Hypothesis Testing

CSCE 587

Background

Assumptions:

- Samples reflect an underlying distribution
- If two distributions are different, the samples should reflect this.
- A difference should be testable

Model Evaluation

- Issue: limited observations
- Questions:
  - What is the appropriate test
  - What level of significance is appropriate

Test Selection

- One sample test:
  - compare sample to population
  - Actually: sample mean vs population mean
  - Does the sample match the population

- Two sample test:
  - compare two samples
  - Sample distribution of the difference of the means

Selecting the Test

Type of problem:

- Comparing a sample to a population
- “Does the sample come from the population?”
- Test: One Sample *t*-test

$$t = \frac{\mu_{sample} - \mu_{population}}{SD_{sample} / \sqrt{\# observations}}$$

One Sample *t*-Test

$$t = \frac{\mu_{sample} - \mu_{population}}{SD_{sample} / \sqrt{\# observations}}$$

Test: One Sample *t*-test

- Used only for tests of the population mean
- Null hypothesis: the means are the same
- Compare the mean of the sample with the mean of the population
- What do we know about the population?

Example

- Historically, the mean score for juniors on a test is 81.3
- Six students take the test resulting in a mean score of 87.6 and a standard deviation of 13.14
- The *t*-value is:

$$t = \frac{87.6 - 81.3}{13.14/\sqrt{6}} = \frac{6.3}{5.354} = 1.176$$

Example

Given t= 1.176, how do we test the hypothesis?

Compare to critical values in t-table.

To do so we need:

- to know the degrees of freedom
- To select a confidence interval
- One-tailed or two-tailed test?

One-Tailed vs Two-Tailed

One-tailed:

- allots all of the alpha in one direction
- (Note: alpha is the probability of rejecting the null hypothesis when it is true)
- Makes sense if we want to know a difference in one direction
- $\mu_{sample} > \mu_{population}$ or $\mu_{sample} < \mu_{population}$

One-Tailed vs Two-Tailed

Two-tailed: we only care if the means are different

- Splits alpha into two tails
- Makes sense if we are only asking if the means are different

$$\mu_{sample} \neq \mu_{population}$$

Example

- The degrees of freedom
  - DoF for one sample test is n-1 = 5 in this case
- To select a confidence interval
  - We select 95% confidence interval
- One-tailed or two-tailed test?
  - One-tail allots all of the alpha in one direction
  - Makes sense if we want to know a difference in one direction
  - Two-tailed: we only care if the means are different
  - We choose two-tailed

Example

Parameters: t= 1.176, DoF=5, 95%, 2-tailed

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Decision

Compare t= 1.176, and t_α = 2.571

Accept or reject the null hypothesis?

If $|t_c| > |t_\alpha|$ ➔ Reject null hypothesis

If p-value < α ➔Reject null hypothesis

Steps We Took

1. Selected the test
2. Selected confidence level
3. Calculated the test statistic
4. Determined the degree of freedom
5. Compared the computed test statistic to the table value
6. Accepted/Rejected the hypothesis

Test Selection

- One sample test:
  - compare sample to population
  - Does the sample match the population
- Two sample test:
  - compare two samples
  - Sample distribution of the difference of the means

Two Sample Tests

- Independent two-sample *t*-test
  - Independently sampled
  - Identically distributed
  - What does this mean?
- Paired two-sample *t*-test
  - Paired, not independently sampled
  - Example: sample, treat, resample same subjects

Independent two-sample *t*-test

Must have same variance

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1 X_2}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Numerator: difference between samples

Denominator: estimated standard error of the difference between the two means

Note the above version allow unequal sample sizes

Independent two-sample *t*-test

Denominator: estimated standard error of the difference between the two means

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_{X_1 X_2}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where: $$S_{X_1 X_2} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

And S_1^2, S_2^2 are the **unbiased** variances of the two samples

Unbiased Estimates of the Mean & Variance

The unbiased estimate of the mean from a sample is calculated using:

$$\bar{x} = \frac{\sum x}{n}$$

The variance if unknown can be calculated by using the formula below and we call it s^2 rather than σ^2

$$s^2 = \frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right) = \frac{1}{n-1}\sum\left(x - \bar{x}\right)^2$$

Degrees of Freedom

One sample test: n1 + n2 – 1

Two sample test: n1 + n2 – 2

Why the difference?

- For a sample: DoF = n-1
- We have two samples: (n1-1) & (n2-1)

Independent t-test, unequal sizes, unequal variances

- Use Welch's *t*-test
- Difference is in the denominator

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Where, S_1^2, S_2^2 are the **unbiased** variances of the two samples