Hypothesis testing: Goodness of Fit

Chi-squared Distribution (X^2)

The last type of hypothesis test we will see is a **goodness of fit** test. This is a test where the hypothesis involves the distribution of the data—whether the data "fits" within the "expected" result. The null hypothesis H_0 is that there is <u>no</u> significant difference between our observed distribution and the expected result; the alternative hypothesis H_1 is that there <u>is</u> a significant difference.

Just as we used a **Z-test** to test a population mean where the standard deviation is known, a **T-test** for a population mean where the population standard deviation is <u>unknown</u>, or a **1-Prop-Z-Test** for a population proportion, we similarly use a **Chi-Squared (X^2) test a distribution**.

This test uses the X^2 distribution, which like the T-Distribution is different depending on how big your sample is and how many degrees of freedom your sample has. The **degree of freedom (df)** is determined by **n – 1 where n is the # of categories**. The idea is that if you know how many categories there are, the size of your sample, and the sample frequency of all but one of the categories, you know the frequency of the last category so the last one has no "freedom."

We use our observations and our expectations, along with the df, to determine p, the probability of getting our observed result. Our probability will get smaller the more our observations "O" differ from our expectations "E", so we begin by calculating how these differ using the formula $X^2 = \sum \frac{(O-E)^2}{E}$. This allows us to find (on a table or TI84 calculator) the probability p.

Here are the steps:

NOTE: Follow the steps but choose <u>either</u> OPTION 1 (for a TI84) OR OPTION 2.
1. Determine H_0 and H_1, the null and alternative hypotheses.
2. Calculate the expected probabilities (for H_0).
3. Determine degree of freedom (df): $n – 1$ where n is the # of categories**.**

OPTION 1: Run a X^2 goodness of fit (GOF) test on a <u>TI-84</u> calculator:
4. Enter the observed probability distribution into L1, and the expected distribution in L2.
5. Run a X^2 GOF-Test from STAT->TESTS->D:X^2GOF-Test.., press ENTER to choose options and obtain the probability p.

                              OR

OPTION 2: Find X^2 and use a table to find p.

4. Calculate $X^2 = \sum \frac{(O-E)^2}{E}$ by hand or enter **o**bserved in L1, **e**xpected in L2, and in L3: (L1-L2)²/L2. Choose 2nd MODE (to QUIT) and then 2nd STAT->MATH->5 to sum(L3). This is X^2.
5. Use the X^2 table. Locate the df row and find the critical values in that row between which your X^2 lies. This gives us the upper and lower limits of the p-values.

Compare
6. Finally, compare the probability p to *alpha.* If $p > alpha$, do not reject H_0. If $p < alpha$, evidence of unlikeliness of result is sufficient to reject H_0.

Example:

A die is being tested to see if it's a fair die; that is, that all six sides will turn up with equal probability. We roll the die 400 times and arrive at the following observed distribution of results to compare them to expectations:
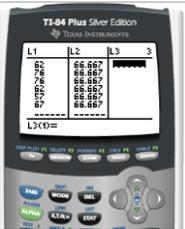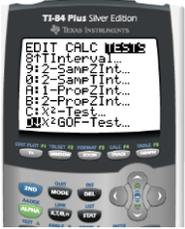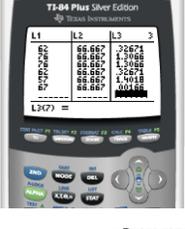
Observations	Chance (Expected Results)
Rolled 1: 62 times **Rolled 2: 76 times** **Rolled 3: 76 times** **Rolled 4: 62 times** **Rolled 5: 57 times** **Rolled 6: 67 times**	If all six sides turn up with equal probability, a die rolled 400 times should be equally divided between the possibilities. This means whether we rolled 1, rolled 2, rolled 3, etc, we would **expect**: $\frac{400\ rolls}{6} = 66.667 \approx 67\ times\ rolled\ for\ each$ *Note that this happened ONLY when we rolled 6.*

<u>Question:</u> Clearly the die did not do as expected but is the die fair at the *alpha* = 0.01 level of significance?

H_0: All numbers are equally likely…. the die is fair.

H_1: All numbers are not equally likely…the die has been altered so that select numbers occur more often.

Solution:

OPTION 1: TI Goodness of Fit test	OPTION 2: Calculate X^2 and use the table
Observed: **L1** Expected: **L2** df: 5 *There are 6 categories, so 6-1=5 degrees of freedom* Rounded results: $X^2 = 4.67$, *p*= 0.4575	Rounded result: $X^2 =$ Sum(L3) = 4.67 *Video Example*

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2							
	0.99	0.95	0.90	0.75	0.50 p	0.25	0.10	0.05
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07

X2 = 4.67

p is near but less than 0.50

Whether we found that the probability that fair die would show the observed results is a fairly high p=0.4575 (about 46%), used a table to find that p is near 0.50 (about 50%), or used a less accurate table and know only that p is between 0.90 and 0.10 (somewhere between 90% and 10%), one thing is clear—the probability p of getting the observed result from fair die is much larger than our level of significance *alpha*=0.01 (1% or less). Even though our die did not always give us the outcomes we expected (67), our p value shows us that our sample was not unusual enough to reject the idea that the die was fair--**there is not sufficient evidence to reject H_0**.

RESOURCES

Mr. Bryce's *X^2* video TI-84 calculator example:
http://www.youtube.com/watch?v=3NBt9-VTjaY&feature=youtu.be

Need to do the calculation by hand and use a table? Look at these examples:
https://www.youtube.com/watch?v=s1xEB9XknBI and https://www.youtube.com/watch?v=VskmMgXmkMQ