Week 12 - Confidence Intervals and Hypothesis Testing

Key Concepts

- Population vs Sample
- Sampling Bias: some members of the intended population are less likely to be included than others in the selected sample.

Key Concepts

- Statistical hypothesis: assumption about the population average.
- Null Hypothesis H_0: assumption that the population average is identical to a specific value.
- Alternative Hypothesis H_1 or H_a: opposite of the null hypothesis. We compare this hypothesis with the null hypothesis to decide whether or not we reject the null hypothesis.

Key Question

What is the chance of observing the test-statistic, this extreme, for this sample (considering its size and its propertiest), purely by chance if H_0 was true?

Types of Alternative Hypothesis

$$H_0 : \mu = k$$

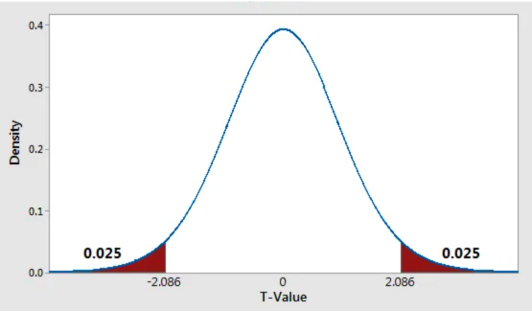
- Two types of alternative hypotheses: one-sided and two-sided.
- One-sided alternative hypothesis is used to determine whether the population average differs from the hypothesized value in a specific direction (e.g. larger than).

$$H_1 : \mu > k \text{ OR } H_1 : \mu < k$$

- Two-sided alternative hypothesis is used to determine whether the population average is either greater than or less than the hypothesized value.

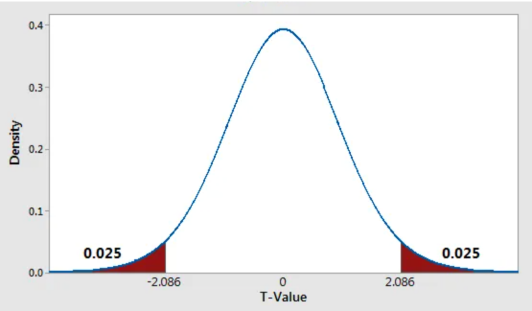
$$H_1 : \mu \neq k$$

Two-Sided Hypothesis Test



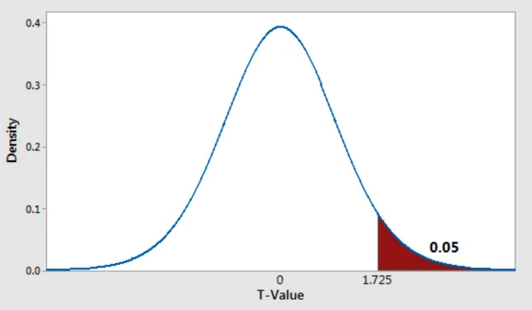
Plots the probabilities of obtaining test statistic values when the null hypothesis is correct for the population. The peak of the distribution occurs at $t = 0$, which represents the null hypothesis. When the null hypothesis is true for the population, obtaining samples that exhibit a large discrepancy becomes less likely, so the probability mass becomes smaller away from the center.

Two-Sided Hypothesis Test



Critical regions are ranges of the distributions where we can reject the null hypothesis. At a significance level of $\alpha = 0.05$ and a two-sided tests, those are the outer 2.5% on each tail. When the test statistic falls in the critical region, the sample data are sufficiently incompatible with H_0 - we can reject it.

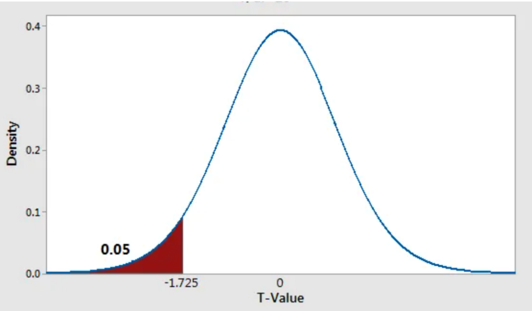
One-Sided Hypothesis Test



$$H_0 : \mu \leq k, H_a : \mu > k$$

If we're in the critical region: H_a is more probable.

One-Sided Hypothesis Test



$$H_0 : \mu \geq k, H_a : \mu < k$$

If we're in the critical region: H_a is more probable.

Test Assumptions

1. Our observations must be independent of each other. For example, if we have people who live in the same household participating in a medical trial, they might be exposed to the same environmental conditions or eat the same food. This can bias our results.

