2014

Calculating P-Values

Isela Guerra
Parkland College

Isela Guerra
Instructor: L. Smith

Calculating P- values

Part 1:

Ex: Employees in an accounting firm claim that the mean salary of the firms accountants is less than that of its competitor's which is $45,000. (Assume that the standard deviation of the firm's accountants is known to be $5,200.) A random sample of 30 of the firm's accountants has a mean salary of $43,500. Test the employees claim at the .01 level of significance.

Step 1: H_0: $\mu = 45000$ H_1: $\mu < 45000$

Step 2: $\alpha = .01$

Using Z-Test

Step 3:

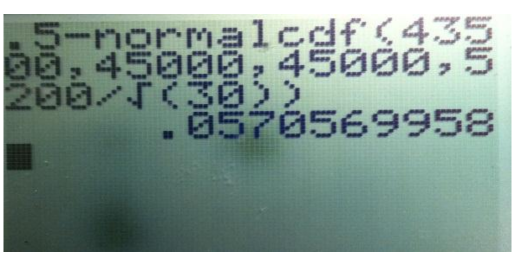
Figure 1

[Because this was a test about the mean, the population standard deviation is known and n≥30, the Z-test was selected. We are attempting to test the claim that the mean salary is less than $45,000. Our null hypothesis is then equal to $45,000. The argument is that the mean is less than $45,000 this becomes our alternative hypothesis.

The level of significance is already given, **α = .01**

Because there is no data set we use the Z- Test, to get there we click **"STATS"** and scroll over to **"TEST"** then scroll down to "**Z-Test**". This asks for the null hypothesis in the form of **"μ_0:"** the standard deviation denoted by **"σ",** the mean **"x̄"**, and **"n"** the sample size. Now, because we want to know if the **mean salary is less than** that of the competing firm we chose **"μ: <μ_0"**. We then click enter on "calculate" and get our p- value of **.0571**.**]**

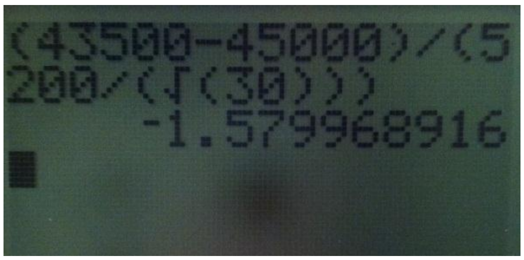
Using Normal CDF



Isela Guerra
Instructor: L. Smith

[We want to know what the area to the left of 43,500 is. To do this we use normalcdf and the syntax within normalcdf is (**x̄, x̄μ, μ, σ/ (n) ^1/2).** Because we want to know the area to the left of $4, 3500 (μ < 45000) we say, **.5-**normalcdf(43500,4500,4500,5200/(30)^1/2. Our p-value is **.0571**]

Using Table and calculating Z value



Z Score Table- chart value corresponds to area below z score.

z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
–3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
–3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
–3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
–3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
–3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
–2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
–2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
–2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
–2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
–2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
–2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
–2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
–2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
–2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
–2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
–1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
–1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
–1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
–1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
–1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668

[To get the z score we use the equation, $$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$, where **x̄** is the sample mean and **μ** is the population mean, and **σ/(n)^1/2** is the population standard deviation divided by the square root of the sample size. This gave us a Z-score of **-1.58** when rounded to three significant figures. To find the corresponding p-value we go to row labeled **Z**, scroll down to **-1.5** and to the right until we go down to the row containing **.008**.]

Step 4: p-value > α

Fail to reject the null hypothesis.

Step 5: There is not significant evidence at the .01 level of significance to conclude that the mean salary of the accountants is less than that of the competitors.

Part2:

Ex: Account empts survey 150 executives it showed that 44% of them say that “little or no knowledge of the company” is most common mistake made by candidates during job interviews (based on data from USA Today). Use a .05 significance level to test claim that less than half of all executives identify that error as being most common job interviewing error.

