What's in the Bag?

Lecture 5
Section 1.4.3

Robb T. Koether

Hampden-Sydney College

Wed, Jan 25, 2012

Outline

Review Quiz

Example (Review Quiz)

1. In a certain hypothesis test, the probability of a Type II error is 0.25. It follows that
   (a) $\alpha = 0.25$.
   (b) $\beta = 0.25$.
   (c) The probability of a Type I error is 0.75.
   (d) We cannot determine the values of α or β from the information given.

Review Quiz

Example (Review Quiz)

2. A statistical experiment is constructed in which $\alpha = 0.10$. It follows that
   (a) $\beta = 0.90$.
   (b) $\beta > 0.10$.
   (c) H_0 has a 0.90 chance of being true.
   (d) We cannot determine the value of β from the information given.

Review Quiz Answers

Example (Review Quiz Answers)

1. (b) $\beta = 0.25$.
2. (d) We cannot determine the value of β from the information given.

Outline

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?
- Were we “kinda” sure?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?
- Were we “kinda” sure?
- “Pretty” sure?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?
- Were we “kinda” sure?
- “Pretty” sure?
- “Really” sure?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?
- Were we “kinda” sure?
- “Pretty” sure?
- “Really” sure?
- “Really, really” sure?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?
- Were we “kinda” sure?
- “Pretty” sure?
- “Really” sure?
- “Really, really” sure?
- Can we quantify that?

How Confident Are We?

- When we reach a conclusion, as in the coin-tossing experiment, how confident are we that we have reached the correct conclusion?
- How sure were we that the coin was fair?
- Were we “kinda” sure?
- “Pretty” sure?
- “Really” sure?
- “Really, really” sure?
- Can we quantify that?
- Yes, we use the “p-value” for that purpose.

The *p*-Value

Definition (*p*-value)

The *p*-value of an observation is the probability of observing a value at least as extreme as that observation, if the null hypothesis is true.

The p-Value

- If the p-value is small, that means that our observation was very unlikely if the null hypothesis is true.
- In which case, we would be inclined to reject H_0 in favor of H_1.

The p-Value

- If the p-value is small, that means that our observation was very unlikely if the null hypothesis is true.
- In which case, we would be inclined to reject H_0 in favor of H_1.
- On the other hand, if the p-value is relatively large, that means that our observation could reasonably have occurred by chance if the null hypothesis is true.
- In which case, we have no reason to reject H_0.

The *p*-Value

Definition (*p*-value)

The *p*-value of an observation is the probability of observing a value at least as extreme as *that observation*, if the null hypothesis is true.

Definition (α)

α is the probability of observing a value at least as extreme as *the critical value*, if the null hypothesis is true.

- Thus, *p*-values and α are calculated in essentially the same way.

The *p*-Value

Definition (*p*-value)

The *p*-value of an observation is the probability of observing a value at least as extreme as *that observation*, if the null hypothesis is true.

Definition (α)

α is the probability of observing a value at least as extreme as *the critical value*, if the null hypothesis is true.

- Thus, *p*-values and α are calculated in essentially the same way.
- The first uses the critical value.

The *p*-Value

Definition (*p*-value)

The *p*-value of an observation is the probability of observing a value at least as extreme as *that observation*, if the null hypothesis is true.

Definition (α)

α is the probability of observing a value at least as extreme as *the critical value*, if the null hypothesis is true.

- Thus, *p*-values and α are calculated in essentially the same way.
- The first uses the critical value.
- The other uses the observed value.

How Strong is the Evidence?

Example (p-value)

- Suppose I give a quiz with 12 questions.
- Each questions asks whether a person in a situation is making a Type I error or a Type II error.
- When I grade the student's quiz, I consider two hypotheses.
  - H_0 : He is guessing at each answer.
  - H_1 : He is not guessing at each answer.

How Strong is the Evidence?

