Contingency Tables

χ^2 Test

How effective are various treatments when dealing with medical issues? The use of **Contingency Tables** attempts to answer this relevant and important question. And the process can generally be used to test the "independence" of various processes, or attributes, on different subjects. Another illustration of how a **Contingency Table** works is the consideration of people who are left-handed versus right-handed.

Sex Versus Handedness

	Right Handed	Left Handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

When considering **two variables** (**sex** and **handedness**) we want to determine whether these **variables** are **independent** of one another. Let us focus on some technical definitions so we can proceed in answering these questions. We consider statistical methods to **analyze the frequency counts**. In our "contingency table" (two-way frequency table.

Definition- Contingency Table: A table consisting of frequency counts for categorical data corresponding to two variables. One variable is used to categorize the rows, while the other variable is used to categorize the columns.

x represents the sex (male or female) and y represents the handedness (left and right).

The word **contingent** refers to dependence on some other factor. The term **contingency table** refers to a test regarding the independency between the variables x and y (rows and columns).

Definition- Test of Independence: when we test the Null Hypothesis that in a contingency table, the row and columns variables are independent. There is no dependency between the roe and column variables.

Requirements

- The sample data is randomly collected.
- The sample data is represented as frequency counts in rows and columns (two-way table).
- For every cell in the **Contingency Table**, the **Expected Frequency E** is at least 5.

Hypothesis Test

H_0: The row and columns variables are Independent.
H_1: The row and columns variables are dependent.

Test Statistic $\chi^2 = \sum \frac{(O-E)^2}{E}$ where O is the observed frequency and E is the expected frequency in a cell and

$$E = \frac{(row\ total)(column\ total)}{(grand\ total)}$$

Critical Values

$$df = (r-1)(c-1)$$

Where r is the number of rows and c is the number of columns. This is also a right tail χ^2 **distribution** test.

TABLE A-4 Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of the Critical Value									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Traditional Method

H_0: The row and columns variables are Independent.
H_1: The row and columns variables are dependent.

Sex Versus Handedness

	Right Handed	Left Handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

		Column 1	Column 2	
		Right Handed	**Left Handed**	**Total**
Row 1	**Male**	E_{11}	E_{12}	52
Row 2	**Female**	E_{21}	E_{22}	48
	Total	87	13	100

$$\boldsymbol{E}$$

		Column 1	Column 2	
		Right Handed	**Left Handed**	**Total**
Row 1	**Male**	45.24	6.76	52
Row 2	**Female**	41.76	6.24	48
	Total	87	13	100

$$\boldsymbol{O - E}$$

		Column 1	Column 2	
		Right Handed	**Left Handed**	**Total**
Row 1	**Male**	-2.24	2.24	52
Row 2	**Female**	2.24	-2.24	48
	Total	87	13	100

$$\boldsymbol{(O - E)^2}$$

		Column 1	Column 2	
		Right Handed	**Left Handed**	**Total**
Row 1	**Male**	5.0176	5.0176	52
Row 2	**Female**	5.0176	5.0176	48
	Total	87	13	100

$$\frac{\boldsymbol{(O - E)\text{^}2}}{\boldsymbol{E}}$$

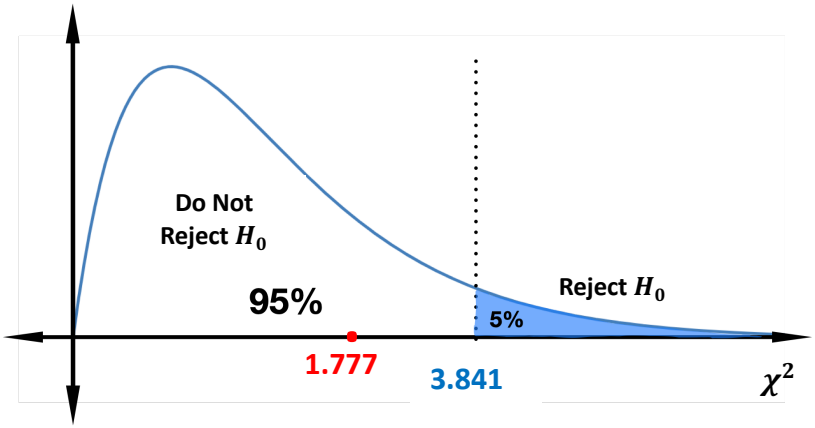
		Column 1	Column 2	
		Right Handed	**Left Handed**	**Total**
Row 1	**Male**	0.111	0.742	52
Row 2	**Female**	0.120	0.804	48
	Total	87	13	100
Sum	**1.777**			

$$\chi^2 = \sum \frac{(O-E)^2}{E} \approx 1.777$$

$$df = (2-1)(2-1) = 1 \text{ using } \alpha = 5\%$$

TABLE A-4 Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of the Critical Value									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169



Conclusion

Do Not Reject H_0

The Sex and Handedness are Independent.