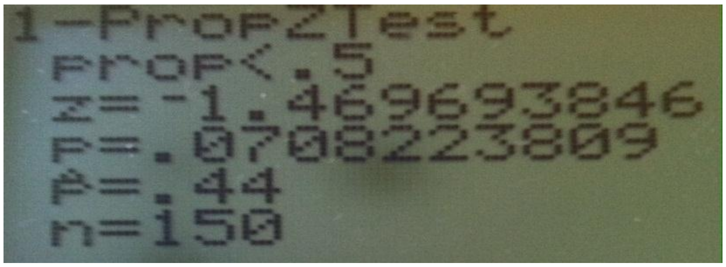
Step 1: H_0: p = .5 H_1: p < .5

Step 2: α = .05

Isela Guerra
Instructor: L. Smith

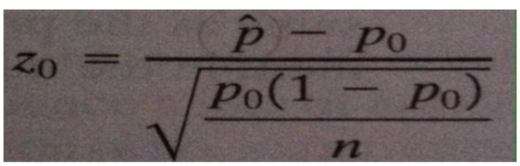
Step 3:

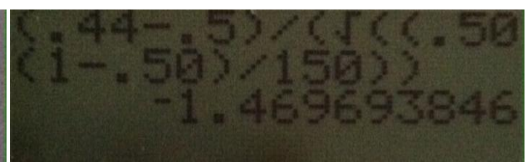
Using 1PropZTest

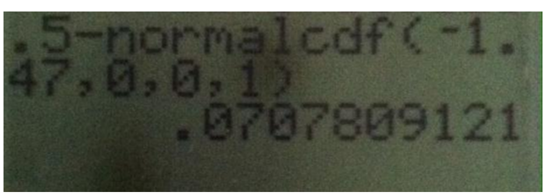


[We use **1-PropZTest** because we have to deal with proportions. We are given the proportion we are to test, **p_0: .5,** we are also given the proportion of those with the given characteristic we are testing for, which is .**44** however to get **X** we multiply **(.44) (150)** and get **66**. We enter these values into 1-PropZTest and get a p- value of **.0708**]

Using Normal CDF



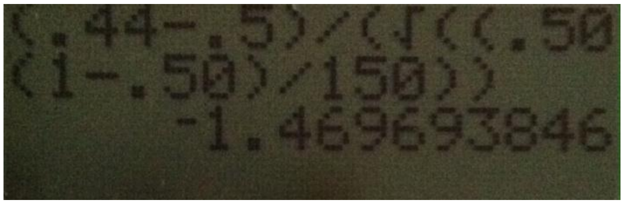




[Because there was no standard deviation given, the normal cdf could not be used until the curve was standardized. To do this the Z score was calculated by conducting the test statistic, the **p** with a **carrot above** it is the population proportion with given characteristic(**.44**), and **p_0** is the population proportion(**.5**). The mean is then equal to zero and the standard deviation is equal to 1. Our value calculated was **-1.47** when rounded we then use this value for the normalcdf and get our p-value of **.0708** when rounded.]

Isela Guerra
Instructor: L. Smith

Using Table and calculating Z value



Negative z-scores:

Z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.0
-3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
-3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
-3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
-3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808

-1.4	0.0681	0.0694	0.0708

[Using the value found from the test statistic, the p-value was located on the table. The calculated z-value is -1.47 when rounded. Go to row labeled Z and stop when **-1.4** is hit, proceed to the right when the row **.07** is reached. The p-value is **.0708**.]

Step 4: p-value > α

Fail to reject the null hypothesis.

Step 5: There is not significant evidence at the .05 level of significance to claim that less than half of all executives identify that error as being most common job interviewing error.

Part 3:

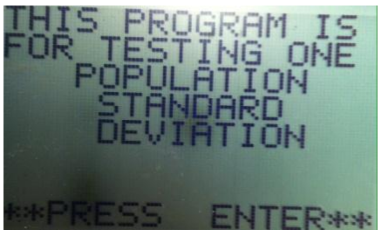
Ex: Tests in past Statistics classes have had scores with a standard deviation equal to 14.1. A sample of 27 scores from the current classes has a standard deviation of 9.3. Use the 0.01 level of significance to test the claim that this current class is more consistent that past classes. Assume that test scores are normally distributed.

