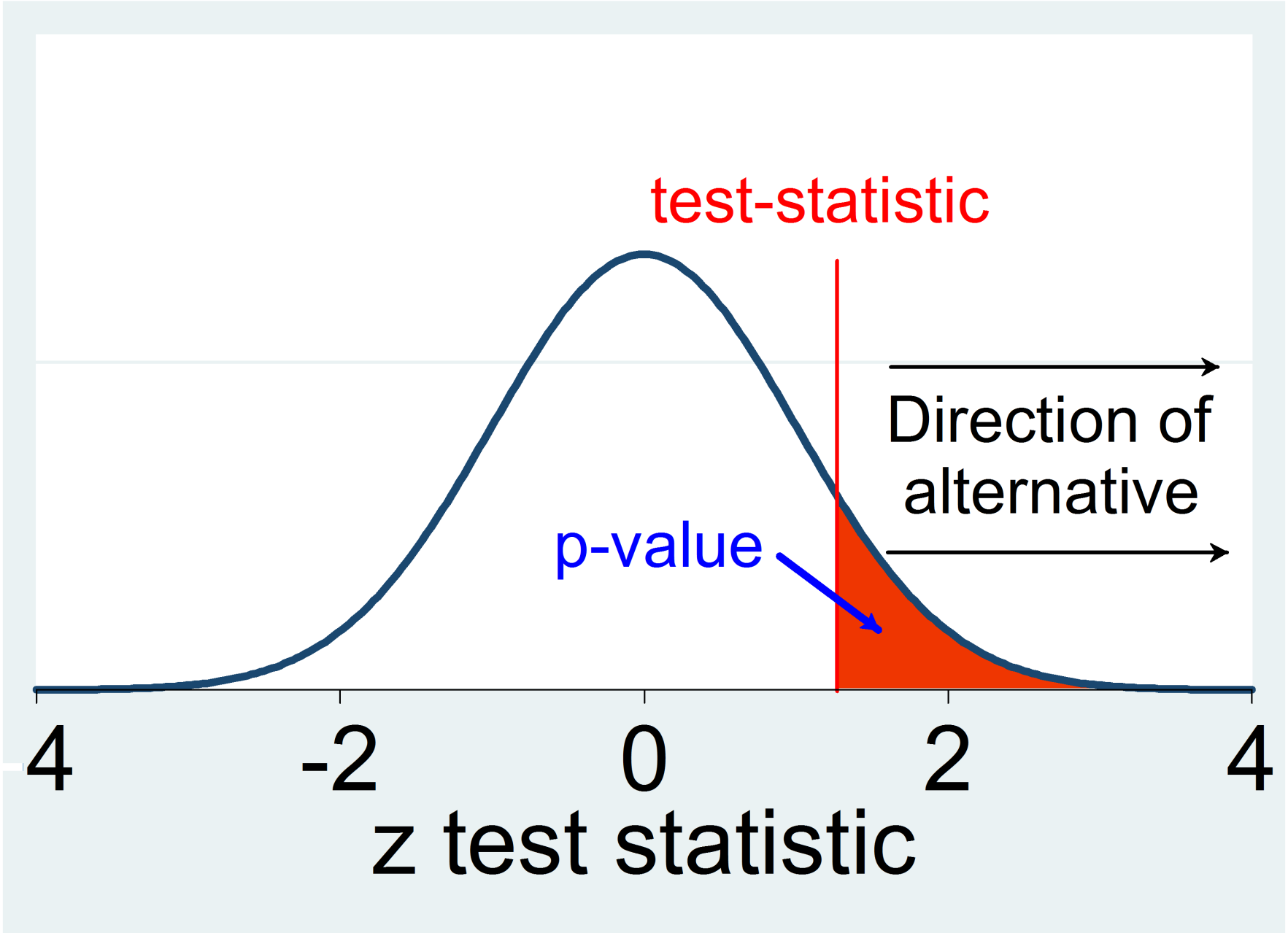
ECO220Y
Hypothesis Tests: *P*-Value Approach
Readings: Chapter 12

