Chi-squared

Confidence Intervals with Standard Deviation and Variance

The NEED to CONTROL variation

(in production- good example)

The Chi-Square Distribution

Of course we again start with a POINT ESTIMATE

- The point estimate for POPULATION VARIANCE (σ^2) is sample variance (s^2) and for POPULATION STANDARD DEV (σ) is sample standard deviation (s)

Use a Chi-Squared distribution to find confidence intervals for variance and St. Dev. (for samples n>1)

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

notation for Chi Squared

There are TWO critical values of each level of confidence (since skewed). We have

for left-tail crit value (the meaty part)

for right tail crit value (the tail part)

Finding the critical values:

df = 19

EX. find the left and right critical values for a 90% confidence interval with sample size 20.

$$X_L^2 = \frac{1+c}{2} = \frac{1+.9}{2} = .95 \xrightarrow{\text{chart}} 10.117$$

$$X_R^2 = \frac{1-c}{2} = \frac{1-.9}{2} = .05 \rightarrow 30.144$$

Now confidence intervals:

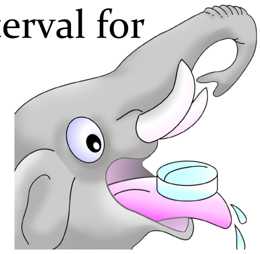
For population variance: $\frac{(n-1)s^2}{X_R^2} < \sigma^2 < \frac{(n-1)s^2}{X_L^2}$

For population standard deviation: $\sqrt{\frac{(n-1)s^2}{X_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{X_L^2}}$

Well holy moley, that looks complicated...........lets see how this plays out with an example

EX.

You randomly select and weigh 30 samples of allergy medicine. The sample standard deviation is 1.20mg. Assuming the weights are normally distributed, construct a 99% confidence interval for population variance and standard deviation.



$X_L^2 = \frac{1+.99}{2} = .995 \rightarrow 13.121$

chart

$\chi_R^2 = \frac{1-.99}{2} = .005 \rightarrow 52.336$

df = 29

$$\frac{(29)(1.2)^2}{52.336} < \sigma^2 < \frac{(29)(1.2)^2}{13.121}$$

$$.7979 < \sigma^2 < 3.18$$

$$.893 < \sigma < 1.78$$

EX.

Water Quality:

As part of a water quality survey, you test the water hardness in 19 randomly selected streams. The sample standard deviation is 15 grains/gallon. Find a 95% confidence interval for the population variance and population standard deviation.

Chi-Squared and Goodness of Fit Test

A High Chi Squared is evidence to Reject Ho

Categorical Data

Conditions for using Chi-Sq:

- Random
- Independent-(check 10% rule: sample<10%pop)
- Expected values at least 5 (this takes the place of "normality" in the previous testing conditions)- LARGE SAMPLE condition- check expected not observed counts

Formula for Goodness of Fit Test:

Tests statistic:

$$\text{Chi-square test statistic} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

X^2 Notation

df=n-1

Writing Ho and Ha's:

Since you can use this to cover multiple variables at the same time, make your alternative (Ha) reflect the failure of just one.

EX: (M&Ms)

H_o: p_{blue}=.24 p_{red}=.13 p_{green}=.16

H_a: at least one of the p's is incorrect.

Ex:

Mars reports that their peanut M&Ms havethe following color distribution:

23% each of blue and orange

15% each of green and yellow

12% each of red and brown

Joey bought a bag and counted each color and got the following counts:

12 blue, 7 orange, 13 green, 4 yellow, 8 red, 2 brown $= 46$

A) State appropriate hypothesis for testing the companies claim about the color distribution of peanut M&Ms

$H_0: p_{BL} = .23, \; p_O = .23 \;\; p_G = .15 \;\; p_Y = .15 \;\; p_R = .12 \;\; p_B = .12$

H_a: at least one color proportion is incorrect

B) Calculate the expected counts for each color assuming the companies claim is correct.

