Hypothesis Testing

• Principles
• Vocabulary
• Problems

Principles

• Game
• I say something is true
• Then we get some data
• Then you decide whether
  – Mr. Larsen is correct, or
  – Mr. Larsen is a lying dog
Risky Game
- Situation #1
- This jar has exactly (no more and no less than) 100 black marbles
- You extract a red marble
- Correct conclusion:
  - Mr. Larsen is a lying dog

Principles
- My statement will lead to certain probability rules and results
- Probability I told the truth is “zero”
- No risk of false accusation

Principles
- Game
- I say something is true
- Then we get some data
- Then you decide whether
  - Mr. Larsen is correct, or
  - Mr. Larsen has inadvertently made a very understandable error
Principles

• My statement will lead to certain probability rules and results
• Some risk of false accusation
• What risk level do you accept?

Risky Game

• Situation #2
• This jar has exactly (no more and no less than) 999,999 black marbles and one red marble
• You extract a red marble
• Correct conclusion:
  – Mr. Larsen is mistaken

Risky Game

• Situation #2 (continued)
• Mr. Larsen is mistaken because if he is right, the one red marble was a 1-in-a-million event.
• Almost certainly, more than red marbles are in the far than just one
Risky Game

• Situation #3
• This jar has 900,000 black marbles and 100,000 red marbles
• You extract a red marble
• Correct conclusion:
  – Mr. Larsen’s statement is reasonable

Risky Game

• Situation #3 (continued)
• Mr. Larsen’s statement is reasonable because it makes
  \( P(\text{one red marble}) = 10\% \).
• A ten percent chance is not too far fetched.

Principles (reworded)

• The statement or “hypothesis” will lead to certain probability rules and results
• Some risk of false accusation
• What risk level do you accept?
Risky Game

• Situation #4
• This jar has 900,000 black marbles and 100,000 red marbles
• A random sample of four marbles has 3 red and 1 black
• If Mr. Larsen was correct, what is the probability of this event?

Risky Game

• Situation #4 (continued)
• Binomial: n=4, x=1, p=0.9
• Mr. Larsen’s statement is not reasonable because it makes \( P(\text{three red marbles}) = 0.0036\).
• A less than one percent chance is too far fetched.

Formal Testing Method
Structure and Vocabulary

• The risk you are willing to take of making a false accusation is called the **Significance Level**
• Called “alpha” or \( \alpha \)
• \( P[\text{Type I error}] \)
Conventional $\alpha$ levels

- Two-tail
- 0.20 0.10
- 0.10 0.05
- 0.05 0.025
- 0.02 0.01
- 0.01 0.005

---

Formal Testing Method
Structure and Vocabulary

- Critical Value
  - similar to $Z_{\alpha/2}$ in confidence int.
  - separates two decision regions
- Critical Region
  - where you say I am incorrect

---

Formal Testing Method
Structure and Vocabulary

- Critical Value and Critical Region
  are based on three things:
  - the hypothesis
  - the significance level
  - the parameter being tested
- not based on data from a sample
- Watch how these work together
Test Statistic for $\mu$

\[
\frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t_{(n-1)df}
\]

Test Statistic for $p$

(np₀ > 5 and nq₀ > 5)

\[
\frac{\hat{p} - p_0}{\sqrt{\frac{p_0q_0}{n}}} \sim N(0,1)
\]

Test Statistic for $\sigma$

\[
\frac{(n-1)s^2}{s_0^2} \sim \chi^2_{(n-1)df}
\]
Formal Testing Method
Structure and Vocabulary

• $H_0$: always is $\leq$ or $\geq$
• $H_1$: always is $\neq >$ or $<$

Formal Testing Method
Structure and Vocabulary

• In the alternative hypotheses, $H_1$: put the parameter on the left and the inequality symbol will point to the “tail” or “tails”
• $H_1$: $\mu, p, \sigma \neq$ is “two-tailed”
• $H_1$: $\mu, p, \sigma <$ is “left-tailed”
• $H_1$: $\mu, p, \sigma >$ is “right-tailed”

