Goodness of Fit

Discrete Random Variables

Does a sample set of data that consists of **observed frequency counts** arranged in a one-way frequency table (row or column) align with the **expected frequency counts** from the known distribution (uniform distribution, normal distribution, or any other distribution). To make that determination, we need to perform a **"Goodness of Fit Test"**.

Goodness of Fit Test- Is a Hypothesis Test that and observed frequency distribution fits (conforms) some claimed distribution.

A Goodness of Fit Test is also known as a Chi Square Test of a Distribution.

Consider the experiment that a bag on M and M's has an equal distribution of colors in every package. That is, the distribution of red, orange, yellow, green, blue, and brown is **uniform**. This is a vital requirement for the Goodness of Fit Test. **We need a random variable** *x* **of data and an assumed probability distribution**.

We require the following:

- Data is randomly selected.
- Sample data consists of frequency counts for each of the categories of data.
- Every category needs an expected frequency that is at least 5.

We need to Create the Following Hypothesis Test.

 $H_0: p_1 = p_2 = p_3 = p_4 = p_5 = p_6$ Uniform Distribution as our claim $H_1:$ At least one p_i is not equal.

 $\alpha = 5\%$

What we are saying is that the assumption in H_0 is that the frequency counts are the same while H_1 indicates that the frequency counts are not the same.

The following is our sample information as well as the expected frequencies. Recall, that $p_i = 1/6$ for all *i*.

M and M Candies		
Sample Size=900	Observed	Expected
Color	0	E
Red	134	150
Orange	152	150
Yellow	126	150
Green	168	150
Blue	162	150
Brown	158	150
Total	900	900

Test Statistic

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

Where **E** is the expected frequency of an outcome found by the assumption of the distribution.

If the expected frequencies are **all equal** as in a uniform distribution, then E=n/k

If the expected frequencies are **not all equal**, then $E = np_i$ for all *i* each group.

$$E = \frac{n}{k} = \frac{800}{6} = 150$$

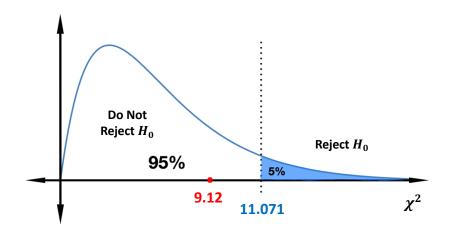
M and M Candies					
Sample Size=900	Observed	Expected			
Color	0	E	0-E	(O-E)^2	((O-E)^2)/E
Red	134	150	-16	256	1.7067
Orange	152	150	2	4	0.0267
Yellow	126	150	-24	576	3.8400
Green	168	150	18	324	2.1600
Blue	162	150	12	144	0.9600
Brown	158	150	8	64	0.4267
Total	900	900		1368	9.12

 $\chi^2 \approx 9.12$

TEXAS INSTRUMENTS TI-84 Plus CE NORMAL FLOAT AUTO REAL DEGREE MP	TEXAS INSTRUMENTS TI-84 Plus CE	TI-84 Plus CE NORMAL FLOAT AUTO REAL DEGREE MP
L1 L2 L3 L4 L5 2 134 150 152 150 150 150 150 168 150 168 150 158 150 158 150 158 150 158 150 158 150 158 150 158 150 158 150 158 150 158 150 150 158 150 150 150 158 150 <td>X2GOF-Test Observed:L1 Expected:L2 df:5 Color: BLUE Draw</td> <td>X²GOF-Test X²=9.12 P=0.1043724308 df=5 CNTRB={1.7066666667 0.026</td>	X2GOF-Test Observed:L1 Expected:L2 df:5 Color: BLUE Draw	X ² GOF-Test X ² =9.12 P=0.1043724308 df=5 CNTRB={1.7066666667 0.026
L2(?)= statplot f1 tblset f2 format f3 calc f4 table f5 y= window zoom trace graph	statplot f1 tblset f2 format f3 calc f4 table f5 y= window zoom trace graph	statplot f1 tblset f2 format f3 calc f4 table f5 y= window zoom trace graph

 $p \approx 0.104; p \ll \alpha$; Accept H_0 The Sample Supports the Claim The Distribution is Uniform

				Are	a to the F	light of the	e Critical V	alue		
Degrees of										
Freedom	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



