How to Determine a *p*-Value When Testing a Null Hypothesis

Based on Deborah J. Rumsey

Presented by Leen-Kiat Soh
CSCE100 Introduction to Informatics

Introduction

- When you **test a hypothesis** about a **population**, you can use your test statistic to decide whether to reject the null hypothesis, H_0
- You make this decision by coming up with a number, called a *p*-value

Null Hypothesis

- In statistics, **a null hypothesis is a statement that one seeks to nullify** (that is, to conclude is incorrect) with evidence to the contrary
  - Most commonly, it is presented as a statement that ***the phenomenon being studied produces no effect or makes no difference***
- An experimenter frames a null hypothesis with the **intent of rejecting it**:
  - i.e., intending to run an experiment which produces data that shows that the phenomenon under study does indeed make a difference
- An example of such a null hypothesis might be the statement,
  - "A diet low in carbohydrates has no effect on people's weight."
  - **To nullify**: in this case, show that a diet low in carbohydrates over some specific time frame does in fact tend to lower the body weight of people who adhere to it

https://en.wikipedia.org/wiki/Type_I_and_type_II_errors

Type I Error

- A *type I error* (or *error of the first kind*) is the **rejection of a true null hypothesis**
- **Usually a type I error leads to the conclusion that a supposed effect or relationship exists when in fact it does not**
- Examples of type I errors include a test that shows
  - a patient to have a disease when in fact the patient does not have the disease
  - a fire alarm going on indicating a fire when in fact there is no fire
  - an experiment indicating that a medical treatment should cure a disease when in fact it does not

https://en.wikipedia.org/wiki/Type_I_and_type_II_errors

Type II Error

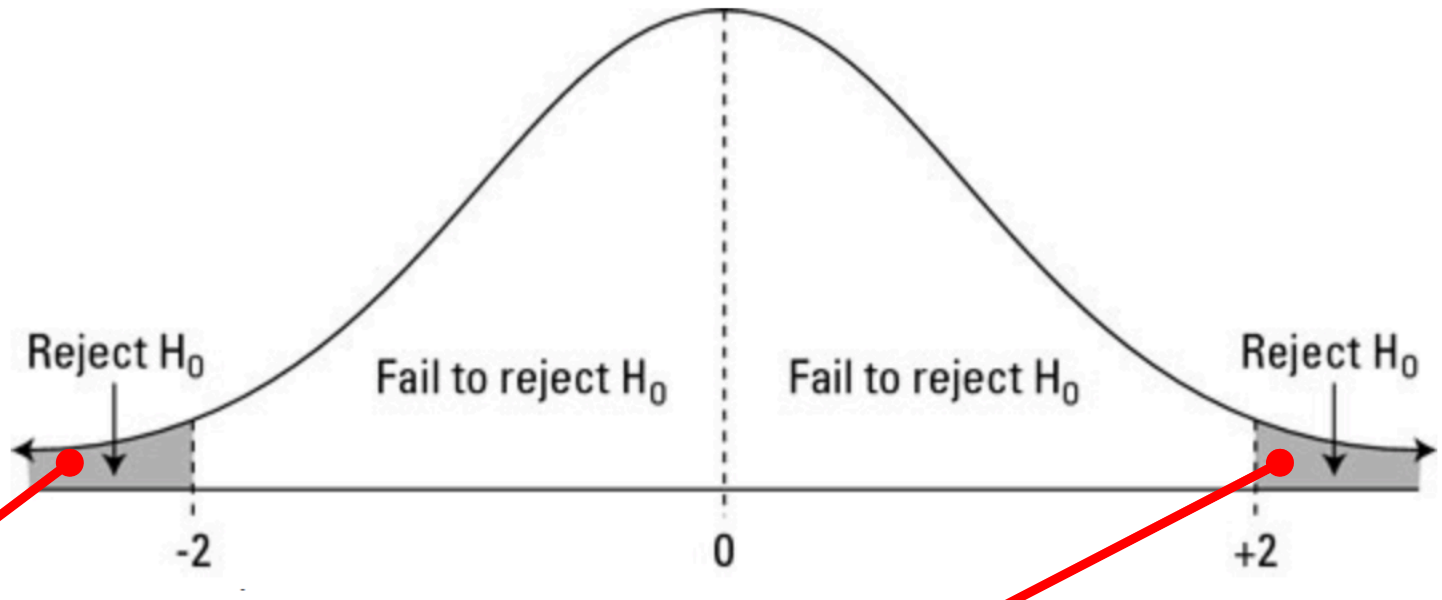
- A *type II error* (or *error of the second kind*) is **the failure to reject a false null hypothesis**
- Some examples of type II errors are
  - a blood test failing to detect the disease it was designed to detect, in a patient who really has the disease
  - a fire breaking out and the fire alarm does not ring
  - a clinical trial of a medical treatment failing to show that the treatment works when really it does