Example (p-value)

- If he is guessing, then how likely is it that he will guess *all* 12 answers correctly?
- The probability turns out to be $\left(\frac{1}{2}\right)^{12} = 0.0002441$, or about 1 in 4000.
- That is the p-value of "12."

How Strong is the Evidence?

Example (*p*-value)

- If he is guessing, then how likely is it that he will guess *all* 12 answers correctly?
- The probability turns out to be $\left(\frac{1}{2}\right)^{12} = 0.0002441$, or about 1 in 4000.
- That is the *p*-value of "12."
- If he answered all 12 questions correctly, would we be inclined to reject the null hypothesis?

How Strong is the Evidence?

Example (*p*-value)

- If he is guessing, then how likely is it that he will guess at least 8 answers correctly?
- That probability turns out to be 0.1938, or about 1 in 5.
- That is the *p*-value of “8.”
- If he answered at least 8 questions correctly, would we be inclined to reject the null hypothesis?

How Strong is the Evidence?

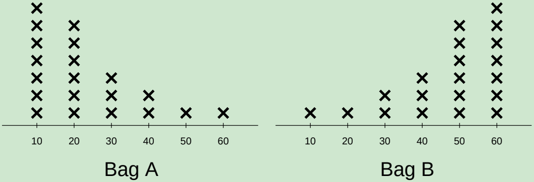
Example (*p*-value)

- What if he answered all 12 questions wrong?

The *p*-value

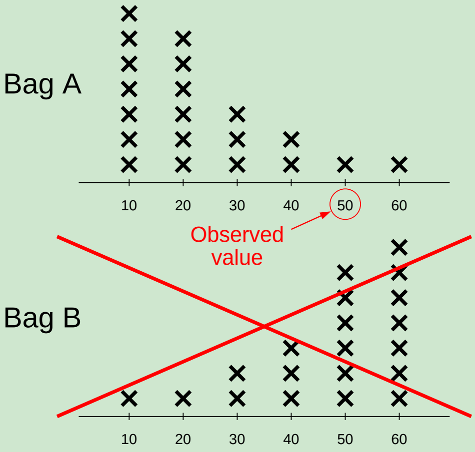
Example (The *p*-Value)

- In the Bag A/Bag B example from yesterday, if the selected token is worth \$50, what is the *p*-value?

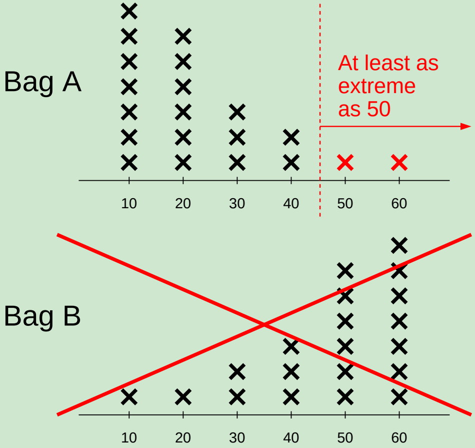


Two Bags

Example (The p-Value)

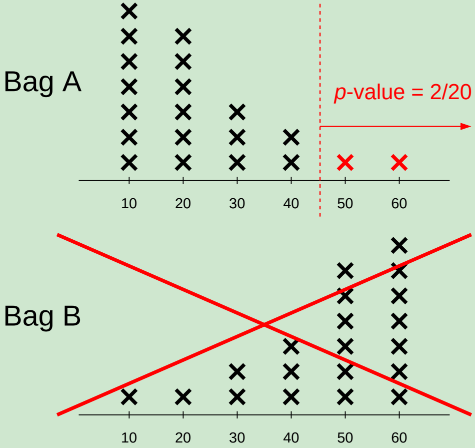


Example (The *p*-Value)



Two Bags

Example (The *p*-Value)



The *p*-value

Practice

- If the selected token is worth \$30, what is the *p*-value?
- Keep in mind,
  - We may compute the *p*-value of any number.
  - The *p*-value does not depend on the decision rule.
  - We always use the null hypothesis to compute the *p*-value.