2. Normality of data. We assume that the sample is derived from a normally distributed data.

3. Adequate sample size. In order to perform a test using the normal distribution and not approximate to the t distribution, our sample size must be greater than 30.

4. In order to use the normal distribution for our hypothesis test, we must assume the population standard deviation is known. If the population standard deviation is not known, then we use the t-distribution for the hypothesis test.

Test Statistic

Our test statistic is equal to the difference between the sample mean and our constant divided by the standard error.

For the z-test (assuming a normal distribution):

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

For the t-test (using a student-t distribution for smaller samples):

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

where we have n-1 degrees of freedom.

Test Significance

- We use test significance to indicate whether we are confident enough to reject the null hypothesis.

- Test significance is represented by the probability score we calculate from the z or t test. (We need to use statistical programs or distribution tables for this!) Typically, we are confident to reject when we are 95% confident by using the 5% significance level ($\alpha = 0.05$).

- The significance that we produce from the test statistic is the probability that we obtained our result due to random chance.

Example

A pharmaceutical company is trying out a medication for lowering blood sugar and managing diabetes. It is known that any level of Hemoglobin A1c below 5.7 is considered normal. The drug company has treated 100 study volunteers with this medication and measured an average Hemoglobin A1c level of 5.1 with a standard deviation of 1.6. They would now like to prove that is significantly below 5.7.

Example

A pharmaceutical company is trying out a medication for lowering blood sugar and managing diabetes. It is known that any level of Hemoglobin A1c below 5.7 is considered normal. The drug company has treated 100 study volunteers with this medication and measured an average Hemoglobin A1c level of 5.1 with a standard deviation of 1.6. They would now like to prove that is significantly below 5.7.

$$H_0 : \mu \geq 5.7$$

$$H_a : \mu < 5.7$$

Example

We do not know anything about the population standard deviation, so even though the sample size is large enough, we will use the t test.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$t = \frac{5.1 - 5.7}{1.6/\sqrt{100}} = -3.75$$

Example

We do not know anything about the population standard deviation, so even though the sample size is large enough, we will use the t test.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$t = \frac{5.1 - 5.7}{1.6/\sqrt{100}} = -3.75$$

Since we have 100 subjects, we use 99 degrees of freedom to compute the test statistic.

T Table

***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
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
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										

The test statistic for 99 degrees of freedom translates to a p-value of less than 0.0005. This is smaller than 0.05 and therefore we reject the null hypothesis.

Example: Municipal Children's Home

Boys of a certain age are known to have a mean weight of $\mu = 85$ pounds. A complaint is made that the boys living in a municipal children's home are underfed.
As one bit of evidence, $n = 25$ boys (of the same age) are weighed and found to have a mean weight of 80.94 pounds. It is known that the population standard deviation $\sigma = 11.6$ pounds (the unrealistic part of this example!).
Based on the available data, what should be concluded concerning the complaint?

Assumption: Samples are drawn from a population which is normally distributed.

Example: Municipal Children's Home

It is assumed that the population mean weight is $\mu = 85$, but we do not have the complete data from the population. Otherwise we would have calculated the actual mean directly. However we have sample data from 25 subjects. So based on this sample data we will try to prove or disprove our assumption, using statistical test.

Example: Municipal Children's Home

Steps for solving this problem:

1. Define the null hypothesis H_0.

Example: Municipal Children's Home

Steps for solving this problem:

1. Define the null hypothesis H_0.

$$H_0 : \mu <= 85$$

2. Define the null hypothesis H_a.

Example: Municipal Children's Home

Steps for solving this problem:

1. Define the null hypothesis H_0.

$$H_0 : \mu <= 85$$

2. Define the null hypothesis H_a.

$$H_a : \mu > 85$$

This is a two-sided test since our null hypothesis would be false on either side of H_a ($H_a : \mu > 85$ and if $H_a : \mu < 85$).

Example: Municipal Children's Home

3. Decide on a test statistic based on the information available.

Example: Municipal Children's Home

3. Decide on a test statistic based on the information available. Assuming data is normally distributed and number of observations are less than 30, we will use a t-test.

4. Level of significance: This defines the rejection region / critical region.

Example: Municipal Children's Home

3. Decide on a test statistic based on the information available. Assuming data is normally distributed and number of observations are less than 30, we will use a t-test.

4. Level of significance: This defines the rejection region / critical region. We will use $\alpha = 0.05$.

5. Calculate the test statistic based on the given information