

Chi-square Goodness of Fit Test

The *Chi-square Goodness of Fit Test* is designed to test whether the observed distribution is significantly different than the null distribution. This also can be viewed as a test of the observed proportions compared to the theoretical proportions.

Null Hypothesis: The population distribution is equivalent to the theoretical distribution.

Group	Theoretical Proportion
Traveled outside U.S.	
Never Traveled outside U.S.	

Alternative Hypothesis: The population distribution is significantly different than the theoretical distribution.

Characteristics of Chi-square Distributions

- ❑ Chi-square tests have a minimum value of zero.
- ❑ Chi-square distributions are positively skewed.
- ❑ The *expected value* of a chi-square test equals the degrees of freedom.
- ❑ The *variance* of a chi-square test equals 2(degrees of freedom).

The **Chi-square goodness of fit** statistic with $C - 1$ **degrees of freedom** should be compared to a critical value from a chi-square table.

In order for the critical values from the chi-square table to be appropriate, the *expected N for each category* should be 5 or larger. However, if the degrees of freedom is 1 then the expected value of each category should be 10 or larger.

Assumption of the Chi-square Goodness of Fit Test

1. Independence of the observations.

Stated in another fashion, each subject must fit into one and only one category and the categories must be chosen so as to include all subjects.

A psychologist sampled 200 college students and found that 75 had traveled outside the United States. Based on the national travel data for adults, 25% of all US citizens had traveled outside the United States. Do college students have different travel patterns than other US citizens? Use an alpha of .05.

Null Hypothesis:

Group	Theoretical Proportion
Traveled outside U.S.	
Never Traveled outside U.S.	

Npar Tests Chi-Square Test Frequencies

Outside U.S.

	Observed N	Expected N	Residual
No	125	150.0	-25.0
Yes	75	50.0	25.0
Total	200		

Test Statistics

	Outside U.S.
Chi-Square ^a	16.667
df	1
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 50.0.

An instructor summarized all of the grades given for the past 10 years to determine whether the grades matched a 10% A, 20% B, 40% C, 20% D, 10% F distribution. The instructor had given 350 As, 500 Bs, 480 Cs, 290 Ds, and 250 Fs during the previous 10 years. Use an alpha of .05 to test the null hypothesis.

Null Hypothesis:

Grade	Theoretical Proportion
A	
B	
C	
D	
F	

NPar Tests Chi-Square Test Frequencies

GradeGiven

	Observed N	Expected N	Residual
F	250	187.0	63.0
D	290	374.0	-84.0
C	480	748.0	-268.0
B	500	374.0	126.0
A	350	187.0	163.0
Total	1870		

Test Statistics

	GradeGiven
Chi-Square ^a	320.642
df	4
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 187.0.

CHI-SQUARE GOODNESS OF FIT TEST BY HAND
--

A psychologist sampled 200 college students and found that 75 had traveled outside the United States. Based on the national travel data for adults, 25% of all US citizens had traveled outside the United States. Do college students have different travel patterns than other US citizens? Use an alpha of .05.

Null Hypothesis (H_0):

Group	Theoretical Proportion

Alpha Level =

Degrees of Freedom = Number of Groups – 1

Chi-square Critical Value =

*Expected N = theoretical proportion * Total Sample Size*

Group	Observed <i>N</i>	Theoretical Proportion	Expected <i>N</i>	Observed <i>N</i> – Expected <i>N</i>	$\frac{(\text{Observed } N - \text{Expected } N)^2}{\text{Expected } N}$
Total Sample Size					chi-square =

Compare the calculated Chi-square to the Chi-square Critical Value.

Interpret the Chi-square Goodness of Fit results.

(Retain Null, $p > .05$). The Chi-square Goodness of Fit test indicated that college students preferences for burgers, pizza, and Mexican food matched the theoretical proportions of 33.3% burgers, 33.3% pizza, 33.3% Mexican, $\chi^2(2, N = 118) = 2.11, p > .05$. Forty percent preferred pizza, 30% preferred burgers, and 30% preferred Mexican.

(Reject Null, $p < .05$). The Chi-square Goodness of Fit test indicated that college students' preferences for burgers, pizza, and Mexican food were different than the theoretical proportions of 33.3% burgers, 33.3% pizza, 33.3% Mexican, $\chi^2(2, N = 118) = 4.81, p < .05$. Forty-five percent preferred pizza, 40% preferred burgers, and 15% preferred Mexican.