Modern Method

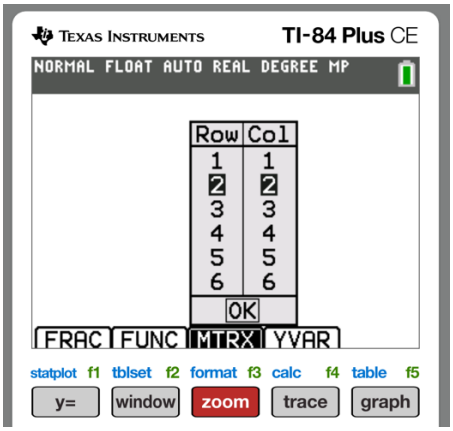
H_0: The row and columns variables are Independent.
H_1: The row and columns variables are dependent.

Sex Versus Handedness

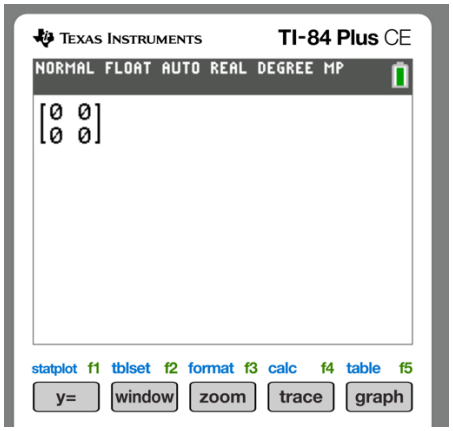
	Right Handed	Left Handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

Create a Matrix

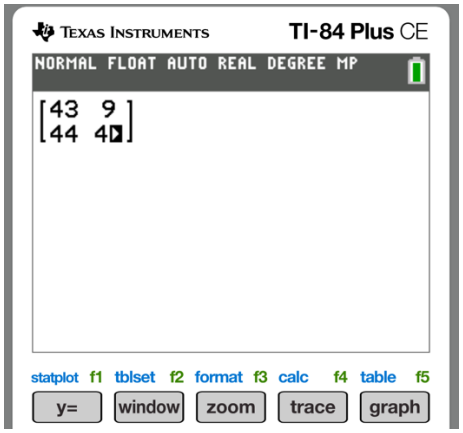
$$alpha > zoom$$



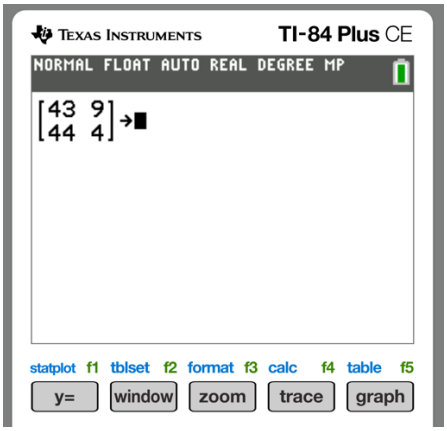
Enter the dimensions, in this case it's a 2x2 matrix



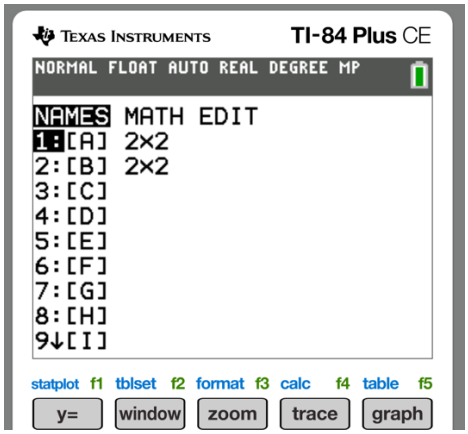
Enter the row and column values and don not Press Enter