Outline

The 12-Questions Quiz

- Reconsider the example of the student taking the quiz with 12 true/false questions.

The 12-Questions Quiz

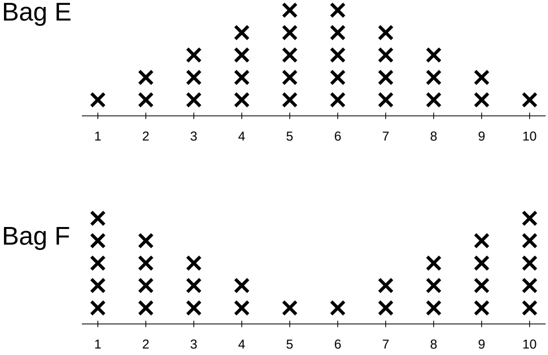
- Reconsider the example of the student taking the quiz with 12 true/false questions.
- If he answered all 12 questions wrong, what should we conclude?

The 12-Questions Quiz

- Reconsider the example of the student taking the quiz with 12 true/false questions.
- If he answered all 12 questions wrong, what should we conclude?
- Isn't that just as strong evidence against H_0 as if he had answered all 12 questions correctly?

A Two-Sided Test

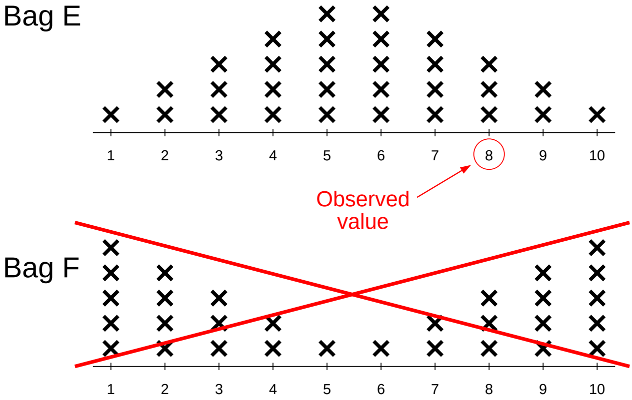
- Now consider Bag E and Bag F.



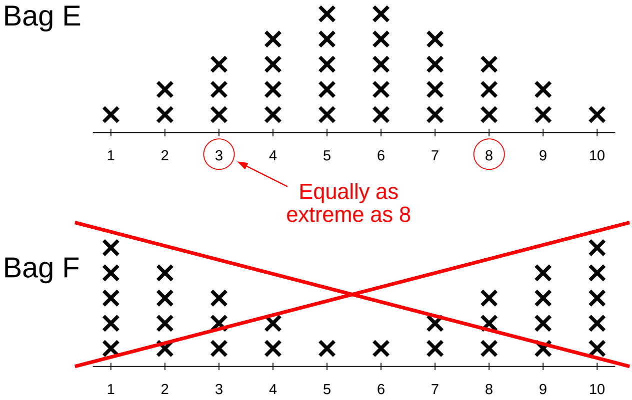
A Two-Sided Test

- If the selected token is worth \$8, what is the *p*-value?
- First, what is the direction of extreme?
- Next, which values are at least as extreme as 8?