Chi-square Test of Independence

The chi-square test of independence is designed to determine whether two categorical variables are independent of each other.

Null Hypothesis: Variable A is not related to variable B.

Alternative Hypothesis: Variable A is related to variable B.

Characteristics of Chi-square Distributions

- ❑ Chi-square tests have a minimum value of zero.
- ❑ Chi-square distributions are positively skewed.
- ❑ The *expected value* of a chi-square test equals the degrees of freedom.
- ❑ The *variance* of a chi-square test equals 2(degrees of freedom).

The **Chi-square goodness of fit** statistic with $(R - 1)(C - 1)$ **degrees of freedom** should be compared to a critical value from a chi-square table.

In order for the critical values from the chi-square table to be appropriate, the *expected N for each category* should be 5 or larger. However, if the degrees of freedom is 1 then the expected value of each category should be 10 or larger.

Assumption of the Chi-square Test of Independence

1. Independence of the observations.

Stated in another fashion, each subject must fit into one and only one category and the categories must be chosen so as to include all subjects.

A researcher is interested in determining if there is a relationship between sex education and unplanned pregnancies. 240 college females were categorized according to whether they (a) had participated in a sex education class and (b) had participated in an unplanned pregnancy. Use an alpha of .05 to test the null hypothesis.

Pregnancy	Sex Education		
	Educated	Not Educated	
Unplanned	50	80	130
Planned	75	35	110
	125	115	240

Null Hypothesis.

Alternative Hypothesis.

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
pregnancy * sexeduc	240	100.0%	0	.0%	240	100.0%

pregnancy * sexeduc Crosstabulation

			sexeduc		Total
			No Sex Ed	Sex Ed	
pregnancy	unplanned	Count	80	50	130
		Expected Count	62.3	67.7	130.0
		% within pregnancy	61.5%	38.5%	100.0%
planned	Count	Count	35	75	110
		Expected Count	52.7	57.3	110.0
		% within pregnancy	31.8%	68.2%	100.0%
Total	Count	Count	115	125	240
		Expected Count	115.0	125.0	240.0
		% within pregnancy	47.9%	52.1%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	21.088 ^b	1	.000		
Continuity Correction ^a	19.914	1	.000		
Likelihood Ratio	21.453	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	21.001	1	.000		
N of Valid Cases	240				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 52.71.

Radelet (1981) studied the effect of racial characteristics on whether individuals convicted of homicide receive the death penalty. The 326 subjects were defendants in homicide indictments in 20 Florida counties during 1976 – 1977. Use an alpha of .05 to test the null hypothesis.

Victim's Race	Death Penalty		
	Yes	No	
White	30	184	214
Black	6	106	112
	36	290	326

Null Hypothesis.

Alternative Hypothesis.

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
victimrace * deathpenalty	326	100.0%	0	.0%	326	100.0%

victimrace * deathpenalty Crosstabulation

			deathpenalty		Total
			No	Yes	
victimrace	Black	Count	106	6	112
		Expected Count	99.6	12.4	112.0
		% within victimrace	94.6%	5.4%	100.0%
White	Count	Count	184	30	214
		Expected Count	190.4	23.6	214.0
		% within victimrace	86.0%	14.0%	100.0%
Total	Count	Count	290	36	326
		Expected Count	290.0	36.0	326.0
		% within victimrace	89.0%	11.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.615 ^b	1	.018		
Continuity Correction ^a	4.768	1	.029		
Likelihood Ratio	6.250	1	.012		
Fisher's Exact Test				.024	.012
Linear-by-Linear Association	5.598	1	.018		
N of Valid Cases	326				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 12.37.

Chi-square Test of Independence By Hand

A researcher is interested in determining if there is a relationship between sex education and unplanned pregnancies. 240 college females were categorized according to whether they (a) had participated in a sex education class and (b) had participated in an unplanned pregnancy. Use an alpha of .05 to test the null hypothesis.

