Advanced Experimental Design PSY 464

Topic 3

Sampling Distributions/Hypothesis Testing, Effect Size, & Probability Basics

1

Agenda

- Descriptive vs. Inferential Statistics
  - Sampling distributions
  - Decision-making
  - Hypothesis testing, statistical significance and new directions
- Probability
  - Basics
  - Combinatorics
  - Bayes theorem
  - Binomial distribution

2

Descriptive vs. inferential

- Descriptive statistics: simply describing a collection of data
- Inferential statistics: making inferences beyond the data collected to some larger population
- Describing samples can be interesting
- Usually interested in estimating population parameters
- Generalize beyond the data at hand

3

Inferential statistics

- larger samples more likely than small to accurately describe a population
- can't measure every population member directly
- Sample mean is unbiased estimator of the population mean
- don't know how close we are to true population mean.
- Every sample is likely to result in a slightly different value

4

Sampling distributions

- Draw all possible random samples of size n from a population of N
- Calculated the mean of each sample
- We have created a sampling distribution of means

5

Hypothesis testing

- Interested in identifying patterns/differences "in nature"
- Data often ambiguous
- Howell's examples
- Example built around sample mean differences – goal is what?

6

Hypothesis testing – ESP?

- ESP example - slips of paper lettered A, B, etc.
- With A & B, only two possible orders, thus 50% chance of being right by chance
- With A, B, C - 6 different orderings:17% chance of getting right. If someone did this would you believe their ESP claim?

7

ESP example (cont.)

- With A, B, C, D – 24 orderings: 4.2% chance of being right
- Do you believe yet?
- A, B, C, D, E – 120 orderings, 0.8% chance
- More the individual was able to do, the more likely we would be to believe the claim
- Less likely to conclude that it was simply luck
- Where do we put the cut off for results being due to chance?

8

Statistical significance (α)

- Dictated by where we set alpha level
- Alpha level: beyond this point unwilling to believe that the result is due to random sampling error alone
- Significance levels of .05 or .01 are the most commonly used significance levels
- We are willing to accept a 5% or 1% probability that our decision is wrong

9

Interpretation of p

- p value associated with our statistical test represents the probability of a result as large or larger if null true
- Decisions about the null hypothesis
  - accepting or rejecting

		True State of Affairs in nature	
		H_0 true	H_0 false
Decision based on data	Retain H_0	Correct decision	Type 2 error ($p= \beta$)
	Reject H_0	Type 1 error ($p=\alpha$)	Correct decision

10

Statistical Signif. (cont.)

- Levels are completely arbitrary
- Usefulness being seriously questioned
- An alternative is to focus on effect sizes or magnitude of differences between groups
- Standard deviation units difference between groups is the simplest
  - Cohen's d

11

Overall Process

1. Begin with a research hypothesis
2. Set up the null hypothesis
3. Construct the sampling distribution of the particular statistic based on the assumption that the null hypothesis is true
4. Collect some data
5. Compare the sample statistic to that distribution
6. Reject or retain the null, depending on the probability, under the null, of a sample statistic as large or larger than the one we have observed.

12

How many tails? p – please?

- One tailed vs. two-tailed tests
  - What is the difference?
  - When do we use each?
- What does the p from our statistical test actually tell us?

13

New ideas in data analysis

- Jones & Tukey approach to hypothesis testing
  - Evaluate two separate directional hypotheses
  - If neither is statistically significant – conclude "not enough information"
  - Controls Type I error better
- Jacob Cohen – Effect size indices & power analysis
  - Practical vs. statistical significance
  - Power of typical study
  - Effect size indices, means and confidence intervals – meta analysis
  - Probability of replicating findings

14

Probability

- Terminology
  - Event
    - Independent events
    - Mutually exclusive events
    - Exhaustive
  - Additive rule
    - If A & B are mutually exclusive events: p(A or B) = p(A)+p(B)
  - Multiplicative rule
    - If A & B are independent events: p(A,B) = p(A)xp(B)
  - Can use both rules in one problem

15

Joint and Conditional Probability

- Joint probability – probability of two or more events in conjunction written p(A ∩ B) or p(A,B).
- Conditional probability – probability of some event A, given the occurrence of some other event B.
  - p(A|B) – probability of A given B
  - p(A|B) does not equal p(B|A)

16

Combinatorics

- Branch of probability theory that deals with how objects can be chosen and put together
- Permutations
- Combinations
- Which is which? Where does order matter?
- N! Factorial notation tells us to do what?