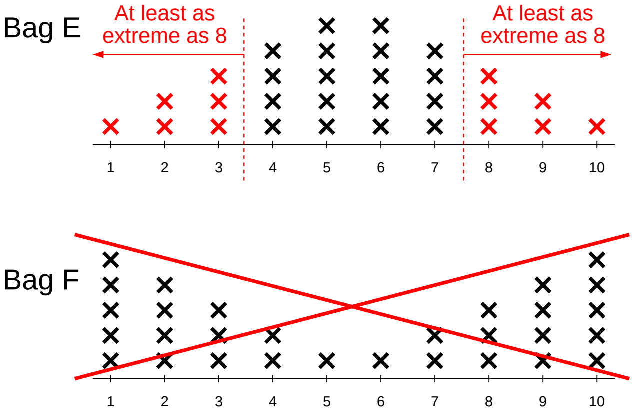
A Two-Sided Test



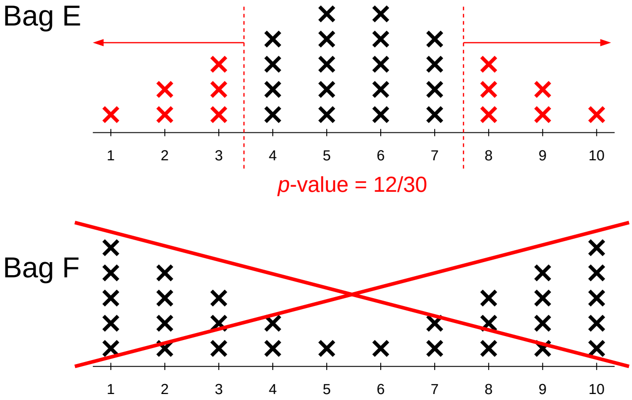
A Two-Sided Test



A Two-Sided Test



A Two-Sided Test



The p-value

- If the selected token is worth \$1, what is the p-value?
- If the selected token is worth \$5, what is the p-value?

Shortcut

- A shortcut:
  - In a two-sided test, if the null distribution is symmetric, then you can compute the probability in one direction, and then double it to get the p-value.

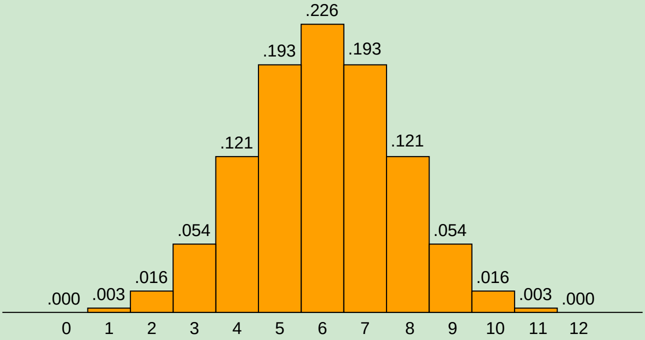
The 12-Question Quiz

Example (The 12-Question Quiz)

- In the true/false quiz, suppose the student answers 8 out of 12 questions correctly.
- What is the p-value of 8?

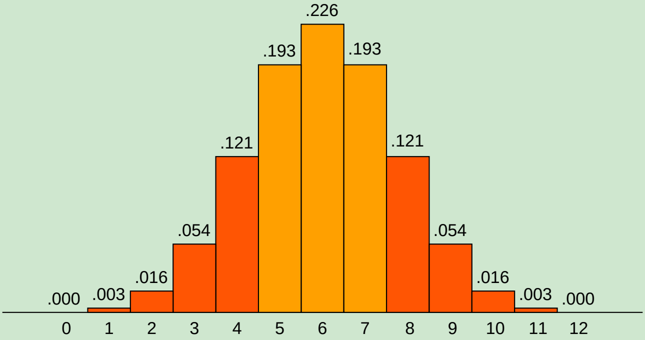
The 12-Question Quiz

Example (The 12-Question Quiz)



The 12-Question Quiz

Example (The 12-Question Quiz)



How Strong is the Evidence?

Example (p-value)

- The direction of extreme is two-sided.
- The probability of 8 *or more* correct answers is

$$0.121 + 0.054 + 0.015 + 0.003 + 0.000 = 0.193.$$

- Therefore, the p-value of 8 is $2 \times 0.193 = 0.386$.

Outline

Assignment

Homework

- Read Section 1.4.3, pages 28 - 47.
- Let's Do It! 1.9, 1.10, 1.11, 1.12, 1.13.
- Page 67, exercises 16, 27, 28, 30 - 32, 37, 43, 45, 50.