$BL = 46(.23) = 10.58$

$Or = 46(.23) = 10.58$

$gr = 46(.15) = 6.9$

$Y = 46(.15) = 6.9$

$R \; (46)(.12) = 5.52$

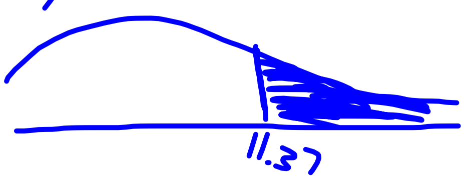
$B \; (46)(.12) = 5.52$

C) Calculate the chi-squared statistic for Joeys sample.

$$\chi^2 = \sum \frac{(Ob - Ex)^2}{Ex} = \frac{(12-10.58)^2}{10.58} + \frac{(7-10.58)^2}{10.58} + \frac{(13-6.9)^2}{6.9} + \frac{(4-6.9)^2}{6.9} + \frac{(8-5.52)^2}{5.52} + \frac{(2-5.52)^2}{5.52} = 11.37_{24}$$

D) Find the P-Value and make a decision about the companies claim.

$\chi^2cdf = (11.3724, \; E99 / 20000, \; 5)$



$P = .0448$

EX.

Are births evenly distributed across the days of the week. The one way table below shows the distribution of births across the days of the week in a random sample of 140 births from local records in a large city.

$140/7 = 20 = \text{Exp V.}$

Days	Sun	Mon	Tues	Wed	Thurs	Fri	Sat
Births	13	23	24	20	27	18	15

Do the data give significant evidence that local births are not equally likely on all days of the week. χ^2 GOF

Conditions: Random → 140 random births
Indep → births are independ
→ 140 < 1400 all births
Large #: Exp V · 20 > 5 ✓

State Hypotheses
$H_0: P_M = P_T = P_W = P_{TH} = P_F = P_S = P_S = 1/7$
H_a: births do not happen equally on days of the week ($p = 1/7$)

Chi Squared Goodness of fit Test. Calc X^2 and then find the P-Value using the calculator (Distr #8 χ^2cdf). State your conclusion

$$\chi^2 = \sum \frac{(Ob - Exp)^2}{Exp} = \frac{(13-20)^2}{20} + \frac{(23-20)^2}{20} + \frac{(24-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(27-20)^2}{20} + \frac{(18-20)^2}{20} + \frac{(15-20)^2}{20} = 7.6$$

$\chi^2\text{cdf}(7.6, \infty, 6) = .2689$

Another Calc Function is under TESTS #D X^2-GOF test- uses data from lists. (put observed values L1 and expected values L2) Lets do it with the Births example

$\chi^2 = 7.6$

EX.

In his book Outliers, Malcomb Gladwell suggests that hockey players birth month has a big influence on his chance to make it to the highest levels of the game. Specifically since Jan 1 is the cutoff date for youth leagues in Canada (where most NHL players come from), players born in January will be competing against kids who are 12 months younger. The older players tend to be bigger, stronger, and more coordinated and hence get more playing time, coaching and are more coordinated and have a better chance of being successful. To see if birth dates are related to success, a random sample of 80 NHL players were selected and their birthdays recorded. Overall, 32 were born in the first quarter, 20 in the second quarter, 16 in the third and 12 in the fourth quarter. Do these data provide convincing evidence that the birthdays of NHL players are not uniformly distributed?(use alpha of 0.05)

Section 14.2:

Inference for relationships/two way tables

Chi squared test for Homogeneity

Two way tables

Formula and Calc functions

Use when comparing two distributions from two samples

$$\boxed{Expected_{ij} = \frac{Row\ Total_i \times Col\ Total_j}{Grand\ Total}}$$

Expected counts are NOT rounded

$$\text{Chi-square test statistic} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

df=(row-1)(col -1)

Calc:

matrix input (2nd x^{-1})

tests: X^2-test (#c)

EX. Calculating expecting counts-formula

AN article in the Journal of the American Medical Association report the results of a study designed to see if the herb St. Johns Wart is effective at treating moderately severe cases of depression. There were 338 subjects being treated for major depression. The subjects were randomly assigned to recieve one of 3 treatments for an eight week period. Here are the results of an experiment comparing the effects of St. Johns Wart, Zoloft, and a placebo

	St Johns Wart	Zoloft	Placebo	Totals
Full response	27 30.42	27 29.35	37 31.2	91
Partial Response	16 18.39	26 17.7	13 18.9	55
No Response	70 64.2	56 61.9	66 65.9	192
totals	113	109	116	338

A) Calculate the $\boxed{\text{expected counts}}$ for the three treatments assuming all three are equally effective.