Winter 2012

Lecture 14

Rejection Region Approach - Shortcomings

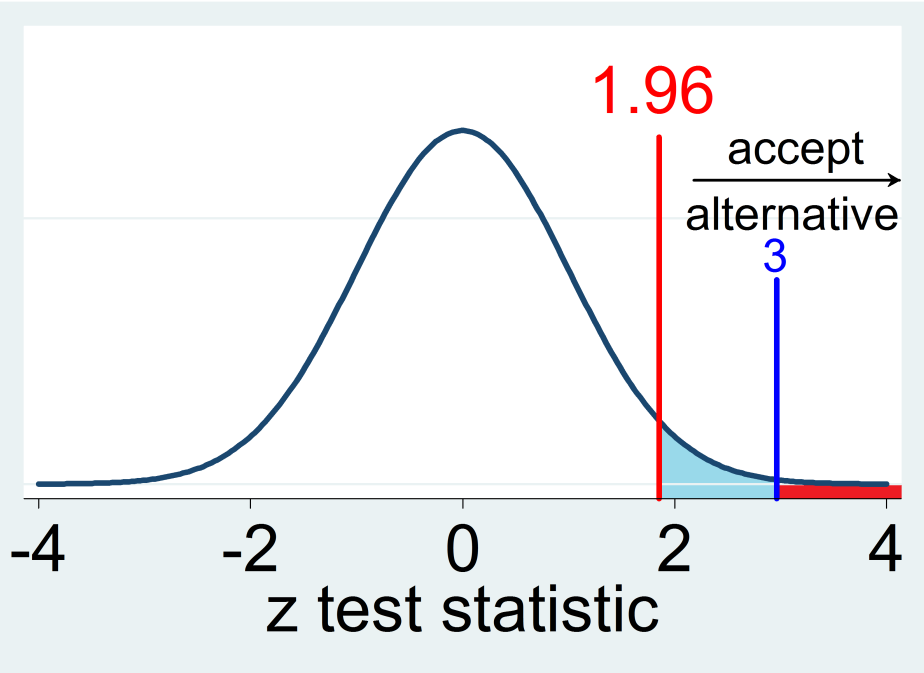
- Rejection region approach has only two outcomes - yes or no.
- However, results from some tests are stronger than from the others.
- Think of the test statistic being very close to or very far from the critical value.
- To take advantage of information available from test statistic, we need a better measure of the statistical evidence supporting alternative hypothesis.
- Solution: P-value approach