	Sex Education		
	Educated	Not Educated	
Pregnancy			
Unplanned	50	80	130
Planned	75	35	110
	125	115	240

Null Hypothesis:

Alpha Level =

Degrees of Freedom = (Number of Rows – 1)(Number of Columns – 1)

Chi-square Critical Value =

Expected N = (Row Total * Column Total) / Total Sample Size

Group	Observed <i>N</i>	Expected <i>N</i>	Observed <i>N</i> – Expected <i>N</i>	$\frac{(\text{Observed } N - \text{Expected } N)^2}{\text{Expected } N}$
Total Sample Size				chi-square =

Compare the calculated Chi-square to the Chi-square Critical Value.

Interpret the Chi-square Test of Independence results.

(Retain Null, $p > .05$). The Chi-square Test of Independence indicated that the likelihood of receiving a ticket for speeding is not significantly related to one's gender, $\chi^2(1, N = 135) = 2.11, p > .05$.

(Reject Null, $p < .05$). The Chi-square Test of Independence indicated that the likelihood of receiving a ticket for speeding is significantly related to one's gender, $\chi^2(1, N = 135) = 5.74, p < .05$.

Chi-square Test of Independence By Hand

Radelet (1981) studied the effect of racial characteristics on whether individuals convicted of homicide receive the death penalty. The 326 subjects were defendants in homicide indictments in 20 Florida counties during 1976 – 1977. Use an alpha of .05 to test the null hypothesis.

Victim's Race	Death Penalty		
	Yes	No	
White	30	184	214
Black	6	106	112
	36	290	326

Null Hypothesis:

Alpha Level =

Degrees of Freedom = (Number of Rows – 1)(Number of Columns – 1)

Chi-square Critical Value =

Expected N = (Row Total * Column Total) / Total Sample Size

Group	Observed N	Expected N	Observed N – Expected N	$\frac{(\text{Observed } N - \text{Expected } N)^2}{\text{Expected } N}$
Total Sample Size				chi-square =

Compare the calculated Chi-square to the Chi-square Critical Value.

Interpret the Chi-square Test of Independence results.

CRITICAL VALUES of the Chi-Square Distribution

DF	Proportion in Right Tail					
	.25	.10	.05	.025	.01	.005
1	1.323	2.706	3.841	5.024	6.635	7.879
2	2.773	4.605	5.991	7.378	9.210	10.597
3	4.108	6.251	7.815	9.348	11.345	12.838
4	5.385	7.779	9.488	11.143	13.277	14.860
5	6.626	9.236	11.070	12.833	15.086	16.750
6	7.841	10.645	12.592	14.449	16.812	18.548
7	9.037	12.017	14.067	16.013	18.475	20.278
8	10.219	13.362	15.507	17.535	20.090	21.955
9	11.389	14.684	16.919	19.023	21.666	23.589
10	12.549	15.987	18.307	20.483	23.209	25.188
11	13.701	17.275	19.675	21.920	24.725	26.757
12	14.845	18.549	21.026	23.337	26.217	28.300
13	15.984	19.812	22.362	24.736	27.688	29.819
14	17.117	21.064	23.685	26.119	29.141	31.319
15	18.245	22.307	24.996	27.488	30.578	32.801
16	19.369	23.542	26.296	28.845	32.000	34.267
17	20.489	24.769	27.587	30.191	33.409	35.718
18	21.605	25.989	28.869	31.526	34.805	37.156
19	22.718	27.204	30.144	32.852	36.191	38.582
20	23.828	28.412	31.410	34.170	37.566	39.997
21	24.935	29.615	32.671	35.479	38.932	41.401
22	26.039	30.813	33.924	36.781	40.289	42.796
23	27.141	32.007	35.172	38.076	41.638	44.181
24	28.241	33.196	36.415	39.364	42.980	45.559
25	29.339	34.382	37.652	40.646	44.314	46.928
26	30.435	35.563	38.885	41.923	45.642	48.290
27	31.528	36.741	40.113	43.195	46.963	49.645
28	32.620	37.916	41.337	44.461	48.278	50.993
29	33.711	39.087	42.557	45.722	49.588	52.336
30	34.800	40.256	43.773	46.979	50.892	53.672
31	35.887	41.422	44.985	48.232	52.191	55.003
32	36.973	42.585	46.194	49.480	53.486	56.328
33	38.058	43.745	47.400	50.725	54.776	57.648
34	39.141	44.903	48.602	51.966	56.061	58.964
35	40.223	46.059	49.802	53.203	57.342	60.275
36	41.304	47.212	50.998	54.437	58.619	61.581
37	42.383	48.363	52.192	55.668	59.893	62.883
38	43.462	49.513	53.384	56.896	61.162	64.181
39	44.539	50.660	54.572	58.120	62.428	65.476
40	45.616	51.805	55.758	59.342	63.691	66.766
41	46.692	52.949	56.942	60.561	64.950	68.053
42	47.766	54.090	58.124	61.777	66.206	69.336
43	48.840	55.230	59.304	62.990	67.459	70.616
44	49.913	56.369	60.481	64.201	68.710	71.893
45	50.985	57.505	61.656	65.410	69.957	73.166
46	52.056	58.641	62.830	66.617	71.201	74.437
47	53.127	59.774	64.001	67.821	72.443	75.704
48	54.196	60.907	65.171	69.023	73.683	76.969
49	55.265	62.038	66.339	70.222	74.919	78.231
50	56.334	63.167	67.505	71.420	76.154	79.490