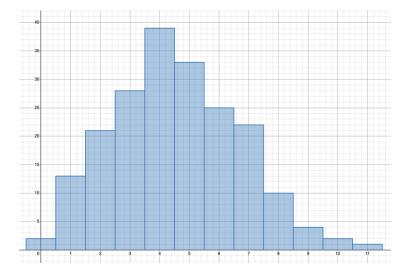
Conclusion Do Not Reject the Null The Distribution is Uniform Let us consider another experiment that consists of Poisson Probability Distribution. If we let x represent the number of 911 emergency calls in an hour where it is known that there is 4.6 calls per hour, the distribution will look as follows.

$$x = 0, 1, 2, 3, 4, \dots, 11$$

Let us now assume that the distribution is uniform and that we should see every outcome as equally likely.

$$p_i = rac{1}{12}$$
 for all i

However, when we randomly collect 200 data values (Poisson Distribution) we see the following histogram. We will conduct a **Goodness of Fit Test** and verify that the distribution is **not Uniform**.



 $H_0: p_1 = p_2 = p_3 = p_4 = p_5 = p_6$ Uniform Distribution as our claim $H_1:$ At least one p_i is not equal.

Using the sample information below we can proceed with $E = np_i$ for all *i* with n = 200.

$$\alpha = 5\%$$

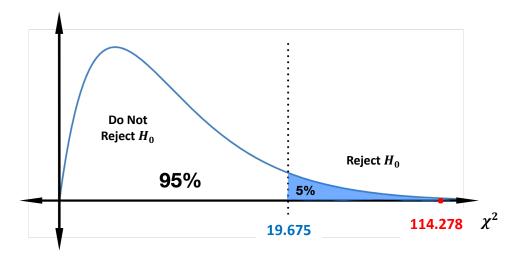
n=200	Observed	Expected			
x	0	E	0-E	(O-E)^2	(O-E)^2/E
0	2	16.667	-14.667	215.121	12.907
1	13	16.667	-3.667	13.447	0.807
2	21	16.667	4.333	18.775	1.126
3	28	16.667	11.333	128.437	7.706
4	39	16.667	22.333	498.763	29.925
5	33	16.667	16.333	266.767	16.006
6	25	16.667	8.333	69.439	4.166
7	22	16.667	5.333	28.441	1.706
8	10	16.667	-6.667	44.449	2.667
9	4	16.667	-12.667	160.453	9.627
10	2	16.667	-14.667	215.121	12.907
11	1	16.667	-15.667	245.455	14.727
Total	200	200.004		1904.667	114.278

 $\chi^2 \approx 114.278$

	as Instruments L Float auto ri		4 Plus CE	TEXAS INSTRUMENTS	TI-84 Plus CE	TEXAS INSTRUMENTS	TI-84 Plus CE
L1 21 28 39 33 25 22 10 4 2 1 1 	L2 L3 16.667 16.677 16.667 16.677 16.777 16.777 16.777 16.7777 17.7777 17.7777 17.77777 17.777777 17.7777777777	<u>L4</u>	<u>_5</u> 2	X2GOF-Test Observed:L1 Expected:L2 df:11 Color: BLUE Calculate Draw	3	X2GOF-T X2=114.2777145 P=2.55383278E df=11 CNTRB={12.9069	19
statplot			f4 table f5 graph	statplot f1 tblset f2 format f3 ca y= window zoom	alc f4 table f5 trace graph	statplot f1 tblset f2 format f y= window zoom	

 $p pprox 0.000; p < lpha; H_0$ is too low, has to go! The Sample Does Not Support the Claim The Distribution is not as Claimed

			Area to the Right of the Critical Value										
Degrees of													
Freedom	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			



Conclusion Reject the Null The Distribution is not as Claimed Let us revisit the Poisson Distribution of a mean 4.6 office hour arrivals in an hour. We will use a **Goodness to Fit Test** with the **assumption that the distribution is a Poisson Distribution**. Let's see how Goodness to Fit addresses works in the case where each group has assigned probabilities using the **5% level of significance**.

x	p(x)			
0	0.010			
1	0.046			
2	0.106			
3	0.163			
4	0.188			
5	0.173			
6	0.132			
7	0.087			
8	0.050			
9	0.026			
10	0.012			
11	0.005			
12	0.002			
Total	0.999			

 $x = 0, 1, 2, 3, 4, \dots, 11$

 $\begin{array}{l} H_0: p_1 = 0.\,010, \ p_2 = 0.\,046, \ p_3 = 0.\,163, \ p_4 = 0.\,188, \ p_5 = 0.\,173, \ p_6 = 0.\,132, p_7 = 0.\,087, p_8 = 0.\,050, p_9 = 0.\,026, p_{10} = 0.\,012, p_{11} = 0.\,005, p_{12} = 0.\,002 \ \ \mbox{Claim} \end{array}$

 H_1 : At least one p_i is not equal to what is in the claim.

