CHI-SQUARE

Dr. Hamza Alduraidi

STEPS IN TEST OF HYPOTHESIS

1. Determine the appropriate test
2. Establish the level of significance: α
3. Formulate the statistical hypothesis
4. Calculate the test statistic
5. Determine the degree of freedom
6. Compare computed test statistic against a tabled/critical value

1. DETERMINE APPROPRIATE TEST

- Chi Square is used when both variables are measured on a nominal scale.
- It can be applied to interval or ratio data that have been categorized into a small number of groups.
- It assumes that the observations are randomly sampled from the population.
- All observations are independent (an individual can appear only once in a table and there are no overlapping categories).
- It does not make any assumptions about the shape of the distribution nor about the homogeneity of variances.

2. ESTABLISH LEVEL OF SIGNIFICANCE

- α is a predetermined value
- The convention
  - $\alpha = .05$
  - $\alpha = .01$
  - $\alpha = .001$

3. DETERMINE THE HYPOTHESIS: WHETHER THERE IS AN ASSOCIATION OR NOT

- H_o : The two variables are independent
- H_a : The two variables are associated

4. CALCULATING TEST STATISTICS

- Contrasts observed frequencies in each cell of a contingency table with expected frequencies.
- The expected frequencies represent the number of cases that would be found in each cell if the null hypothesis were true (i.e. the nominal variables are unrelated).
- Expected frequency of two unrelated events is product of the row and column frequency divided by number of cases.

$$F_e = F_r F_c / N$$

$$\text{Expected frequency} = \frac{\text{row total} \times \text{column total}}{\text{Grand total}}$$

4. CALCULATING TEST STATISTICS

Continued

$$\chi^2 = \sum \left[\frac{(F_o - F_e)^2}{F_e} \right]$$

4. CALCULATING TEST STATISTICS

Continued

Observed frequencies

$$\chi^2 = \sum \left[\frac{(F_o - F_e)^2}{F_e} \right]$$

Expected frequency

Expected frequency

5. DETERMINE DEGREES OF FREEDOM

$$df = (R-1)(C-1)$$

Number of levels in row variable

Number of levels in column variable

6. COMPARE COMPUTED TEST STATISTIC AGAINST A TABLED/CRITICAL VALUE

The computed value of the Pearson chi- square statistic is compared with the critical value to determine if the computed value is *improbable*

The critical tabled values are based on sampling distributions of the Pearson chi-square statistic

If calculated χ^2 is greater than χ^2 table value, reject H_o

χ^2

	$P(X \le x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21

DECISION AND INTERPRETATION

- If the probability of the test statistic is less than or equal to the probability of the alpha error rate, we reject the null hypothesis and conclude that our data supports the research hypothesis. We conclude that there is a relationship between the variables.
- If the probability of the test statistic is greater than the probability of the alpha error rate, we fail to reject the null hypothesis. We conclude that there is no relationship between the variables, i.e. they are independent.