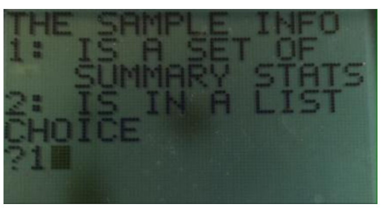
Step 1: $H_0: \mu = 14.1$ $H_1: \mu < 14.1$

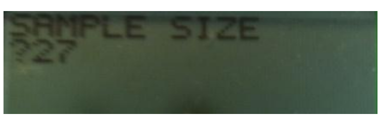
Step 2: $\alpha = .01$

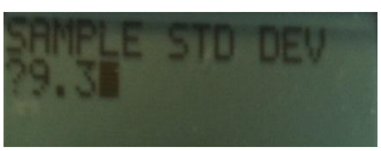
Step 3:

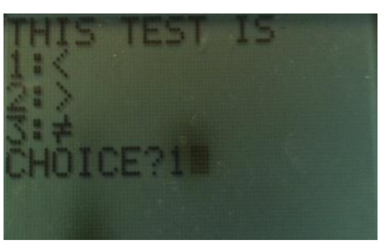
Using Test1 SD

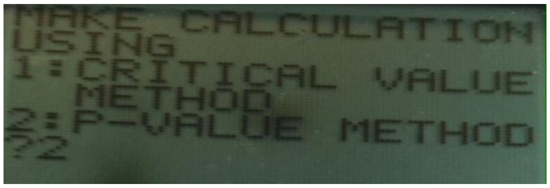


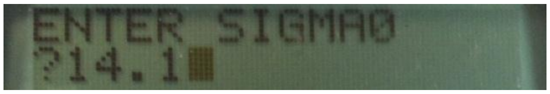


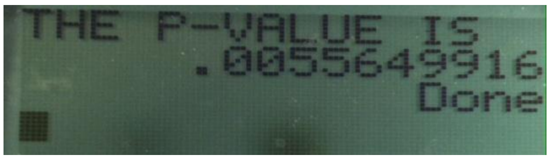








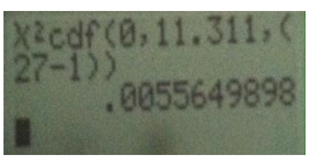




[This test can be found under programs. Because we were given a summary of the stats we select choice **2** under "**The sample info**" We press enter and then enter **27** for the **sample size** and again press **enter**. We then move along to the **sample** standard deviation and enter 9.3 and press enter then select "**<**" for less than because we are testing to see if 14.1 is **<** current classes. We would then like to calculate the p-value, after clicking **ente**r we then press **2** to select "**P-value Method**". After, we enter the **population** standard deviation of **14.1**. The p-value calculated is **.00556**.]

Using Normal Cdf (X^2 Chi square)

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$



[To calculate the **p-value** using **X^2**(chi Square) **cdf**, we first calculate the **critical value**. To do this we use the formula **X^2= ((n-1) $s^{2)}$/σ^2**. Where (**n-1**) is the degrees of freedom, **s** is the sample standard deviation and **σ** is the population standard deviation, however when we square this it then becomes the population variance (**σ^2**)**.** After calculating, we get the critical value of **11.311** when rounded. We then use this value in X^2cdf (**0, critical value, degrees of freedom**). This then gives us the p- value; we calculated a p-value of .00556 when rounded to three significant figures.]

Using Chi Square table

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
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.452	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	18.144	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.196	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.772	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	100.43	104.21
80	51.172	53.540	57.153	60.391	64.278	96.578	101.88	106.63	112.33	116.32
90	59.196	61.754	65.647	69.126	73.291	107.57	113.15	118.14	124.12	128.30
100	67.328	70.065	74.222	77.929	82.358	118.50	124.34	129.56	135.81	140.17

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
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.452	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

[When using the table we see degrees of freedom to the top left. We scroll down to 26 degrees of freedom, and we then see the p-value to the top right of **.005**.]

Step 4: p-value ≤ α

Reject the null hypothesis

Step 5: There is enough evidence at the .01 level of significance to conclude that this current class is more consistent that past classes.