https://en.wikipedia.org/wiki/Type_I_and_type_II_errors

p-value

- A *p*-value is a probability associated with your critical value
- The critical value depends on **the probability you are allowing for a Type I error**
- It measures the chance of getting results at least as strong as yours if the claim (H_0) were true

Test Statistic

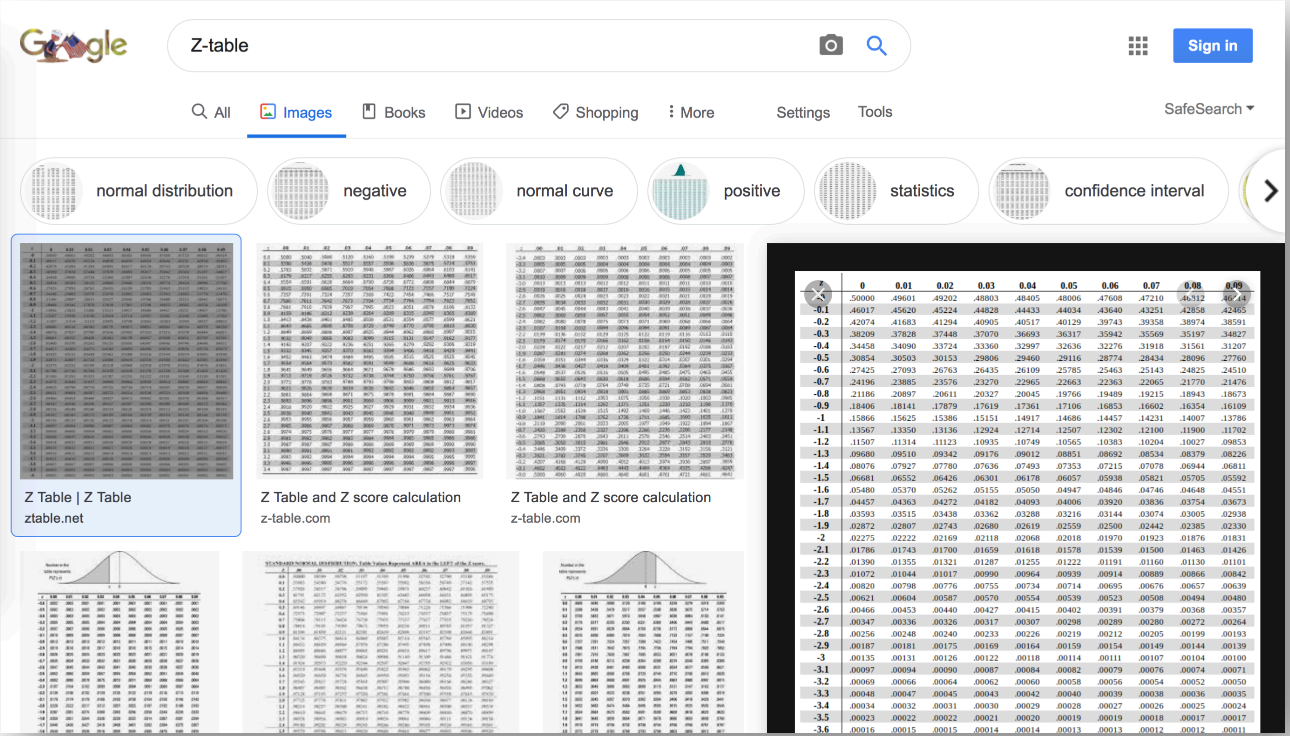


- If the alternative hypothesis, H_a, is the less-than alternative, you reject H_0 only if the test statistic falls in the left tail of the distribution (below –2)
- If H_a is the greater-than alternative, you reject H_0 only if the test statistic falls in the right tail (above 2)

p-value: How to compute it?