CRITICAL VALUES of the Chi-Square Distribution

DF	Proportion in Right Tail					
	.25	.10	.05	.025	.01	.005
51	57.401	64.295	68.669	72.616	77.386	80.747
52	58.468	65.422	69.832	73.810	78.616	82.001
53	59.534	66.548	70.993	75.002	79.843	83.253
54	60.600	67.673	72.153	76.192	81.069	84.502
55	61.665	68.796	73.311	77.380	82.292	85.749
56	62.729	69.919	74.468	78.567	83.513	86.994
57	63.793	71.040	75.624	79.752	84.733	88.236
58	64.857	72.160	76.778	80.936	85.950	89.477
59	65.919	73.279	77.931	82.117	87.166	90.715
60	66.981	74.397	79.082	83.298	88.379	91.952
61	68.043	75.514	80.232	84.476	89.591	93.186
62	69.104	76.630	81.381	85.654	90.802	94.419
63	70.165	77.745	82.529	86.830	92.010	95.649
64	71.225	78.860	83.675	88.004	93.217	96.878
65	72.285	79.973	84.821	89.177	94.422	98.105
66	73.344	81.085	85.965	90.349	95.626	99.330
67	74.403	82.197	87.108	91.519	96.828	100.554
68	75.461	83.308	88.250	92.689	98.028	101.776
69	76.519	84.418	89.391	93.856	99.228	102.996
70	77.577	85.527	90.531	95.023	100.425	104.215
71	78.634	86.635	91.670	96.189	101.621	105.432
72	79.690	87.743	92.808	97.353	102.816	106.648
73	80.747	88.850	93.945	98.516	104.010	107.862
74	81.803	89.956	95.081	99.678	105.202	109.074
75	82.858	91.061	96.217	100.839	106.393	110.286
76	83.913	92.166	97.351	101.999	107.583	111.495
77	84.968	93.270	98.484	103.158	108.771	112.704
78	86.022	94.374	99.617	104.316	109.958	113.911
79	87.077	95.476	100.749	105.473	111.144	115.117
80	88.130	96.578	101.879	106.629	112.329	116.321
81	89.184	97.680	103.010	107.783	113.512	117.524
82	90.237	98.780	104.139	108.937	114.695	118.726
83	91.289	99.880	105.267	110.090	115.876	119.927
84	92.342	100.980	106.395	111.242	117.057	121.126
85	93.394	102.079	107.522	112.393	118.236	122.325
86	94.446	103.177	108.648	113.544	119.414	123.522
87	95.497	104.275	109.773	114.693	120.591	124.718
88	96.548	105.372	110.898	115.841	121.767	125.913
89	97.599	106.469	112.022	116.989	122.942	127.106
90	98.650	107.565	113.145	118.136	124.116	128.299
91	99.700	108.661	114.268	119.282	125.289	129.491
92	100.750	109.756	115.390	120.427	126.462	130.681
93	101.800	110.850	116.511	121.571	127.633	131.871
94	102.850	111.944	117.632	122.715	128.803	133.059
95	103.899	113.038	118.752	123.858	129.973	134.247
96	104.948	114.131	119.871	125.000	131.141	135.433
97	105.997	115.223	120.990	126.141	132.309	136.619
98	107.045	116.315	122.108	127.282	133.476	137.803
99	108.093	117.407	123.225	128.422	134.642	138.987
100	109.141	118.498	124.342	129.561	135.807	140.169