Theory of Inference III

Dajiang Liu @ PHS 525

Feb-16, 2016

Review from Last Time – Measure Uncertainty in Hypothesis Testing

- Hypothesis testing may not be flawless
- Errors can be made
  - Two types of errors: Type I Error and Type II Error

	Not Reject H_0	Reject H_0
H_0 is true	Okay	Type 1 Error
H_A is true	Type II Error	Okay

Type I and II Errors

- Type I Error: When null hypothesis is true, but incorrectly reject the null hypothesis

- Type II Error: When null hypothesis is not true, but fail to reject the null.

- Example:
  - Mendel's experiment: Pod color: yellow or green. Null hypothesis: the fraction of yellow pod is 0.25
  - What is a type I error & type II error

Significance Level

- Ideally, we want to minimize both type I and II errors

- However this is not often meaningful:
  - Rejecting all the null hypothesis will make type II errors zero, but type I errors 1

- Strategy used:
  - Control for the level of type I errors (say 5%), and minimize type II errors

- Significance level controls for type I errors
  - For example, we want to limit the type I error <5%, we use a hypothesis testing with significance level of 5%.

Measuring Significance in Hypothesis Testing: P-value

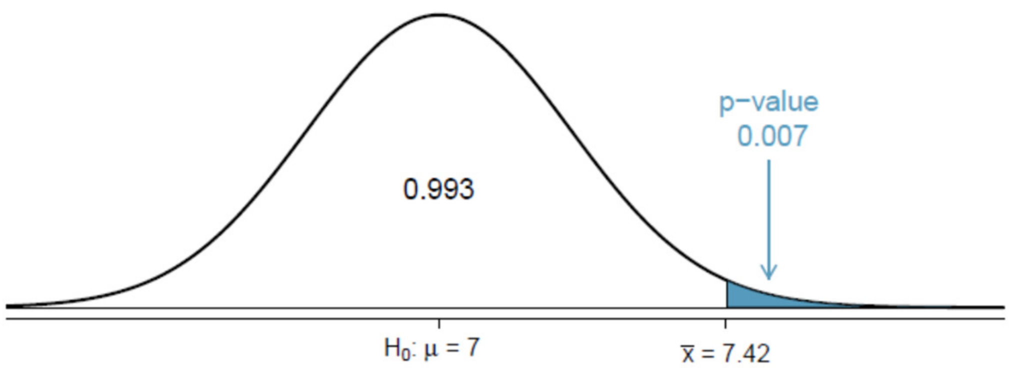
- Confidence interval is a coarse/simple way of performing hypothesis testing

- In practice, we want to measure how strong an evidence may be against the null hypothesis

- P-value measures the probability of observing a dataset that is more favorable to the alternative hypotheses than the current observation, given that the null hypothesis is true

P-value Example – Sleep Data

Figure 4.14: Distribution of a night of sleep for 110 college students. These data are moderately skewed.



How to Compute P-value – Testing for Sample Mean

For testing the null hypothesis that $H_0: \mu = \mu_0$

- Step 1: Compute sample mean value

$$\bar{X} = \frac{(X_1 + X_2 + \cdots + X_n)}{n}$$

- Step 2: Compute standard deviation for the sample

$$\sigma = \frac{\sqrt{(X_1 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}}{\sqrt{n}}$$

- Step 3: Compute standard error for the sample mean estimate

$$se(\bar{X}) = \sigma/\sqrt{n}$$

- Step 4: Estimate z-score

$$Z = (\bar{X} - \mu)/se(\bar{X})$$

- Step 5: If alternative hypothesis is $H_A: \mu > \mu_0$ $\text{PVALUE} = P(Y > Z)$
- If alternative hypothesis is $H_A: \mu < \mu_0$ $\text{PVALUE} = P(Y < Z)$
- If alternative hypothesis is $H_A: \mu \neq \mu_0$ $\text{PVALUE} = 2 * P(Y > |Z|)$

