PSY 3307

Hypothesis Testing & The z-Test

Jonathan A. Pedroza, MS, MA

Cal Poly Pomona

2021-09-14

Agenda

- The Role of Inferential Statistics in Research
- Setting Up Inferential Procedures
- Performing the z-Test
- Interpreting **STATISTICALLY** Significant and Nonsignificant Results
- Summary of the z-Test
- The One-Tailed Test
- Statistics in Research Literature: Reporting the Results
- Errors in Statistical Decision Making

First

- None of this information will be on Exam 1
- This information, however, will be on Exam 2

Role of Inferential Statistics

- Sampling error is always a worry when trying to compare samples to a population in a sampling distribution
- **Inferential Statistics** procedures for deciding whether sample data represents particular relationships in a given population
- **Parametric Statistics/Tests** inferential statistical tests that require assumptions; used when computing the mean
    - data should be close to normally distributed
    - scores of variables should be on an interval or ratio scale

- **Nonparametric Statistics/Tests** inferential statistics that do not require stringent assumptions; used with the median and mode
    - severely skewed data for interval/ratio scales
    - scores of variables are nominal or ordinal scores
    - not always the case for all nominal or ordinal scores, there are parametric tests that handle these scores (e.g., logistic regression & multinominal regression)
- Parametric and nonparametric tests are for deciding whether there is a relationship between variables or if sampling error is misleading us into thinking there is a relationship

Setting up Inferential Procedures

For experiments, there are 4 steps in the process of conducting an experiment

1. Create experimental hypothesis/es
2. Design and conduct the experiment
3. Translate the experimental hypothesis/es into statistical hypothesis/es
4. Test the statistical hypothesis/es

- **Experimental Hypotheses** two statements describing the predicted relationship that may be demonstrated in the study
    - Don't worry about knowing this

JP's Guide To Actually Learning Hypotheses

- Do your research using peer-reviewed articles and other resources (e.g., meta-analyses, systematic reviews)
- Create a hypothesis/es relevant to your study's design
- Conduct your experiment/collect your data
- Conduct your statistical analyses
- Present findings

- **One-tailed test** when researchers predict the direction of scores in the DV will change
  - Females will have higher test scores than males.
- **Two-tailed test** when researchers don't predict the direction of scores in the DV will change

One-Sample Experiment

- testing to see if a sample's average outcome measured is significantly different from the population's average outcome
  - Giving an experimental drug for blood pressure to the sample and comparing it to the population's average blood pressure

The Null Hypothesis

- **Null Hypothesis** is a statement declaring that there is not a relationship between variables
  - Often seen as `H0`
- Starting point for research, as we don't know if a relationship will be present without any additional information
- **Research Hypothesis** is a statement declaring that there is a relationship between variables *Alternative Hypothesis* often see as `H1`

Examples

- There will be no differences between males and females in IQ test scores
- Non-Latina/o White adults will have more access to parks than Latina/o adults
- There will be a positive relationship between age and wisdom
- There will not be a relationship between hand size and social support

Logic of Hypothesis Testing

- At least for the current statistical testing we will be conducting (z-test)
- We are interested whether our DV score in our sample is representative of the population (HO) or if our sample's DV score is significantly different from the population DV score (H1)
- This means we will be using:
    - probability
    - region of rejection
    - criterion
    - critical value

Example

- We are interested in knowing if CPP students eat more or less fast food than the average CSU student H0: H1:
- The average number of a times a week that a CSU student eats fast food is 3.7, sd is 1.2
- Sampling from CPP, we find out the average CPP student eats fast food 2.4 times a week
- We would then conduct a z-test to see if our sample is statistically significantly different from CSU students
- Before doing any analyses, we would create our criterion for what counts as a significant finding

The z-Test

1. Randomly selected a sample
2. DV is somewhat normally distributed in population
    - and is ratio or interval scale
3. Know population mean of raw scores under another condition of the IV
4. Know population standard deviation

