10.1: Unpaired z-Test

We have two populations and two sample sets, one from each population :

	Sample Mean	Sample std. dev.
From population 1	$\overline{x}_1$	s_1
From population 2	$\overline{x}_2$	s_2

The population means are μ_1 and μ_2 and just as with the single population test, there are 3 possible hypothesis tests :

Two Tailed	Right Tailed	Left Tailed
$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 \leq \mu_2$	$H_0 : \mu_1 \geq \mu_2$
$H_1 : \mu_1 \neq \mu_2$	$H_1 : \mu_1 > \mu_2$	$H_1 : \mu_1 < \mu_2$
or	or	or
$H_0 : \mu_1 - \mu_2 = 0$	$H_0 : \mu_1 - \mu_2 \leq$	$H_0 : \mu_1 - \mu_2 \geq 0$
$H_1 : \mu_1 - \mu_2 \neq 0$	$H_1 : \mu_1 - \mu_2 > 0$	$H_1 : \mu_1 - \mu_2 < 0$

In the second row the hypotheses are written in terms of a difference. Irrespective of which way you write the hypotheses, give population 1 priority. Write population 1 first. That way you won't mess up your signs or your interpretation.

The test statistic to use, in all cases[1] is

$$z_{\text{test}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where n_1 = sample set size from population 1 and n_2 = sample set size from population 2. This test statistic is based on a distribution of sample means as shown in Figure 10.1.

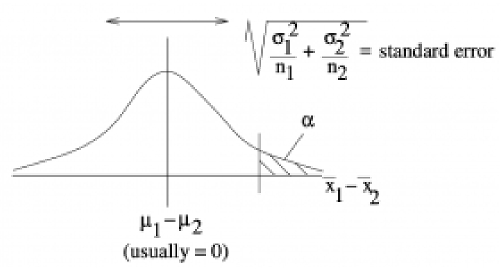


Figure 10.1 : The distribution of the difference of sample means $\bar{x}_1 - \bar{x}_2$ under the null hypothesis $H_0 : \mu_1 - \mu_2 = 0$. A one-tail example is shown here. The test statistic of Equation 10.1 follows from a z-transformation of this picture.

Example 10.1 : A researcher hypothesizes that the average number of sports colleges offer for males is greater than the average number of sports offered for females. Samples of the number of sports offered to each sex by randomly selected colleges is given here :

Males (pop. 1)	Females (pop. 2)
$n_1 = 50$	$n_2 = 50$
$\bar{x}_1 = 8.6$	$\bar{x}_2 = 7.9$
$s_1 = 3.3$	$s_2 = 3.3$

At $\alpha = 0.10$ is there enough evidence to support the claim?

Solution :

1. Hypotheses.

$$H_0 : \mu_1 \leq \mu_2 \qquad H_1 : \mu_1 > \mu_2 \text{ (claim)} \tag{10.1.1}$$

Note that $\bar{x}_1 > \bar{x}_2$ $(8.6 > 7.9)$ so $H_1 : \mu_1 > \mu_2$ is true on the face of it. If H_1 is not true on the face of it then H_1 is just plain false without the need for any statistical test. With the hypotheses direction set correctly, the question becomes: Is $\bar{x}_1$ *significantly* greater than $\bar{x}_2$? The term “statistically significant” corresponds to “reject H_0“.

2. Critical statistic.

From the ***t* Distribution Table**, one-tailed test at $\alpha = 0.10$ we find

$$z_{\text{crit}} = 1.282 \tag{10.1.2}$$

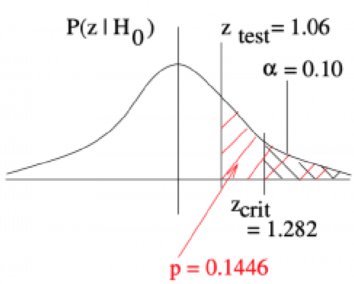
Note that $z_{critical}$ is positive because this is a right-tailed test. For left tailed tests make z_{crit} negative. For two-tailed tests you have $\pm z_{\text{crit}}$.

3. Test statistic.

$$\begin{aligned} z &= \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\ &= \frac{(8.6 - 7.9)}{\sqrt{\frac{3.3^2}{50} + \frac{3.3^2}{50}}} \\ &= 1.06 \end{aligned}$$

Using the **Standard Normal Distribution Table**, we can find the p-value. Since $A(z) = A(1.06) = 0.3554$, $p = 0.05 - 0.3554 = 0.1446$.

4. Decision.



Do not reject H_0 since z_{test} is not in the rejection region. The p-value reflects this :

$$(p = 0.1446) > (\alpha = 0.10) \tag{10.1.3}$$

5. Interpretation.

There is not enough evidence, at $\alpha = 0.10$ under a z-test, to support the claim that colleges offer more sports for males than females.

---

1. You could specify a non-zero null hypothesis, e.g. $H_0 : \mu_1 - \mu_2 = k$, in which case you would have $z_{\text{test}} = \frac{(\bar{x}_1 - \bar{x}_2) - k}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$. We won't consider that case in this course. ↵