Using the original sample information below we can proceed with $E = np_i$ for all *i* with n = 200.

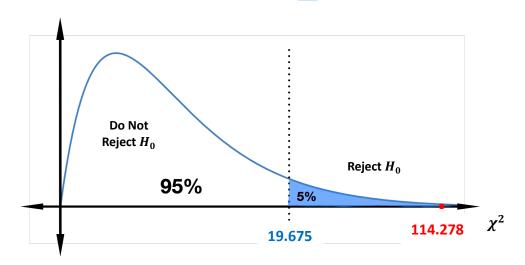
n=200	Observed	Expected			
х	0	E	0-E	(O-E)^2	(O-E)^2/E
0	2	2.010	-0.010	0.000	0.000
1	13	9.248	3.752	14.080	1.523
2	21	21.270	-0.270	0.073	0.003
3	28	32.614	-4.614	21.285	0.653
4	39	37.506	1.494	2.233	0.060
5	33	34.505	-1.505	2.265	0.066
6	25	26.454	-1.454	2.114	0.080
7	22	17.384	4.616	21.307	1.226
8	10	9.996	0.004	0.000	0.000
9	4	5.109	-1.109	1.230	0.241
10	2	2.350	-0.350	0.123	0.052
11	1	0.983	0.017	0.000	0.000
Total	200	199.4274733		64.710	3.903

$$\chi^2 \approx 3.903$$

· · · · ·	AS INSTRUMENTS L FLOAT AUTO RE	 Plus CE	TEXAS INSTRUMENTS TI-84 Plus CE NORMAL FLOAT AUTO REAL DEGREE MP	TEXAS INSTRUMENTS TI-84 Plus CE
L1 28 39 33 25 22 10 4 2 1 	L2 L3 21.27 32.614 37.506 34.505 26.454 17.384 9.996 5.109 2.35 0.983	2	X2GOF-Test Observed:L1 Expected:L2 df:11 Color: BLUE Draw	X ² GOF-Test X ² =3.90237153 p=0.9726862592 df=11 CNTRB={4.975124378E-5 1
statplot y=		table f5 graph	statplot f1 tblset f2 format f3 calc f4 table f5 y= window zoom trace graph	statplot f1 tblset f2 format f3 calc f4 table f5 y= window zoom trace graph

 $p \approx 0.973$; $p \ll \alpha$; Accept H_0 The Sample Supports the Claim The Distribution is as Claimed

				Are	a to the H	Right of the	e Critical V	alue		
Degrees of										
Freedom	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



Conclusion Reject The Null The Distribution is not as Claimed.

Police Calls

St. Vegas released the number of Police Calls for every day of the week in the month of October for 31 days. Monday (128), Tuesday (158), Wednesday (172), Thursday (176), Friday (382), Saturday (398), Sunday (348). Use the 1% level of significance to test the claim that the distribution is not Uniform.

 H_0 : $p_1 = 0.1$, $p_2 = 0.1$, $p_3 = 0.1$, $p_4 = 0.1$, $p_5 = 0.2$, $p_6 = 0.2$, $p_7 = 0.2$ Claim

 H_1 : At least one p_i is not equal to what is in the claim.

Using the original sample information below we can proceed with $E = np_i$ for all *i* with n = 1762

х	p(x)	0	E	О-Е	(O-E)^2	(O-E)^2/E
Monday	0.1	128	176.2	-48.2	2323.24	13.185
Tuesday	0.1	158	176.2	-18.2	331.24	1.880
Wednesday	0.1	172	176.2	-4.2	17.64	0.100
Thursday	0.1	176	176.2	-0.2	0.04	0.000
Friday	0.2	382	352.4	29.6	876.16	2.486
Saturday	0.2	398	352.4	45.6	2079.36	5.901
Sunday	0.2	348	352.4	-4.4	19.36	0.055
Total	1	1762				23.607
		Sample Size				