Part4

EX: A firm wishes to estimate the economic composition of a community where they are opening a branch office. They have examined data concerning this and similar communities and have obtained a hypothetical distribution of income that they feel to be accurate. To check this accuracy, they take a random sample of 500 families and family incomes. They have hypothesized the following distribution of income levels:

Category	1	2	3	4	5
Income level	3000^{+}	25000-29999	20000-24999	15000-19999	Under 15000
Proportion	0.4	0.2	0.2	0.1	0.1

The sample yields the following results:

Category	1	2	3	4	5	Total
Number	166	97	134	61	42	500

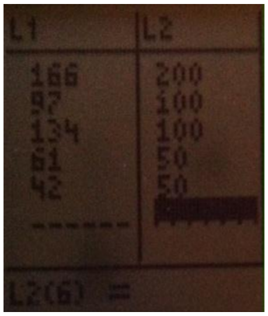
Do these results differ from the hypothesized distribution?

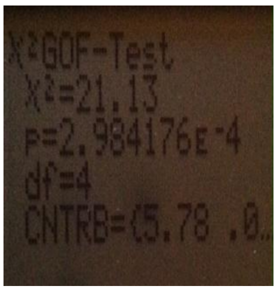
Step 1: H_0: $p_1=0.4$, $p_2=0.2$, $p_3=0.2$, $p_4=0.1$, $p_5=0.1$ H_1: The proportions differ from those specified.

Step 2: $\alpha=0.05$

Step 3:

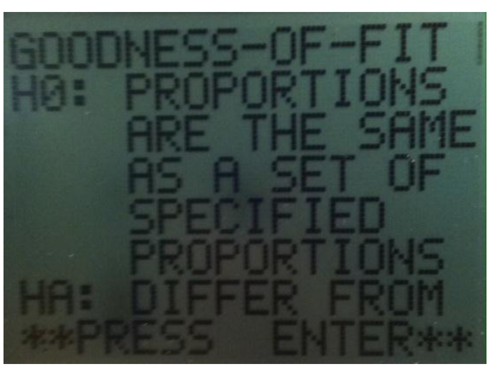
Using X^2- Goodness of Fit Test



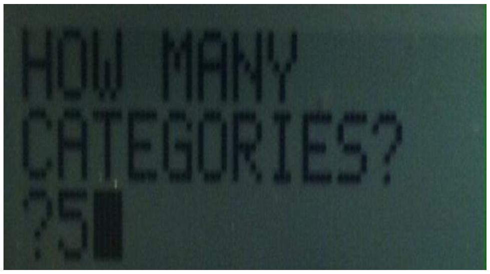


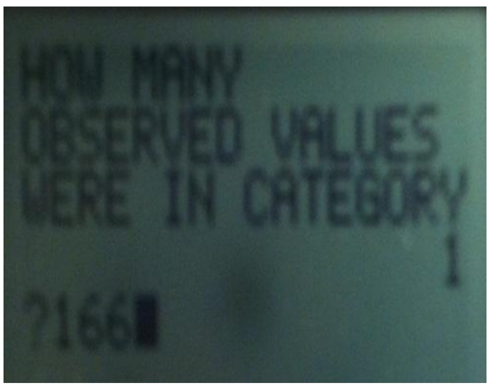
[To perform this test we place the **numbers observed** from each of the categories into $\mathbf{L_1}$, we then calculate the **proportions** which will go into $\mathbf{L_2}$. To do this we take the given proportions for each number and multiply them by the total. For example, (.4) (500) = 200. You can see that 200 is the first number on the list of L_2. We then go to **X^2 GOF- Test** and enter L_1 in observed and L_2 in expected, we then scroll down to **df** (degrees of freedom, found by (n-1)) and enter 4, then enter on calculate and record our **p-value** of **$2.98X10^{-4}$**.]