Use of P-value to Perform Hypothesis Testing

- Set up a null hypothesis
  - A statement of “no difference”

- Small p-values means that observing something more extreme is unlikely, assuming the null hypothesis is true
  - Small p-values indicate support for alternative hypothesis

- Reject the null hypothesis if p-values are smaller than the threshold level

TIP: Draw "P-values"

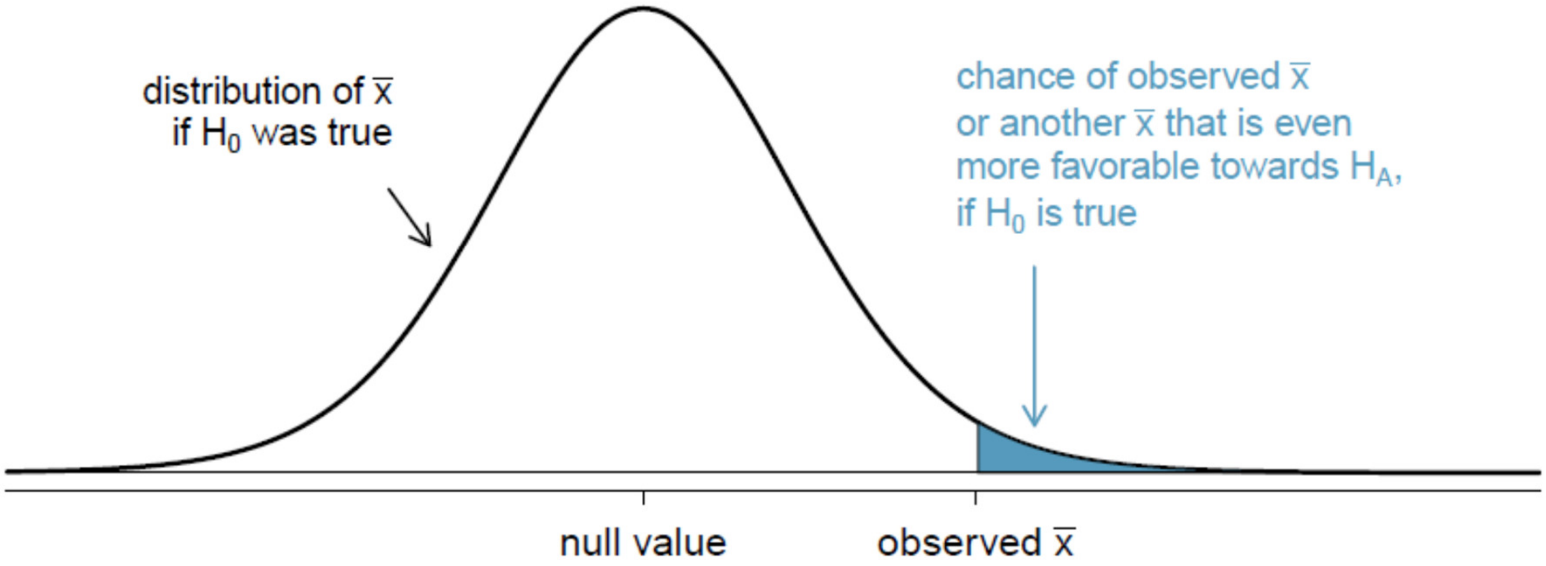


Figure 4.16: To identify the p-value, the distribution of the sample mean is considered as if the null hypothesis was true. Then the p-value is defined and computed as the probability of the observed $\bar{x}$ or an $\bar{x}$ even more favorable to H_A under this distribution.