17

Permutations

- Each unique arrangement of objects
- A B different from B A

$$P_r^N = \frac{N!}{(N-r)!}$$

- Six experimental tasks, all similar, can only present three to any one subject.
- Want to provide all possible orders, how many subjects would be needed?

18

Permutations (cont.)

$$P_3^6 = \frac{6!}{(6-3)!} = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1} = 6 \cdot 5 \cdot 4 = 120$$

- Anyone having trouble w/math of this?
- Permutations are used any time that the ordering makes a difference

19

Combinations

- Order of selection makes no difference
- A B is the same as B A

$$C_r^N = \frac{N!}{r!(N-r)!}$$

- Cash 5 lottery game, 39 balls, match 5 to win jackpot

20

Combinations (cont.)

$$C_5^{39} = \frac{39!}{5!(39-5)!} = \frac{39 \cdot 38 \cdot 37 \cdot 36 \cdot 35}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} =$$

$$\frac{39 \cdot 38 \cdot 37 \cdot 7 \cdot 6}{4} = 575757$$

- Jackpot on 2-10 was $100,000
- Good deal?

21

A Problem to work through

- 10 members in a club, needs to elect four officers.
- President, Secretary, Treasurer, Parlimentarian
- Order matters here.
- How many different arrangements of officers can we derive from the club membership?
- Two different ways to solve…

22

Bayes theorem

- Accumulating information to generate probability estimates – integrating all available information
- Howell's example – Can see snow on the mountains. Is it April? (He's in CO rockies)
- Need to determine p(April |Snow)

23

Need to find p(April |Snow)

- p(April)=.083 p(NApril)=.917
- p(Snow|April)=.20 p(Snow|NApril)=.10

$$p(April \mid Snow) = \frac{p(Snow \mid April)\, p(April)}{p(Snow \mid April)\, p(April) + p(Snow \mid NApril)\, p(NApril)}$$

$$= \frac{(.2)(.083)}{(.2)(.083) + (.1)(.917)} = \frac{.0166}{.0166 + .0917} = \frac{.0166}{.1083} = .1533$$

24

Generic formula

- H – hypothesis
- NH – not hypothesis
- D – Data

$$p(H \mid D) = \frac{p(D \mid H)\,p(H)}{p(D \mid H)\,p(H) + p(D \mid NH)\,p(NH)}$$

- Confident in ability to apply this?

25

Gigerenzer and Hoffrage 1995 (and others)

- 1% of 40 y.o. women who have routine screening have breast cancer
- 80% of women with breast cancer will have a positive mammogram
- 9.6% of women without breast cancer will have a positive mammogram
- A 40 year-old woman has a positive mammogram in a routine screening
  - What is the probability that she actually has breast cancer?

26

Alternative presentation

- 10 out of 1000 women at age forty who participate in routine screening have breast cancer.
- 800 out of 1000 women with breast cancer will get a positive mammogram
- 96 out of 1000 women without breast cancer will also get a positive mammogram
- If 1000 women in this age group undergo a routine screening, about what fraction of women with positive mammograms will actually have breast cancer?

27

Common error

- Ignoring base rate – that is the overall proportion of women with cancer
- Also tend to ignore the rate of false positives and focus only on those with cancer who have the positive mammogram
- Many seem to make the incorrect p(cancer|+)=p(+|cancer) connection
- Correct answer requires all 3 pieces of information – p(Cancer), p(+|NCancer), and p(+|Cancer)

28

Gizmo mechanic problem

- You are a gizmo mechanic
- When a gizmo stops working, it is due to a blocked hose 30% of the time
- If a gizmo's hose is blocked, there is a 45% chance that prodding the gizmo produces sparks
- If the hose is unblocked, there is only a 5% chance that prodding=sparks
- You prod the gizmo and it produces sparks, what is the probability this gizmo has a blocked hose?

29

Binomial Distribution

- Testing proportions
- Can test a variety of hypotheses
- Howell presents sign test – simple nonparametric test
- SPSS binomial test, assess observed proportion against some hypothesized value

30