$\frac{(91)(113)}{338} = 30.42$ $\frac{(91)(109)}{338} = 29.35$ $\frac{(91)(116)}{338} = 31.2$

$\frac{(55)(113)}{338} = 18.39$ $\frac{(55)(109)}{338} = 17.7$

L1=obs

L2=Exp

B) Calculate the chi squared test statistic

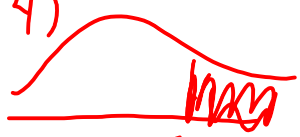
$x^2 = \sum \frac{(obs - Exp)^2}{Exp} = \frac{(27-30.42)^2}{30.42} +$ $\boxed{L3}$ $(L1-L2)^2/L2$

1 Var Stat

$\sum x$ 8.72

$\chi^2cdf(8.72, \infty, 4)$

.0685



8.7

C) Write the appropriate hypotheses.

Ho: No diff in the distributions for treatments for depression using St. J. wart, Zoloft, placebo

Ha: There IS a difference in the distributions for treatments using St. Jw, Zoloft or placebo

D) Verify the conditions

→ Random → random assignment of trmt.

→ ind. → results will be ind. per indiv.

338 < 10% all people suff major dep

→ large #'s → all exp counts ≥ 5

E) Interpret the P-value and make a decision

$\alpha = .05$

$p = .0685 \not< .05 \Rightarrow$ FTR H_0

EX: SuperPowers Revisited.(calc ex)

EX: Super Powers Revisited.(calc ex)
Separate random samples form the UK and the US who completed a survey in a survey in a recent year were selected. For each student we recorded the superpower he or she would most like to have. The results of the sampling is in the chart.
Is there convincing evidence that the distributions of superpower preference are different for survey takers in the two countries.

	UK	US	Total
Fly	54	45	99
Freeze Time	52	44	96
Invisibility	30	37	67
Super Strength	20	23	43
Telepathy	44	66	110
Total	200	215	415

State the appropriate hypotheses
Verify the conditions (Random, expected counts(5), 10%sampling)
Calculate the chi squared test statistic (x²-cdf)
Conclusions

Ho: There is no difference in distribution of super power choices of children from UK + US

Ha: There is a difference in the distributions of super power choices in kids from UK + US

Conditions
random samples
indep.
Large #'s → Exp. Val ≥ 5 ✓

$\chi^2 = 6.29$

$p = .1784$

with a high Pvalue of .1784 we fail to reject the Ho (no diff) and thus can not support there is a difference in the distributions of ~~supper~~ super power choices btwn kids in UK + US.

Ex. Health Care

Canada has universal health care. The US does not but often offers more elaborate treatment to patients who have access. How do the two systems compare in treating heart attacks? Researchers compared random samples of US and Canadian heart attack patients. One key outcome was the patients own assessment of their quality of life relative to what it had been before the heart attack. Here are the data for the patients who survived the year. Is there a significant difference between distributions of quality of life ratings? Use a significance level of 0.01 and complete an appropriate test.

Quality of Life	Canada	US
Much Better	75	541
Somewhat Better	71	498
About the Same	96	779
Somewhat Worse	50	282
Much Worse	19	65
TOTAL		

Chi Squared Test for Independence

Another common situation that would use a two way table is when a SINGLE random sample of individuals is chosed from a SINGLE population and then classified based on TWO categorical variables. Our goal would be to analyze the relationship between the variables.

For Example: Are people who are prone to sudden anger more likely to develop heart disease.

Setting up Ho and Ha (referring to above)

The null is the "no association" and the alternative is "there is an association"

No association between two variables means that knowing the value of one variable does not help up predict the ~~calue~~ value of the other. That is, the variables are INDEPENDENT. There are two ways to set this up:

Ho: There is no association between anger levels and heart disease status in the population of people with normal blood pressure.

Ha: There is an association between anger levels and heart disease status in the population of people with normal blood pressure.

OR

Ho: Anger and heart disease status are independent in the population of people with normal blood pressure.

Ha: Anger and heart disease status are not independent in the population of people with normal blood pressure.