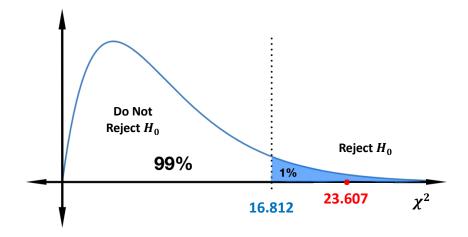
 $\chi^2 \approx 23.607$

•	as Instrumi L FLOAT A			84 Plus C	E	TEXAS INSTRUMENTS T	TI-84 Plus CE REE MP	TEXAS INSTRUMENTS TI-84 Plus C
L1 128 158 172 176 382 398 348 	L2 176.2 176.2 176.2 352.4 352.4 352.4	L3	L4 	Ls	2	X2GOF-Test Observed:L1 Expected:L2 df:6 Color: BLUE Draw		X2GOF-Test X2=23.60726447 p=6.166307982E-4 df=6 CNTRB={13.18524404 1.879
L2(8)= statplot 1 y=		2 format		f4 table e grapi		statplot f1 tblset f2 format f3 cal y= window zoom t	ic f4 table f5 race graph	statplot f1 tblset f2 format f3 calc f4 table y= window zoom trace grap

 $p \approx 0.000; p < \alpha; p$ value is low, H_0 has to go! The Sample Does Not Support the Claim The Distribution is not as Claimed

lpha=1% and df=6

	Area to the Right of the Critical Value												
Degrees of													
Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.00			
1	_	_	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.8			
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.5			
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.8			
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.8			
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.7			
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.5			
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.2			
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.9			
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.5			
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.1			
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.7			
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.2			
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.8			
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.3			
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.8			
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.2			
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.7			
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.1			
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.5			
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.9			
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.4			
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.7			
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.1			
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.5			
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.9			
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.2			
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.6			
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.9			
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.3			
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.0			
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.7			
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.4			
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.9			
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.2			
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.3			
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.2			



Conclusion Reject the Null The Distribution is not as Claimed.

World Series Games

The table below illustrates the number of games played in 105 Major League Baseball World Series. The table also includes the expected proportions for the number of games played in the MLB World Series, assuming that in each series each team has the same chance of winning. Use the **1% level of significance** to test the claim that the actual number of games fit the distribution indicated by the expected proportions.

Games Played	World Series Contests	Expected Proportion
4	21	0.125
5	23	0.25
6	23	0.3125
7	38	0.3125
Total	105	1

H_0 : $p_1 = 0.125$, $p_2 = 0.25$, $p_3 = 0.3125$, $p_4 = 0.3125$ Claim

H_1 : At least one p_i is not equal to what is in the claim.

Using the original sample information below we can proceed with $E = np_i$ for all i with n = 105

Games Played	iames Played O E		E	0-E	(O-E)^2	(O-E)^2/E	
4	21	0.125	13.125	7.875	62.0156	4.7250	
5	23	0.25	26.25	-3.25	10.5625	0.4024	
6	23	0.3125	32.8125	-9.8125	96.2852	2.9344	
7	38	0.3125	32.8125	5.1875	26.9102	0.8201	
Total	105	1				8.8819	

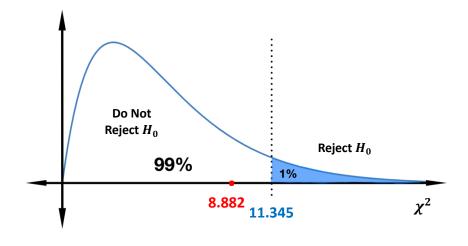
 $\chi^2 \approx 8.882$

TEXAS INSTRUMENTS TI-84 Plus CE NORMAL FLOAT AUTO REAL DEGREE MP	TEXAS INSTRUMENTS TI-84 Plus CE NORMAL FLOAT AUTO REAL DEGREE MP	TEXAS INSTRUMENTS TI-84 Plus CE NORMAL FLOAT AUTO REAL DEGREE MP
L1 L2 L3 L4 L5 2 21 13.125 23 26.25 23 32.813 38 32.813	X2GOF-TestObserved:L1Expected:L2df:3Color:BLUECalculate Draw	X2GOF-Test X2=8.881904762 p=0.0309029398 df=3 CNTRB={4.725 0.402380952
L2(5)= statolot f1 tblset f2 format f3 calc f4 table f5		
statplot f1 tblset f2 format f3 calc f4 table f5 y= window zoom trace graph	statplot f1 tblset f2 format f3 caic f4 table f5 y= window zoom trace graph	statplotf1tblsetf2formatf3calcf4tablef5y=windowzoomtracegraph