Formal Testing Method
Structure and Vocabulary

• Example of Two-tailed Test
  – $H_0$: $\mu = 100$
  – $H_1$: $\mu \neq 100$
Formal Testing Method
Structure and Vocabulary

• Example of Two-tailed Test
  – $H_0: \mu = 100$
  – $H_1: \mu \neq 100$
• Significance level, $\alpha = 0.05$
• Parameter of interest is $\mu$

Formal Testing Method
Structure and Vocabulary

• Example of Two-tailed Test
  – $H_0: \mu = 100$
  – $H_1: \mu \neq 100$
• Significance level, $\alpha = 0.10$
• Parameter of interest is $\mu$

Formal Testing Method
Structure and Vocabulary

• Example of Left-tailed Test
  – $H_0: p \geq 0.35$
  – $H_1: p < 0.35$
Formal Testing Method
Structure and Vocabulary
• Example of Left-tailed Test
  – H₀: p ≥ 0.35
  – H₁: p < 0.35
• Significance level, α = 0.05
• Parameter of interest is “p”

Formal Testing Method
Structure and Vocabulary
• Example of Left-tailed Test
  – H₀: p ≥ 0.35
  – H₁: p < 0.35
• Significance level, α = 0.10
• Parameter of interest is “p”

Formal Testing Method
Structure and Vocabulary
• Example of Right-tailed Test
  – H₀: σ ≤ 10
  – H₁: σ > 10
Formal Testing Method  
Structure and Vocabulary  
• Example of Right-tailed Test  
  – H₀: σ ≤ 10  
  – H₁: σ > 10  
• Significance level, α = 0.05  
• Parameter of interest is σ

Formal Testing Method  
Structure and Vocabulary  
• Example of Right-tailed Test  
  – H₀: σ ≤ 10  
  – H₁: σ > 10  
• Significance level, α = 0.10  
• Parameter of interest is σ

Claims  
• is, is equal to, equals =  
• less than <  
• greater than >  
• not, no less than ≥  
• not, no more than ≤  
• at least ≥  
• at most ≤
Claims

• is, is equal to, equals
• $H_0$: =
• $H_1$: $\neq$

Claims

• less than
• $H_0$: $\geq$
• $H_1$: $<$

Claims

• greater than
• $H_0$: $\leq$
• $H_1$: $>$
Claims

• not, no less than
• $H_0$: ≥
• $H_1$: <

Claims

• not, no more than
• $H_0$: ≤
• $H_1$: >

Claims

• at least
• $H_0$: ≥
• $H_1$: <
Claims

• at most
• $H_0: \leq$
• $H_1: >$

Structure and Vocabulary

• Type I error: Deciding that $H_0$: is wrong when (in fact) it is correct
• Type II error: Deciding that $H_0$: is correct when (in fact) is is wrong

Structure and Vocabulary

• Interpreting the test result
  – The hypothesis is not reasonable
  – The Hypothesis is reasonable
• Best to define reasonable and unreasonable before the experiment so all parties agree
Traditional Approach to Hypothesis Testing

Test Statistic

- Based on Data from a Sample and on the Null Hypothesis, $H_0$:
- For each parameter ($\mu$, $p$, $\sigma$), the test statistic will be different
- Each test statistic follows a probability distribution

Traditional Approach

- Identify parameter and claim
- Set up $H_0$ and $H_1$:
- Select significance Level, $\alpha$
- Identify test statistic & distribution
- Determine critical value and region
- Calculate test statistic
- Decide: “Reject” or “Do not reject”
Next three slides are repeats of slides 19-21

Test Statistic for $\mu$
(small sample size: $n$)

$$\frac{\bar{x} - \mu_0}{s} \sim t_{(n-1)df}$$

Test Statistic for $p$
($np_0 > 5$ and $nq_0 > 5$)

$$\frac{\hat{p} - p_0}{\sqrt{\frac{p_0q_0}{n}}} \sim N(0,1)$$
Test Statistic for $\sigma$

\[
\frac{(n - 1)s^2}{\sigma_0^2} \sim \chi^2_{(n-1)df}
\]