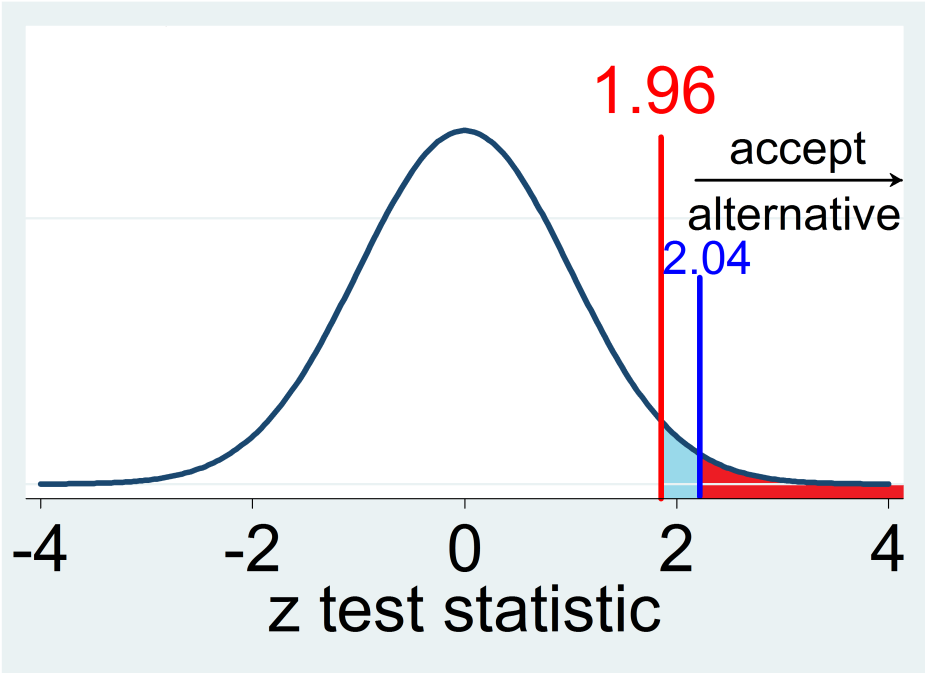

test-statistic
Direction of alternative
p-value
-4
-2
0
2
4
z test statistic

P-Value

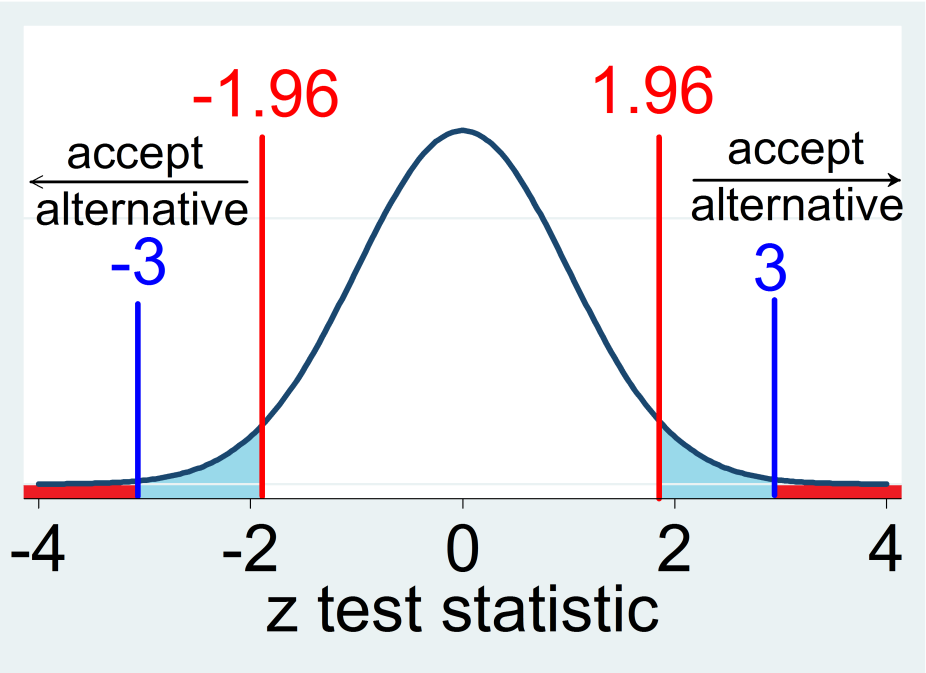
- Definition: The probability of observing a sample statistic at least as extreme as the one actually observed (in the direction of H_A) given H_0 is true
- Example: p-value$=P(\hat{p} > 0.6 | H_0$ is true$)$
- Small p-value:
  - Such an event is highly unlikely if H_0 is true
  - Cast doubt upon the validity of H_0
  - Small enough p-value gives us reason to reject H_0 and supports H_A
- P-value tells us exactly how likely we are to make a Type I error if we reject H_0
- For P-value, smaller is better (in support of alternative hypothesis)

