Chi-squared test for homogeneity

Post graduation activities of graduates from 4 colleges at an upstate NY University

	Agriculture	Arts & Sciences	Engineering	Social Science	Total
Employed	379	305	243	125	**1052**
Grad School	186	238	202	96	**722**
Other	104	123	37	58	**322**
Total	**669**	**666**	**482**	**279**	**2096**

because class sizes different proportions are easier to compare

	Agriculture	Arts & Sciences	Engineering	Social Science	Total
Employed	56.7%	45.8%	50.4%	44.8%	**50.2**
Grad School	27.8	35.7	41.9	34.4	**34.4**
Other	15.5	18.5	7.7	20.8	**15.4**
Total	**100**	**100**	**100**	**100**	**100**

We want to test whether the students' choices are the same across all four colleges?

Chi-squared test for homogeneity

sameness

*Test to determine whether several populations are similar or same in some characteristics

Conditions:

Counted data condition

Independence Assumption

Randomization Condition (check only if you are generalizing results)

Expected cell frequency condition

df = (rows-1)(columns-1)

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp}$$

H_0: The distribution does not change from group to group

H_A: Groups do not have the same distribution

Follow up with chi-squared -> pvalue-> conclusion

Where are the expected values going to come from?

	Agriculture	Arts & Sciences	Engineering	Social Science	Total
Employed	**335.777**	334.271	241.920	140.032	**1052**
Grad School	230.448	229.414	166.032	96.106	**722**
Other	102.776	102.315	74.048	42.862	**322**
Total	**669**	**666**	**482**	**279**	**2096**

Table 26.3 Expected values for the '06 graduates.

For example, overall, 1052, or about 50.2%, of the 2096 students who responded to the survey were employed. If the distributions are homogeneous (as the null hypothesis asserts), then 50.2% of the 669 Agriculture school graduates (or about 335.8 students) should be employed. Similarly, 50.2% of the 482 Engineering grads (or about 241.96) should be employed.

Follow with step by step example then practice.