Two-sample z-test for means

Use case

We have two independent (i.e., disjoint) populations. We want to determine whether there is a statistically significant difference between the means of each population. We know with certainty the standard deviation of each population,
σ_1 and
σ_2. We have a sample of size
n_1 from population 1, and a sample of size
n_2 from popluation 2. Population 1 yields sample mean
$\bar{x}_1$, and population 2 yields sample mean
$\bar{x}_2$.

Preconditions

To apply a two-sample z-test, one should verify that sample means are approximately normally distributed. This is the case if either of the following conditions holds:

1. The population is normally distributed.
2. The sample size is large enough such that the Central Limit Theorem takes effect.

One can enhance robustness by planning
$n_1 \approx n_2$.

Hypotheses

Null hypothesis,
H_0**:**
$\mu_1 - \mu_2 = 0$ (i.e., there is no difference between the two populations.)

Alternative hypothesis,
H_a**:**
$\mu_1 - \mu_2 \neq / > / < 0$ (i.e., there is a difference between the two populations.)

Test statistic

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

p-value

Where Z is a standard normal random variable,

1. For
   $H_a : \mu_1 - \mu_2 < 0$:
   $p = P(Z \leq z)$
2. For
   $H_a : \mu_1 - \mu_2 > 0$:
   $p = P(Z \geq z)$
3. For
   $H_a : \mu_1 - \mu_2 \neq 0$:
   $p = 2P(Z \leq -|z|)$

p-value (Python)

1. For
   $H_a: \mu_1 - \mu_2 < 0$: `p = stats.norm.cdf(z)`
2. For
   $H_a: \mu_1 - \mu_2 > 0$: `p = 1 - stats.norm.cdf(z)`
3. For
   $H_a: \mu_1 - \mu_2 \neq 0$: `p = 2 * stats.norm.cdf(-abs(z))`

Associated confidence interval

$$\text{C\% confidence interval} = \bar{x}_1 - \bar{x}_2 \pm z^\star \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\text{choose } z^\star \text{ s.t. area on standard normal distribution from } (-z^\star, z^\star) = \text{C}$$

Related tests

If you want to perform inference on the mean of a single population, a one-sample z-test for means is more appropriate.

If you don't have access to one or both of the population standard deviation values, a two-sample t-test for means is more appropriate.

Considerations

A two-sample z-test conducted at confidence level
α will reject the null hypothesis if and only if the value corresponding to the null hypothesis,
0, is completely outside of the
$C = 1 - \alpha$ confidence interval for the true difference between means.