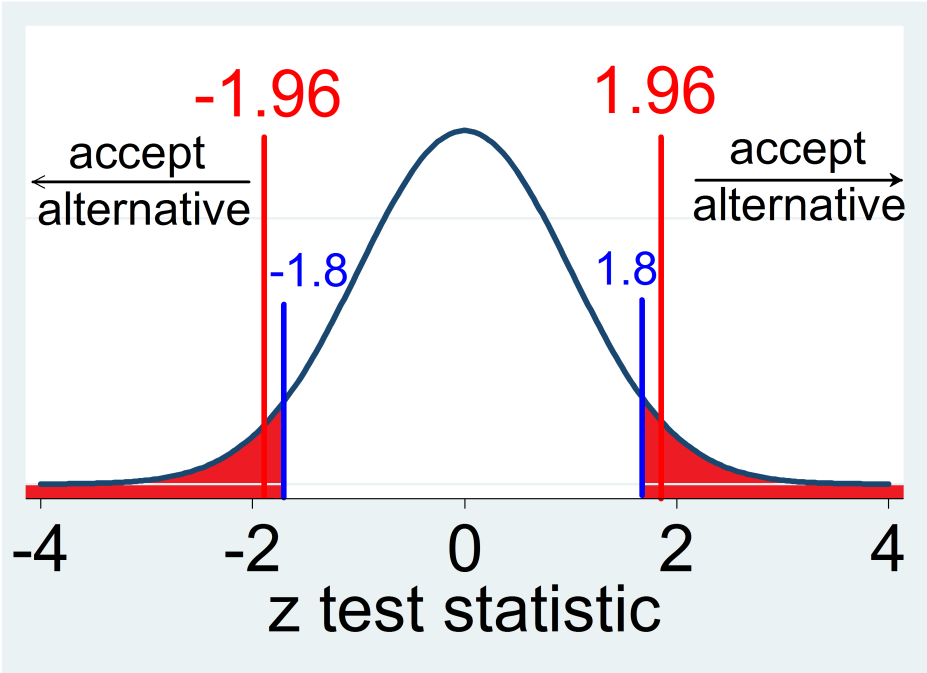
One-Sided Test





Two-Sided Test





How small does the P-Value have to be?

- How small the p-value have to be to infer that H_A is true?
- P-value between 0 and 0.01 implies overwhelming evidence
- P-value between 0.01 and 0.05 implies strong evidence
- P-value between 0.05 and 0.10 implies weak evidence
- P-value greater than 0.10 means no evidence in favour of H_A

Statistical significance

- Pick significance level before calculating p-value!
- If p-value falls below significance level, we say that the results from the test are statistically significant
  - Significant: has meaning, is important
  - Economically significant: the effect is large enough for decision makers to consider it to be important
  - Statistically significant: an effect that is not likely equal to zero given the data; an effect that is not likely observed due to chance (sampling error)
- Do not confuse statistical significance and economic, or practical significance
- Always report p-value together with your conclusion about the results of the test

Conventional Significance Levels

Significance Level, α	One-Sided Test	Two-Sided Test
1%	2.33	2.58
2%	2.06	2.33
5%	1.645	1.96
10%	1.28	1.645

Type I error and p-value

- Type I error: accept false H_A, or P(reject $H_0|H_0$ is true)
- Type II error: fail to reject false H_0, or P(fail to reject $H_0|H_A$ is true)
- With rejection region approach, we set the probability of Type I error by choosing appropriate significance level, α
- With p-value approach, we can calculate the exact probability of Type I error
- P-value $=$ probability of Type I error!
- P-value is the chance to reject H_0 when it is true

Testing Hypothesis

- Formulate null and alternative hypotheses
- Pick significance level, α
  1. Rejection Region
     - ★ Calculate test statistic
     - ★ Find rejection region using statistical tables
     - ★ Compare test statistic to rejection region
  2. *P*-value
     - ★ Calculate test statistic
     - ★ Compute *p*-value using statistical table
     - ★ Compare *p*-value to significance level, α
- Interpret results, and draw conclusion
- A picture speaks a thousand words, especially in hypothesis testing

M&M's In-Class Experiment

- Count M&M's of each color in a randomly selected bag.
- Compare with the "official" frequencies.
- Test the discrepancies (if any).
- All steps are described in your M&M's handout.