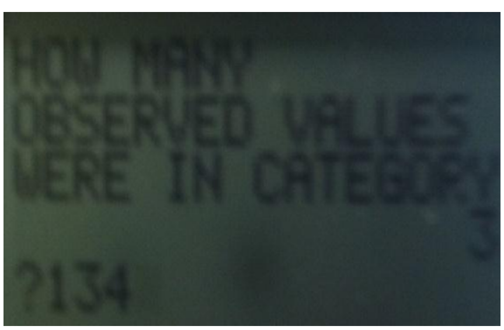
Using GOF program



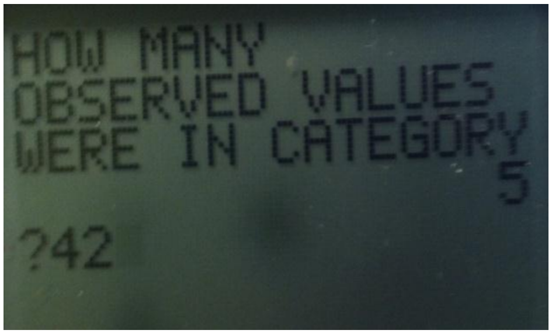
Isela Guerra
Instructor: L. Smith

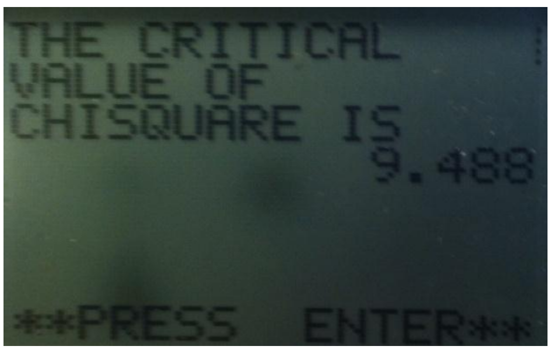


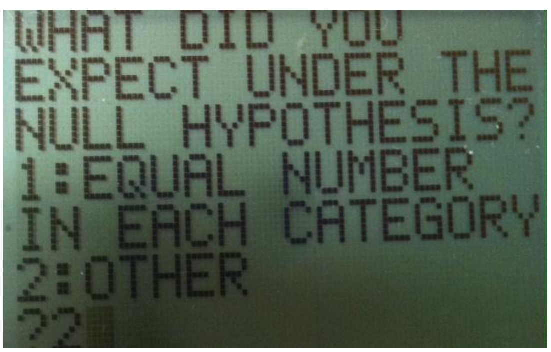


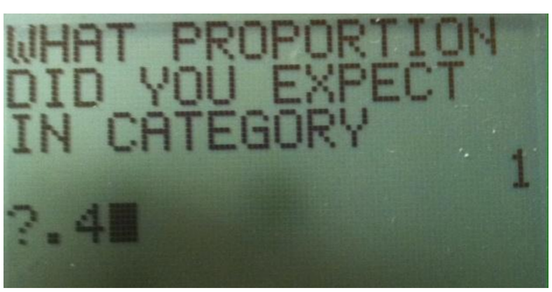


Isela Guerra
Instructor: L. Smith

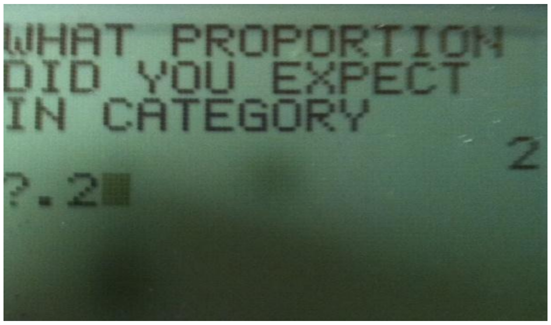


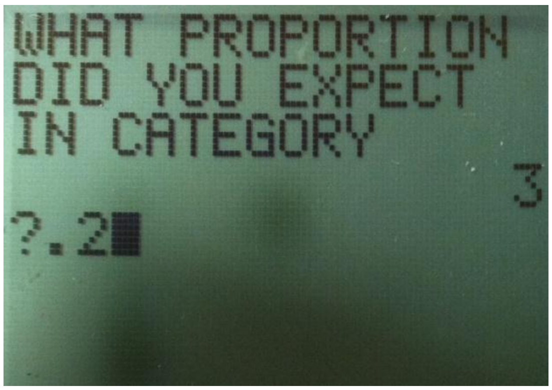


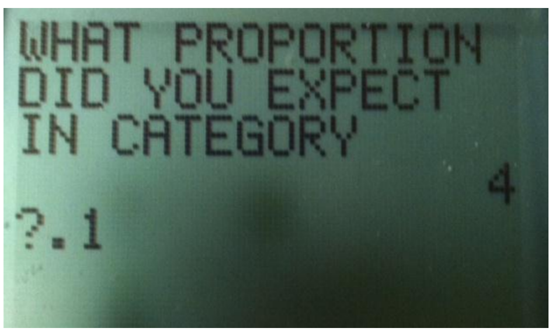


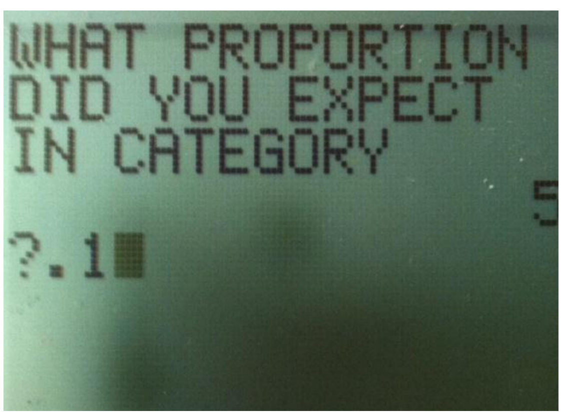


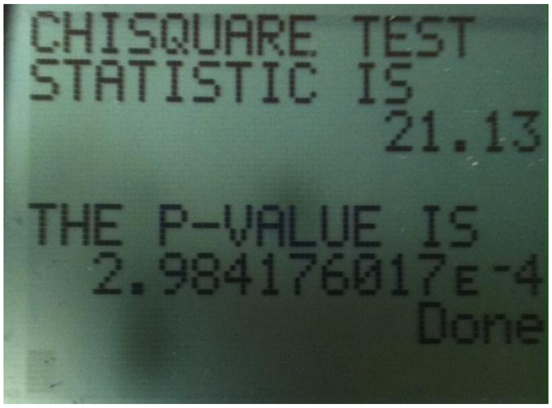
Isela Guerra
Instructor: L. Smith







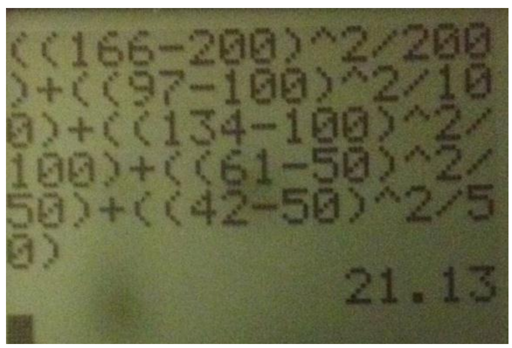




[To perform the **GOF**, we go down to programs and enter the information it asks for, however, keep in mind that when it asks for the null hypothesis we are not expecting them to be all equal, rather equal to the proportions given, so we enter 2 and then enter each individual proportion to its corresponding category. At the end of this program we are given a **p- value** of **2.98X10^{-4}**.]

Using Table

$$\chi^2 = \text{SUM}\frac{(\text{observed frequency - expected frequency})^2}{\text{expected frequency}}.$$



[To find the p-value using the table, first we must calculate the chi-square test statistic: X^2=SUM (observed frequency-expected frequency) ^{2}/expected frequency). We take the observed frequency and subtract it from the proportion we expect multiplied by the total number and square the quantity, we then divide it by the expected frequency. After calculating the summation we get a chi-square test statistic value of **21.13.** When we take a look at the table we will see that we have a **p-value less than .001** and will **not be displayed** on the table itself.]

Step 4: p-value ≤ α

Reject the null hypothesis.

Step 5: There is significant evidence at the .05 level of significance to conclude that the proportions are different.