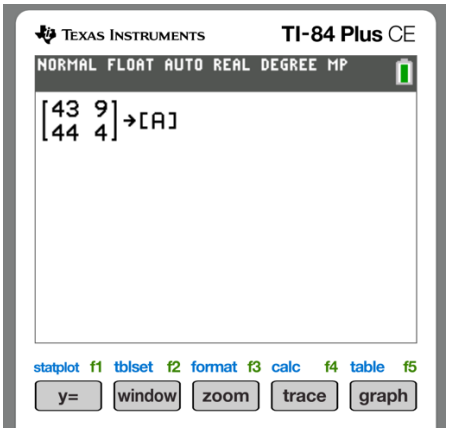
Store the Matrix
Move the arrow to the right and use the $sto \rightarrow$ key



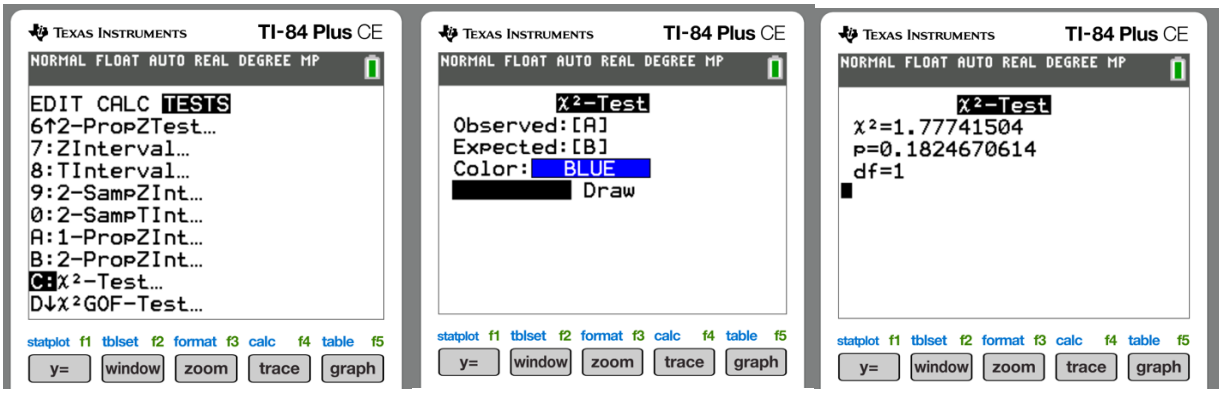
Now name the Matrix
$2nd > x^{-1}$



Press Enter



Conduct a χ^2 *Test*



$p \approx 0.182; p \prec \alpha$; **Accept H_0**

The Sample Supports the Claim

The Sex and Handedness are Independent

Splint or Surgery?

Treatment versus Success

	Successful Treatment	Unsuccessful Treatment	Total
Splint Treatment	60	23	83
Surgery treatment	67	6	73
Total	127	29	156

Traditional Method using 1% Level of Significance

H_0: The row and columns variables are independent.
H_1: The row and columns variables are dependent.

	Successful Treatment	Unsuccessful Treatment	Total
Splint Treatment	E_{11}	E_{12}	83
Surgery treatment	E_{21}	E_{22}	73
Total	127	29	156

$$\boldsymbol{E}$$

		Column 1	Column 2
		Successful Treatment	**Unsuccessful Treatment**
Row 1	**Splint Treatment**	67.5705	15.4295
Row 2	**Surgery treatment**	59.4295	13.5705

$$\boldsymbol{O - E}$$

		Column 1	Column 2
		Successful Treatment	**Unsuccessful Treatment**
Row 1	**Splint Treatment**	-7.5705	7.5705
Row 2	**Surgery treatment**	7.5705	-7.5705

$$\boldsymbol{(O - E)^2}$$

		Successful Treatment	Unsuccessful Treatment
Row 1	**Splint Treatment**	57.31247025	57.31247025
Row 2	**Surgery treatment**	57.31247025	57.31247025

