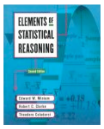
Testing statistical hypotheses about μ when σ is known: the one sample z-test

Minium, Clarke & Coladarci, Chapter 11



Statistical Inference: accounting for chance in sample results

- **Statistics are used to help us make decisions**
  - Can someone identify his favorite beer?
  - Let's assume that he can't (i.e., we assume he's guessing)
  - We'll change our minds only if he gets a "*significant*" number correct
    - let's say 8 or more out of 10 because that has a probability of about .05 of occurring by chance (i.e., if he's guessing)
  - We do the test, count the number correct, then decide if we have to change our minds

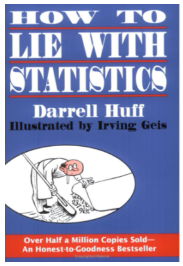
Number Correct	Probability
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977

Statistical Inference: accounting for chance in sample results

- **Restate the question as a null hypothesis and an alternative hypothesis**
- **Determine characteristics of the appropriate sampling distribution**
- **Specify the**
  - significance level required (α)
  - and the corresponding cutoff value of the test statistic
- **Determine your sample's score (*X*) and**
- **Decide whether to reject the null hypothesis**

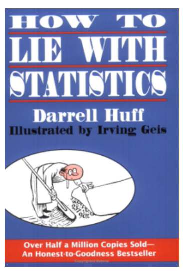
An example problem

- This is a classic book about the misuse of statistics (first published in 1954).
- I think students who read this book will do better in PSYC315 than those who don't.
- I have grades from many years of statistics classes and so I know that the grades are normally distributed with an average of 68% and a standard deviation of 14; i.e., $\mu = 68$ and $\sigma = 14$.
- Q: How do I test my theory?



An example problem

- A: I choose 49 PSYC315 students at random and ask them to read the book
- I record their marks at the end of the year
- Q: how do I make a decision?
- A: I test the null hypothesis.
  - I assume that reading the book has no effect on students' grades; this is the null hypothesis.
  - The alternative hypothesis is that reading the book does have an effect on students' grades.
  - I will only change my mind if the average grade of the 49 students is is "*significantly*" greater than 68%.



The statistical hypotheses: H_0 and H_1

- The Null Hypothesis (H_0)
  - is the hypothesis that is assumed to be true and formally tested
  - determines the sampling distribution to be employed
  - is the hypothesis about which the final decision is to reject or retain.
- The Alternative Hypothesis (H_1)
  - typically represents the underlying research question of the investigator
  - specifies the alternative *population* condition that is supported or asserted upon rejection of H_0

The statistical hypotheses: H_0 and H_1

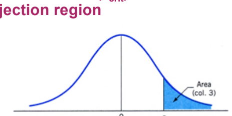
- **So, in this case**
  - H_1: μ > 68
    - the mean of the population of students reading Huff's book is greater than 68.
  - H_0: μ = 68
    - the mean of the population of students reading Huff's book is 68.
  - Note that H_0 and H_1 are expressed in terms of population parameters

Choosing the appropriate sampling distribution

- **In this example the appropriate sampling distribution is the sampling distribution of the means**
- **The population being sampled has μ = 68, σ = 14 and sample size (*n*) is 49**
- **Therefore,**
  - $\mu_{\bar{X}}$ =
  - $\sigma_{\bar{X}}$ =
  - Once we know what the mean of the sample is we can compute
  - $z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$

Choosing a significance level and cutoff score

- **The level of significance (α) specifies how rare the sample results must be to cause us to reject H_0 as untenable. α is typically set at .05 (and sometimes .01)**
- **5% of the standard normal distribution lies above z = 1.64, which is our cutoff score (z_{crit}) and the area above z_{crit} is called the rejection region**



- **So, we'll reject H_0 if our observed mean ($\bar{X}$) is more than 1.64 standard deviations above 68.**

Computing $\bar{X}$ and z then making a decision

- At the end of term we find that the average grade of the 49 students is $\bar{X}$ = 72%
- Therefore, *z* = (72 - 68)/2 = 2 ($z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$)
- Since 2 > z_{crit} = 1.64 (i.e., it falls in the rejection region) we reject H_0
  - in fact there is a probability of approximately .02 of obtaining $z \geq 2$ when only chance is operating
- We conclude that reading Huff's book leads to improved marks in PSYC315.

