

Testing the Difference in Two Independent Proportions

SPSS has terrific functionality for performing a number of complex statistical tests, yet it offers no clear-cut way to test the difference in two independent (or dependent) proportions. This technical tip provides a quick look at a template that was developed in Microsoft Excel to test the difference in two independent proportions. The example that follows tests the difference in the 1-year (freshman-to-sophomore) retention rates for two independent first-time, full-time (FTFT) fall student cohorts.

Just four values (two per cohort) need to be filled in on the template to complete the test. Figure 1 illustrates the first stage of the template: the 95% confidence interval calculation for each cohort's retention rate (**note**: only the calculations for cohort A are shown). The two fields that are shaded represent those fields where data needs to be entered. The original cohort size and the number of students retained after one year are entered in the shaded cells (**note**: explanatory comments are attached to some cells). All other fields are formula-driven. The same two values for the second cohort (cohort B in this example) are also entered in the space provided on the template (not shown).

mdapp: 1-year retention rate								
First-Time Full-Time Cohort A								
orig cohort	n retained	ret rate	1-ret rate	sq rt ret	sq rt n	SE	95% CI	
							lower	upper
1,089	869	79.80%	20.20%	0.40	33.00	0.01	77.41%	82.18%

Figure 1. Constructing the 95% Confidence Intervals (CI)

Figure 2 shows the second stage of the template where the test of the difference in the two retention rates (independent proportions) is completed based on the values entered for each of the two cohorts from stage one. All fields are formula-driven. In this case, the null hypothesis that there is no difference in the two retention rates is rejected.

Null hypothesis: there is NO difference in the two independent proportions.							
Pooled estimate for overall proportion	p_{s1}	p_{s2}	z statistic ¹	df	absolute value: critical value	z - crit val	Test Decision
0.76	0.80	0.73	4.06	2,860	1.96	2.10	REJECT NULL

Figure 2. Test of the Difference between Proportions (Level of Significance=.05)

Figure 3 shows a high-low chart constructed in SPSS based on the upper and lower limits for the 95% confidence intervals for each cohort. The non-overlapping intervals for the cohorts provide a visual confirmation of the decision to reject the null hypothesis of no difference. Overlapping intervals would have confirmed a decision to retain the null hypothesis of no difference.

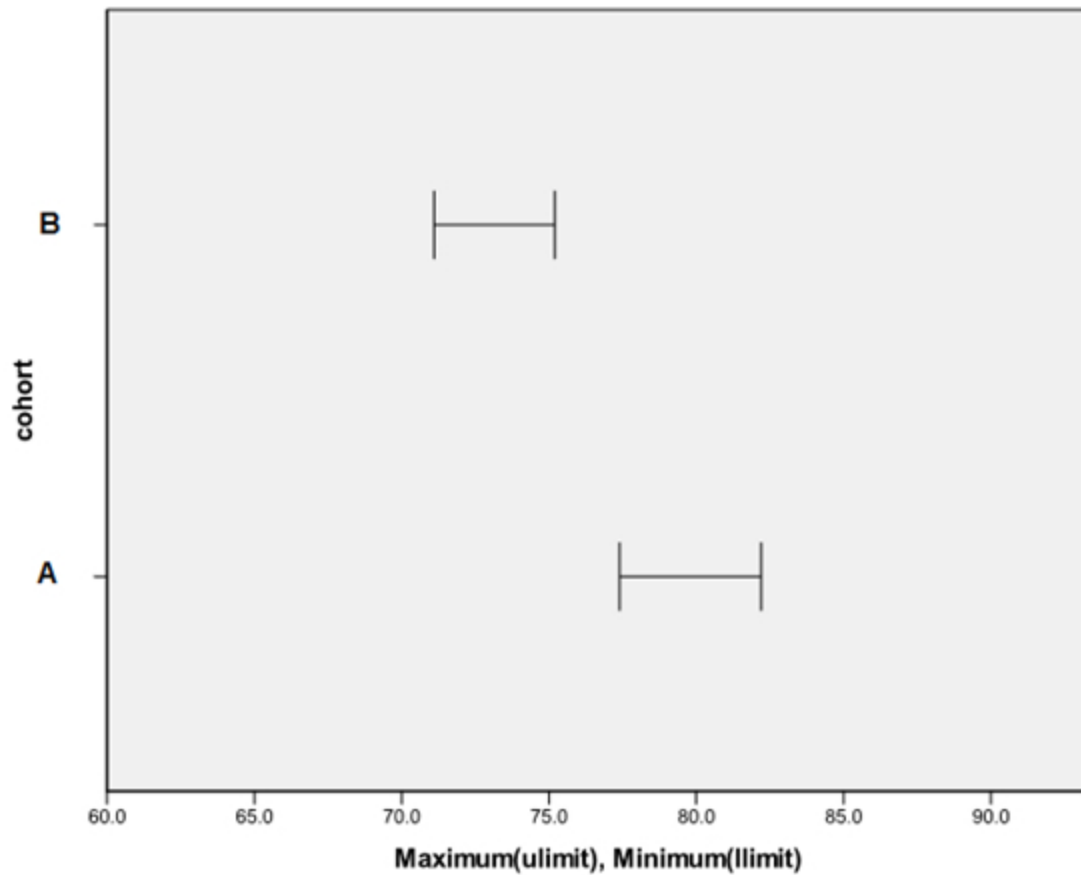


Figure 3. 95% Confidence Intervals for 1-Year Retention Rates: Cohort A versus Cohort B

If the template calculations had in fact resulted in a decision to retain the null hypothesis of no difference then a result such as that shown in the “Test Decision” field in Figure 4 would have occurred.

Pooled estimate for overall proportion	p_{s1}	p_{s2}	z statistic ¹	df	absolute value: critical value	z - crit val	<u>Test Decision</u>
0.47	0.49	0.46	1.68	2,534	1.96	-0.28	RETAIN NULL

Figure 4. Test of the Difference between Proportions (Level of Significance=.05)

[Request a copy](#) of the Microsoft Excel template shown above.