Sampling Distribution of Two-tailed Test

1. Create sampling distribution of means from population raw scores
   - This will be what our H0 states
2. Identify what the population mean is for H0
3. Select an alpha
   - **alpha** greek letter for criterion probability (e.g., .05)
4. Identify regions of rejection
   - One-tailed or two-tailed test
5. Determine critical value
   - two-tailed is $z = 1.96$

Compute the z-score

$$ z{obt} = \frac{\overline{X} - \mu}{\sigma{\overline{X}}} $$

- Need the standard error of the mean before getting the z-score
- z is now obtained from the data, that is why it is z obt in the formula above
- From the previous example
  - mu is 3.7
  - sigma is 1.2
- We would calculate xbar from our sample, let's just call it 2.4

```r
# standard error calculation
1.2/sqrt(100)
```

```
## [1] 0.12
```

```r
# se is .12

(2.4 - 3.7)/.12
```

```
## [1] -10.83333
```

```r
# z-score is -10.83
```

Comparing Obtained z-value to Critical Calue

- We know that the critical value is +-1.96
- Since our value is outside of the critical value in the region of rejection
    - we can reject that CPP students eat fast food as much as the rest of the CSU system students
- We have rejected the null hypothesis

Interpretation of STATISTICALLY Significant & Nonsignificant Results

- If you reject the null, you have statistically significant findings
    - H1 is supported, there is a relationship/there are differences
- If you retain the null, you don't have statistically significant findings
    - H0 supported, there is no relationship/there are no differences

What Does a Statistically Significant z-test Finding Represent

- Significant MEANS NOTHING
- Statistically significant indicates you reject the null and your sample is different from the population (for z-tests)
- We can't prove that H0 is false
- Book: We do not prove that our IV caused DV scores to change
- We don't know the exact population mean represented by our sample
    - potential sampling error

What Does a Nonsignificant z-test Finding Represent

- our sample is representative of the population
    - CPP students eat fast food as much as CSU students
- Don't say insignificant
    - best way of stating this is: "there was no evidence supporting that CPP students are statistically different from CSU students"
- We didn't find evidence of a statistically significant difference/relationship
- We simply have failed to reject the null hypothesis

z-Test Summary

- Create hypothesis/es
- Set up sampling distribution, select alpha, location region of rejection, determine critical value
- compute xbar, standard error and z-score obtained from population mean and standard deviation
- compare obtained z-score to critical value
  - statistically significant = reject H0
  - nonsignificnat = retain H0
  - don't draw conclusions

The One-Tailed Test

- predict scores in a direction
- interested in whether DV scores are higher or lower, not both
- the alpha still needs to be determined prior to analyses

Example

- H0:
- H1: CSUDH students will eat more fast food than CSU students
- CSUDH students eat 4.1 times a week

```r
# standard error calculation
1.2/sqrt(100)
```

```
## [1] 0.12
```

```r
# se is .12

(4.1 - 3.7)/.12
```

```
## [1] 3.333333
```

```r
# z-score is 3.33
```

Errors in Statistical Decision Making

- Make sense of the sampling error issues we've been talking about
- **Type I Errors** rejecting the null when the null hypothesis is true
  - false positive
- Know that with an alpha of 5%, 1 out of 20 statistical tests will be due to chance
  - at 1%, 1 out of 100 statistical tests will be due to chance

- **Note** This is why you don't test every single question of a scale. You won't know whether your finding is actually an extreme score or if you got your statistically significant finding by chance.
- Book: APA format to report $p < .05$
    - With current computing power, you should always report the exact p value
- **Type II Errors** retaining the null when the null hypothesis is false
    - false negative

- Summary
  - When null is true
  - and we reject the null = type I error
  - and we retain the null = accurate conclusion
  - when null is false
  - and we reject the null = accurate conclusion
  - and we retain the null = type II error

Power

- **Power** is the probability that we will detect a relationship correctly by avoiding a type II error
- Power can be determined a priori (before analyses) or post hoc (after analyses)
- Ways to improve power
    - increase sample size
    - potentially use a one-tailed test over a two-tailed test
    - use parametric tests over nonparametric tests