(notice I did not say dependent)

Conditions:

- Random: sample is random from population of interest (includes the 10% rule)
- Large Counts: expected counts are at least 5

Example:

(anger/heart disease)

An observational study followed a random sample of 8474 people with normal blood pressure for about four years. All of the individuals were free of heart-disease at the beginning of the study. Each person took the Spielburg Trait Anger Scale test, which measures how prone a person is to sudden anger. Researchers also recorded whether each individual developed coronary heart disease (CHD)(this includes people who had heart attacks. Here is the summary data (use alpha=.05):

		Anger Level			
		Low	Medium	High	Total
CHD	Yes	53	110	27	190
	No	3057	4621	606	8284
`	Total	3110	4731	633	8474

Lets determine the EXPECTED counts and put them in a table. (remember don't round to whole number) (if the Ho is right, what would we expect)

		Anger Level			
		Low	Medium	High	Total
CHD	Yes	69.73	106.08	14.19	190
	No	3040.27	4624.92	618.81	8284
`	Total	3110	4731	633	8474

$\frac{(190)(3110)}{8474}$

Calc: Low anger in the study and yes CHC: 190/8474 times 3110

Continue with rest of calcs.

$(2-1)(3-1) = 1 \cdot 2 = 2$

Calculate the test statistic X² Using formula → df=(row-1)(col -1)

$\frac{(53-69.73)^2}{69.73} + \frac{(110-106.08)^2}{106.08}$

$$\text{Chi-square test statistic} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

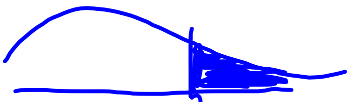
What are your degrees of freedom?

(use this to read the Chi Squared chart and get your critical value)

$\chi^2 = 16.08$

Is your test statistic in your rejection region or fail to reject region?

$\alpha = .05$



$\chi^2 = 5.99$

What can you conclude?

Ex. Snow Mobiles and Yellowstone

A random sample of 1526 visitors to Yellowstone National Park were asked two questions:

1. Do you belong to an environmental club?

2. What is your experience with a snow mobile: own, rent, never use?

The results are below. Do the data provide convincing evidence of an association between environmental club status and type of snow mobile use in the population of winter visitors to Yellowstone?

Environmental Club Status

		Not Member	Member	Total
SnM Exp	Never Used	445	212	657
	Rented	497	77	574
	Owned	279	16	295
	Total	1221	305	1526

H_0: Environmental club status is independ of snow mobile experience.

H_a: Env. Club Status is NOT indep of snow mob Exp.

χ^2-Test for Independance

$\chi^2 = 116.59 \quad df = 2$

$p = 4.82 \times 10^{-26} \approx 0$

With p value approx 0, there is suffic. evidence to Reject the H_0 (indep) and support the Ha that Environmental club status is not independ. of Snow Mobile experience

Choosing the right Chi- squared test

<u>Ex: Scary movies and fear</u>

Are men and woman equally likely to suffer the lingering fear from watching scary movies as children? Researches asked a random sample of 117 college students to write narrative accounts of their exposure to scary movies before the age of 13. More than one fourth of the students said that the fright symptoms are still present when they are awake. The table has the results. Determine the type of Chi-squared test you should use and then carry out the test.

		Gender		
		Male	Female	Total
Freight	Yes	7	29	36
Symptoms	No	31	50	81
Total		38	79	117

OTHER

- What if we wanted to compare two proportions?.....

  calculating a *2-PropZtest* = X^2*-test* for P value

- What is cell counts are less than 5? be clever- combine two rows (rename the row)

AP Notes:

While you can use the calculator on the AP to do the calculation and tests, it is still important to note WHAT test you are doing and to jot down the values you are inputting. If you just put the numbers down, you may risk getting no credit particulatly if there is an error....at least showing your input, the scorer can see where you may have gone wrong and give some credit.

AP Notes:

While you can use the calculator on the AP to do the calculation and tests, it is still important to note WHAT test you are doing and to jot down the values you are inputting. If you just put the numbers down, you may risk getting no credit particulatly if there is an error....at least showing your input, the scorer can see where you may have gone wrong and give some credit.

Difference between Homogeneity and Independence

The best explanation for this is that a ***test for homogeneity*** *is for when you are* ***comparing 2 or more samples*** *to see if they have the same distribution, while* ***independence*** *is when you are testing for independence* ***between 2 variables contained within a single sample.***

*This distinction is not very rigorous, but it's a fairly negligible which you choose to do, because the math is exactly the same. If you're taking something like the AP exam, it matters (in this case, you should be able to discern which one to use from the question wording.) In real life it doesn't make a difference. (*dang)*

Chi Squared goodness of fit *is specific to verifying a specific quantity (proportion or mean-expected value) in one sample. (M and M sample)*