- To find the *p*-value for your test statistic:

1. Look up your test statistic on the appropriate distribution
    - E.g. in this case, on the standard normal (*Z*-) distribution (see the following *Z*-table).



p-value: How to compute it? 2

2. Find the **probability that *Z* is beyond (more extreme than) your test statistic**:
   a. If H_a contains a less-than alternative, find the probability that *Z* is less than your test statistic (that is, look up your test statistic on the *Z*-table and find its corresponding probability) → this is the *p*-value
   b. If H_a contains a greater-than alternative, find the probability that *Z* is greater than your test statistic (look up your test statistic on the *Z*-table, find its corresponding probability, and subtract it from one) → this is the *p*-value
   c. If H_a contains a not-equal-to alternative, find the probability that *Z* is beyond your test statistic and double it
      - There are two cases

p-value: How to compute it? 3

2. Find the **probability that *Z* is beyond (more extreme than) your test statistic** …
   c. If H_a contains a not-equal-to alternative, find the probability that *Z* is beyond your test statistic and double it … There are two cases …
      - If your test statistic < 0, (1) find the probability that *Z* is less than your test statistic (look up your test statistic on the *Z*-table and find its corresponding probability) and (2) double this probability to get the *p*-value.
      - If your test statistic > 0, (1) find the probability that *Z* is greater than your test statistic (look up your test statistic on the *Z*-table, find its corresponding probability, and *subtract it from one*) and (2) double this result to get the *p*-value.

Example

- Suppose you are testing a claim that the percentage of UNL personnel (faculty, staff, and students) who are husker football fans is 25%
- Suppose out of your sample of 100 UNL personnel, 20% are husker football fans
  - Thus: the sample proportion is p=0.20
  - The standard error for your sample percentage is the square root of p(1-p)/n which equals 0.04 or 4% (z-score)
- You find the **test statistic** by taking the proportion in the sample who are husker football fans, 0.20, subtracting the claimed proportion of all UNL personnel who are husker football fans, 0.25, and then dividing the result by the standard error, 0.04
  - Test statistic (standard score) = (0.20 – 0.25)/0.04 = –1.25
  - Your sample results and the population claim in H_0 are 1.25 standard errors apart
    - Your sample results are 1.25 standard errors below the claim
- When testing H_0: $p = 0.25$ versus H_a: $p < 0.25$, you find that the *p*-value of -1.25 by finding the probability that Z is less than -1.25
  - ***p*-value = 0.1056**

Example

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
-0.1	.46017	.45620	.45224	.44828	.44433	.44034	.43640	.43251	.42858	.42465
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-1	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-2	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842

- Suppose you are testing a claim th staff, and students) who are husk
- Suppose out of your sample of 10
  - Thus: the sample proportion is p=0
  - The standard error for your sample or 4% (z-score)
- You find the **test statistic** by takin football fans, 0.20, subtracting the husker football fans, 0.25, and the
  - Test statistic (standard score) = (0.2 0.25)/0.04 = -1.25
  - Your sample results and the population claim in H_0 are 1.25 standard errors apart
    - Your sample results are 1.25 standard errors below the claim
- When testing H_0: $p = 0.25$ versus H_a: $p < 0.25$, you find that the *p*-value of -1.25 by finding the probability that Z is less than -1.25
  - ***p*-value = 0.1056**

Example, Cont'd …

- When you look this number up on the above *Z*-table, you find a probability of **0.1056 (*p*-value)** of Z being less than this value
  - Significant?
  - < 0.1, < 0.05, < 0.01
  - ***Not significant***
- Note, if you had been testing the two-sided alternative, the *p*-value would be $2 * 0.1056$, or 0.2112
  - ***Even less significant***

How does this relate to your Final Project?

- Do you have a null hypothesis that you want to nullify?
  - "X% of the tweets that support Y have non-aggressive tone"
    - How do you determine X?
  - "The number of positive tweets is the same as the number of negative tweets"
    - Or "The % of positive tweets is 50%"