$$\frac{\boldsymbol{(O - E)\text{^}2}}{\boldsymbol{E}}$$

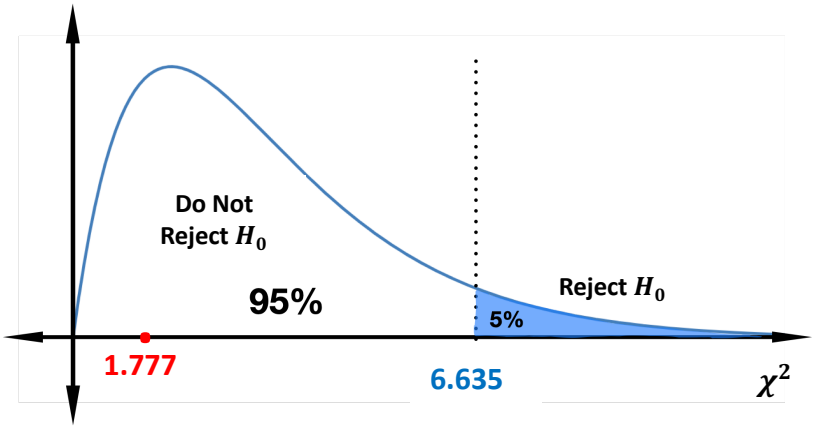
		Column 1	Column 2
		Successful Treatment	**Unsuccessful Treatment**
Row 1	**Splint Treatment**	0.848187748	3.71447359
Row 2	**Surgery treatment**	0.96437746	4.223313087
	Sum	**9.750**	

$$\chi^2 = \sum \frac{(O-E)^2}{E} \approx 9.750$$

$$df = (2-1)(2-1) = 1 \text{ using } \alpha = 1\%$$

TABLE A-4 Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of the Critical Value									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169



Conclusion

Do Not Reject H_0

The Treatment and Success are Independent.

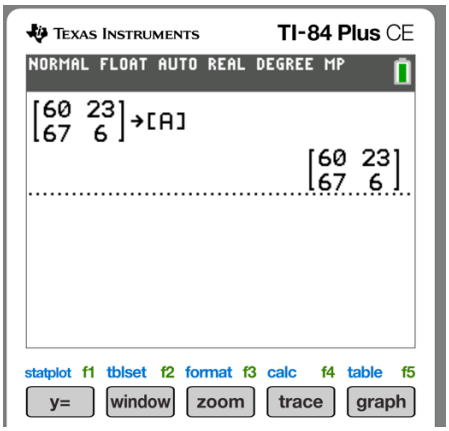
However, surgery is more likely to be successful.

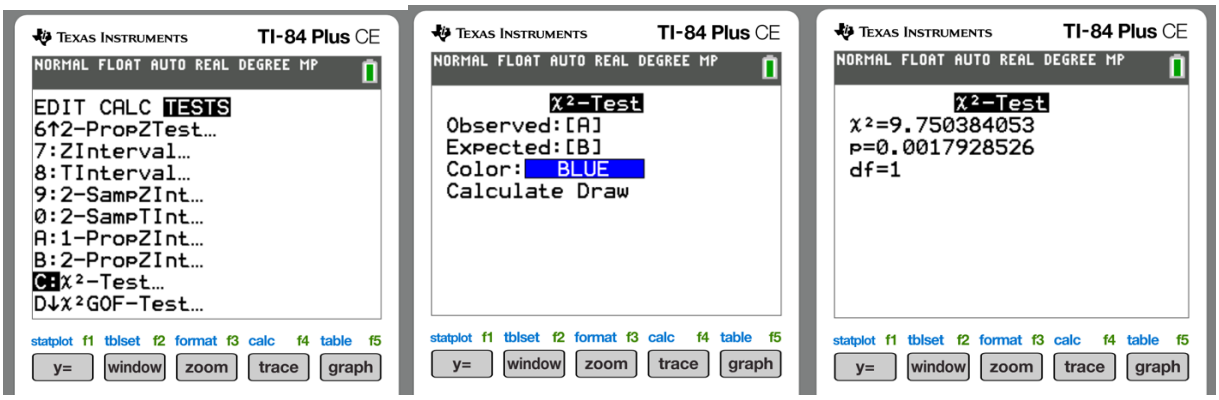
Modern Method

H_0: The row and columns variables are Independent.
H_1: The row and columns variables are dependent.

Treatment versus Success

	Successful Treatment	Unsuccessful Treatment	Total
Splint Treatment	60	23	83
Surgery treatment	67	6	73
Total	127	29	156





$p \approx 0.002$; $p < \alpha$; p value is low, H_0 has to go!
The Sample suggests the results are dependent.
Surgery is a better treatment option.