 $p \approx 0.031$; $p \ll \alpha$; Accept H_0 The Sample Supports the Claim The Distribution is as Claimed

lpha=1% and df=3

	Area to the Right of the Critical Value											
Degrees of												
Freedom	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	18.54		
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	34.26		
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		



Conclusion Do Not Reject the Null **The Distribution is as Claimed.**

IQ Scores

We know IQ Scores are Normally Distributed ($\mu = 100$; $\sigma = 15$) and the "Goodness of Fit Test" applies to discrete random variables. The Normal Distribution is a continuous random variable that can be made to approximate a discrete random variable when we look at the outcomes over mutually exclusive (non-intersecting) intervals. Consider the following example where we shall consider the 1% level of significance to Test the Claim IQ Scores fit the distribution by expected proportions.

	Expected Proportion
x	p(x)
Less than 70	0.023
Between 70 and 79.99	0.068
Between 80 and 89.99	0.161
Between 90 and 99.99	0.247
Between 100 and 109.99	0.247
Between 110 and 119.9	0.161
Between 120 and 129.9	0.068
At least than 130	0.023
Total	0.998

 $\begin{array}{l} H_0: p_1=0.\,023, p_2=0.\,068, \ p_3=0.\,161, \ p_4=0.\,247\\ p_5=0.\,247, \ p_6=0.\,161, \ p_7=0.\,068, \ p_8=0.\,023 \ \ \mbox{Claim} \end{array}$

 H_1 : At least one p_i is not equal to what is in the claim.

Using the original sample information below we can proceed with $E = np_i$ for all *i* with n = 60

	Expected Proportion					
х	p(x)	0	E	O-E	(O-E)^2	(O-E)^2/E
Less than 70	0.023	3	1.380	1.620	2.624	1.902
Between 70 and 79.99	0.068	4	4.080	-0.080	0.006	0.002
Between 80 and 89.99	0.161	7	9.660	-2.660	7.076	0.732
Between 90 and 99.99	0.247	12	14.820	-2.820	7.952	0.537
Between 100 and 109.99	0.247	14	14.820	-0.820	0.672	0.045
Between 110 and 119.9	0.161	10	9.660	0.340	0.116	0.012
Between 120 and 129.9	0.068	6	4.080	1.920	3.686	0.904
At least than 130	0.023	4	1.380	2.620	6.864	4.974
Total	0.998	60	60			9.107

 $\chi^2 \approx 9.107$

· ·	as Instrum L FLOAT f			84 Plus MP	CE	TEXAS INSTRUMENTS	TI-84 Plus CE	TEXAS INSTRUMENTS	TI-84 Plus CE
L1 3 4 7 12 14 10 6 4 	L2 1.38 4.08 9.66 14.82 14.82 9.66 4.08 1.38	L3	L4 	L5	2	X2GOF Observed:L1 Expected:L2 df:7 Color: <u>BLUE</u> Dra		X2GOF=T X2=9.10744102 P=0.2450340227 df=7 CNTRB={1.90173	
L2(9)= statplot y=		f2 format v zoom		f4 table e graj		statplot f1 tblset f2 formal y= window zoo		statplot f1 tbiset f2 format f y= window zoom	

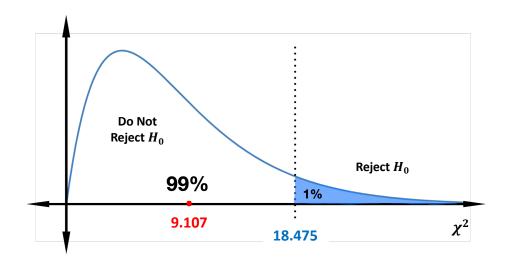
 $p \approx 0.245$; $p \ll \alpha$; Accept H_0 The Sample Supports the Claim The Distribution is as Claimed

Critical Values

Chi-Square Distribution where df = k - 1k is the number of categories Always a **Right Tail Test**

lpha=1% and df=7

				Are	a to the R	ight of the	e Critical V	alue		
Degrees of										
Freedom	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	20.27
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	29.81
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	32.80
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	41.40
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	49.64
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	52.33
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	66.760
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



Conclusion Do Not Reject the Null The Distribution is as Claimed