Summary of the steps

- Specify H_0, H_1, α and z_{crit}
- Select the sample and calculate
  - $\bar{X}$
  - $\sigma_{\bar{X}}$
  - $z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$
- Determine the probability of obtaining the *z* or greater under the null hypothesis
- Make a decision regarding H_0
- ***It is important to remember that in this example we know*** μ ***and*** σ***, and this permits us to compute the SEM and hence a z-score.***

Decision Errors

- Type 1 error: rejecting H_0 when it is actually true
  - The level of significance, α, gives the probability of rejecting H_0 when it is actually true.
- Type 2 error: failing to reject H_0 when it is actually false
  - To calculate the probability of a Type 2 error requires more information than we have at the moment.
  - We'll deal with this when we discuss the concept of power.

One tailed vs Two tailed tests

- A one tailed test has one rejection region because we are making a prediction about the direction of our effect.
- We use two tailed tests when we are testing whether an effect exists but we are not sure of the direction of the effect; i.e., we don't know if the sample mean will be above or below the population mean.

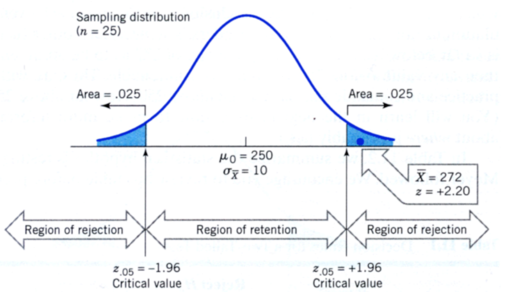
Another example; Home schooling, from the book.

- Q: Does home schooling make a difference?
- We know that the average score of "school-schooled" 4th graders on a standardized test is 250 with a standard deviation of 50; the test is known to produce a normal distribution of scores
  - i.e., $\mu = 250$ and $\sigma = 50$
- We'll choose a sample of 25 home-schooled 4th graders and compute their average score then try to decide if it is *significantly* different from 250

Another example; Home schooling, from the book.

- The steps in conducting the test
- Specify H_0, H_1, α and z_{crit}
- Select the sample and calculate
  - $\bar{X}$
  - $\sigma_{\bar{X}}$
  - $z = \dfrac{\bar{X} - \mu}{\sigma_{\bar{X}}}$
- Determine the probability of obtaining a *z* as extreme as the one observed under the null hypothesis
- Make a decision regarding H_0.
- The following slide summarizes our hypothesis testing situation and the outcome

Another example; Home schooling, from the book



p = .0139 + .0139 = .0278

Important considerations

- **The nature and role of H_0 and H_1**
  - H_0 can be tested directly because it provides the specificity necessary to locate the appropriate sampling distribution. H_1 does not.
- **Caution:**
  - when we compute the probability of obtaining the observed z under H_0 (e.g., .001), THIS DOES NOT MEAN THAT THE NULL HYPOTHESIS HAS A PROBABILITY OF .001 OF BEING TRUE!!!
  - Rather, it means that assuming that H_0 is true, the observed results has a probability of .001 of occurring by chance alone.

Important considerations

- When we "Reject H_0" it sounds as though we are claiming it is false; but this is not the case.
- We are saying that the result is unusual in the sense that it has a low probability of occurring when H_0 is true; Sir Ronald A. Fisher used the term "statistically significant" to mean statistically unusual.
- So, when we "reject H_0" we are concluding that something other than chance is responsible for this unusual result.
  - Of course we recognize that this conclusion might be wrong.
- **Rejection vs Retention of H_0**
  - Retention of H_0 merely means that there is insufficient information to reject it and thus that it *could* be true. It does not mean that it *must* be true, or even that it *probably* is true.

Important considerations

- **Statistical significance vs Importance**
  - A result may be statistically significant and yet completely unimportant
  - Consider that the SEM depends on sample size (*n*). Therefore, if sample size is very large then even small differences can be statistically significant.

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

- **Effect size (again)**

$$EffectSize = \frac{\bar{X} - \mu}{\sigma}$$