

Cheat Sheet:

Hypothesis Testing for

Population Means

Essential Concepts

One-Sample Hypothesis Test for Means

- The null hypothesis (H_0) is a statement about the population that either is believed to be true or is used to put forth an argument unless it can be shown to be incorrect beyond a reasonable doubt.
 - $H_0 : \mu = \mu_0$, μ_0 is the null value.
- The alternative hypothesis (H_A) is a claim about the population that is contradictory to H_0 and what we conclude when we reject H_0 .
 - $H_A : \mu < \mu_0$, μ_0 is the null value.
 - $H_A : \mu > \mu_0$, μ_0 is the null value.
 - $H_A : \mu \neq \mu_0$, μ_0 is the null value.
- Conditions for a One-Sample t -Test
 1. The sample is a random sample from the population of interest or it is reasonable to regard the sample as if it is random. It is reasonable to regard the sample as a random sample if it was selected in a way that should result in a sample that is representative of the population.
 2. For each population, the distribution of the variable that was measured is approximately normal, or the sample size for the sample from that population is large. Usually, a sample of size 30 or more is considered to be "large." If a sample size is less than 30, you should look at a plot of the data from that sample (a dotplot, a boxplot, or, if the sample size isn't really small, a histogram) to make sure that the distribution looks approximately symmetric and that there are no outliers.

- Test statistics: t -statistic: $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$

The distribution of t -scores depends on the degrees of freedom, that is, $df = n - 1$.

Two-Sample Hypothesis Test for Means

- Independent samples are two samples where the individuals selected for the first sample do not influence the individuals selected for the second sample.
- A hypothesis test for comparing two population means is often referred to as a two-sample t -test.
- The null hypothesis (H_0): It is a statement about the population that either is believed to be true or is used to put forth an argument unless it can be shown to be incorrect beyond a reasonable doubt.
 - Null hypothesis: $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$
- The alternative hypothesis (H_A): It is a claim about the population that is contradictory to H_0 and what we conclude when we reject H_0 .
 - Alternative hypothesis:
 - $H_A : \mu_1 < \mu_2$ or $H_A : \mu_1 - \mu_2 < 0$
 - $H_A : \mu_1 > \mu_2$ or $H_A : \mu_1 - \mu_2 > 0$
 - $H_A : \mu_1 \neq \mu_2$ or $H_A : \mu_1 - \mu_2 \neq 0$
- Conditions for a t -Test

1. The sample should be randomly selected or reasonably representative of the population.
2. For each population, the distribution of the variable that was measured is approximately normal, or the sample size for the sample from that population is large. Usually, a sample of size 30 or more is considered to be "large." If a sample size is less than 30, you should look at a plot of the data from that sample (a dotplot, a boxplot, or, if the sample size isn't really small, a histogram) to make sure that the distribution looks approximately symmetric and that there are no outliers.

- standard error of $\bar{x}_1 - \bar{x}_2$: $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

- The test statistic to compare two population means is calculated using the following formula:

$$t = \frac{\text{estimate of parameter-null hypothesis value}}{\text{standard error}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Paired samples or Dependent samples are samples that are chosen in a way that results in the observations in one sample being paired with the observations in the other sample.
- The mean of the differences is equal to the difference in means: $\mu_d = \mu_{\text{after}} - \mu_{\text{before}}$
- In summary, where k is the value of the null hypothesis, we have:

Alternative Hypothesis for Independent Samples	Alternative Hypothesis for Dependent Samples
$H_A : \mu_1 - \mu_2 > k$	$H_A : \mu_d > k$
$H_A : \mu_1 - \mu_2 < k$	$H_A : \mu_d < k$
$H_A : \mu_1 - \mu_2 \neq k$	$H_A : \mu_d \neq k$

- The notations for the summary statistics used to compare paired populations/samples are shown in the following table. We will use d to represent the difference variable.

Summary Statistics	Notation
Population Mean of Difference	μ_d
Sample Mean of Difference	\bar{d}
Population Standard Deviation of Difference	σ_d
Sample Standard Deviation of Difference	s_d

- The test statistic for the dependent (paired) t-test is calculated using the following formulas:

$$\text{standard error of the difference} = \frac{s_d}{\sqrt{n}}$$

$$\text{test statistic } (t) = \frac{\text{estimator} - \text{null value}}{\text{standard error of estimator}} = \frac{\bar{d} - \text{null value}}{\text{standard error of difference}}$$

Key Equations

***t*-statistic**

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

standard error of difference of means

$$\bar{x}_1 - \bar{x}_2: \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

standard deviation of difference of means

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Glossary

independent sample

random samples from each population

paired samples, dependent samples

samples that are chosen in a way that results in the observations in one sample being paired with the observations in the other sample