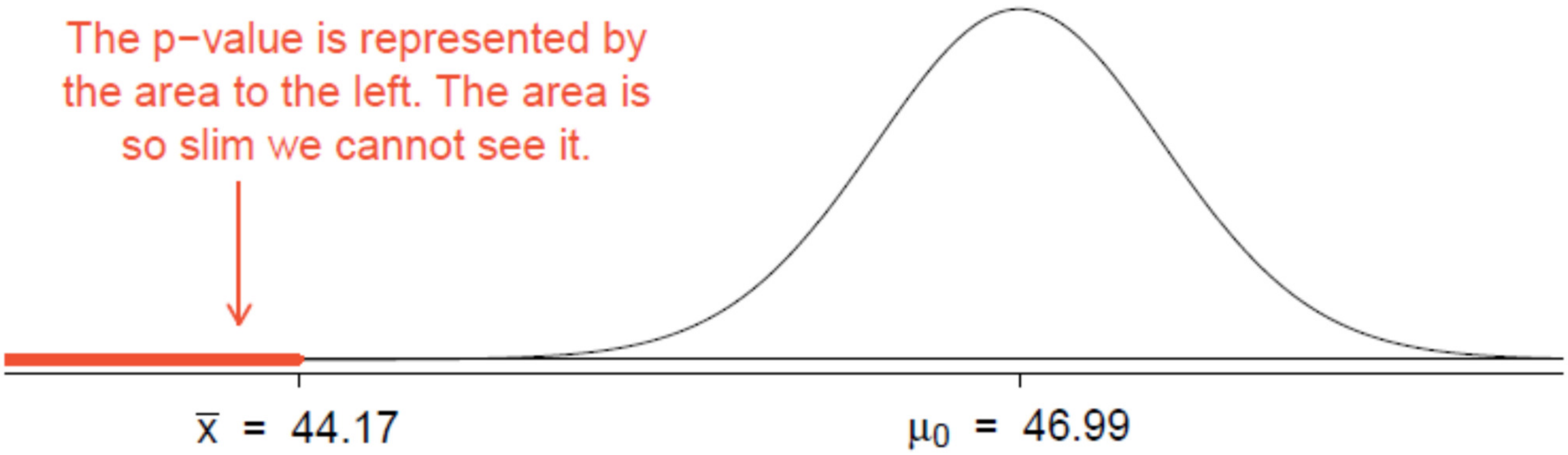
Example: E-bay Auction

- Ebay dataset: Mario Kart dataset in the folder
- Hypothesis: Is consumer paying less on eBay than on Amazon?
- Data: Amazon sells Nintendo Wii games for $46.99
- Question:
  - How to formulate the null/alternative hypothesis?
- Among 52 E-bay auctions, the mean value is $44.17, with a sample standard deviation of $4.15
- What is the Z-score statistic?
- Do we reject the null at $\alpha = 0.001/0.01/0.05$?

R codes

- #load data:
- dat=read.table('marioKart.txt',header=T,strngsAsFactors=T,sep='\t')
- #summarize dataset:
- summary(dat$totalPr)
- #calculate sample standard deviation:
- sd(dat$totalPr)
- #how many samples are there? (each row for a sample)
- nrow(dat)
- # standard error for the mean estimator
- se=sd(dat$totalPr)/sqrt(nrow(dat))
- # Z- score
- z=(49.88-46.99)/se
- # p-value for testing H_A: u<46.99
- pnorm(z,lower.tail=T)

Visualizing P Values



Two Sided Hypothesis and P-values

- How many hours does a college student sleep?

- One sided hypothesis testing:

$H_0: \mu = 7$ v.s. $H_A: \mu > 7$

- Sometimes, there is interest in testing two-sided hypothesis
  - Perform two sided hypothesis instead
- Two sided hypothesis:
  - $H_0: \mu = 7$ v.s. $H_A: \mu \neq 7$

College Student Sleep Data

- Data: surveyed 72 students, with mean value of $\bar{X} = 6.83$ and sample standard deviation of 1.8

- What is the z-score?

- What is the p-value for testing $H_0: \mu = 7$ and $H_A: \mu \neq 7$

Two Sided P-values

- Two sided p-values=left tail + right tail