CHI-SQUARE GOODNESS-OF-FIT, PART ONE

0:00 / 9:18 1.0x

SPEAKER: MICHAEL J. MAHOMETA, Ph.D.

When it comes to Hypothesis testing for categorical variables

- and only categorical variables - we have a few choices

depending upon our data and our question of interest.

If we have only ONE categorical variable,

we can ask a simple question about the distribution of responses

Help

- say "do you prefer red, black, white, or some other color car?"

Now to begin answering these categorical response questions we'll need to learn

a new statistical test - the chi-square test - and a new distribution -

the chi-square distribution.

And just for good measure, let's also start

by asking a more interesting question than about car color.

Now it's a perfectly reasonable question - although slightly disgusting -

and disturbing when you think about disease transmission.

What does a person do when they sneeze or cough?

Do they use a tissue?

Do they use their arm or elbow?

Do they use their hands?

Or, unfortunately, do they use nothing at all?

Well researchers at a southeastern university

measured these events across several classrooms for students.

They used teaching assistants to record the person and the event.

And they had a question - was there a

Help

pattern

to the behavior that was not random?

Now here's the data in a classic representation that we've seen before

- a Table of Counts.

We can see using our table, and the corresponding percentages,

that the pattern is indeed not random.

Had the behaviors BEEN random, then the percentages

should all be the same across the four types of behaviors

- we would EXPECT 25% of the behavior counts to be in each cell.

25% because we would equally distribute 100%

among the 4 cells of behavior possibilities.

But, is there a way that we can formally TEST this result?

Can we determine if this observed pattern is different - significantly so

- from what we would expect (a random distribution of responses)?

Well you bet, with another Hypothesis Test.

This time, because we're examining count data,

we'll use what's called a chi-square Goodness-of-Fit test, which

Help

compares the OBSERVED counts from a Table of Counts

to calculated EXPECTED counts.

Now an interesting side note: this test and it's new distribution

was developed by Karl Pearson, and then expanded on by Sir Ronald Fischer.

Notice those two names tend to come up a lot in our introductory stats classes.

As always with Hypothesis Tests, we've got some assumptions to cover.

Number ONE: We need a random sample.

Number TWO: That sample needs to be made up of independent observations.

Number THREE is actually two assumption that I tend to lump together:

No expected count can be less than one; and no more than 20%

of the cells with an expected count can be less than five.

On to the hypothesis test.

The first step, the Null and Alternative hypothesis statements.

We can write these as formal models using symbol notation,

but I tend to add some plain English where I can:

The Null states that the probability is one quarter for each of the 4

Help

possible behavioral responses, while our Alternative Hypothesis is

just the opposite: that not all probabilities are in fact one quarter.

The SECOND STEP: we establish our alpha, and it's the typical 0.05 level.

On to the THIRD STEP of hypothesis testing: Collect the data

and/or run our analysis.

The chi-square Goodness-of-Fit test formula looks like this.

Let's break it down so we know how it works.

We take each cell's observed count, and subtract from it

the cell's expected count.

We then square that difference, and then divide

by the expected count for that cell.

We do that procedure for all cells in the Table of Counts

and then sum the result.

This will give us our chi-square statistic.

Now let's do this with our data.

As shown earlier, here's our Table of Counts with our observed data.

Now, if we had a random distribution to the behavior following a respiratory event, we would see a 25% expected percentage in each cell.

Now, the chi-square formula uses COUNTS, so we

need to come up the the expected counts for each cell.

Help

Since we have 381 total independent events, 25% of 381

is 95.25 in each cell.

It's perfectly fine that our expected counts are not a whole number.

Now, let's go ahead and break down the computation into successive rows.

We take the observed count and subtract out our expected count for each cell.

Then we square that difference.

And then we divide out the expected count for the cell.

Finally, we sum all those results.

With our data, we have a chi-square statistic of 216.76.

Now, for the big question: How do we know if that statistic represents

that "significant departure" of the observed counts from the expected

counts - how can we tell if we Reject or Fail to Reject our Null hypothesis?

216's a pretty big number, but we don't have context for it.

Help

Well to answer that question, we turn to the chi-square distribution.

This distribution is driven by the degrees-of-freedom

and has this basic shape.

With one degree of freedom, we have this shape, With 5 degrees of freedom,

we have this shape.

And, with 10 degrees of freedom, we have this shape and so on.

You'll notice that this distribution is in fact directional -

we can ONLY start with a chi-square of zero and go up from there.

So, for the chi-square distribution, the critical value is always positive,

and there's always just one, with the upper 5% in the tail being cut off.

For the current data, we have 3 degrees of freedom, as we have 4 cells.

Using our degrees of freedom logic, we know that if we have to total 381,

how many of the four cells can be ANY reasonable count?

Well only three - one cell is needed to reach that 381 total.

Formally, the degrees of freedom for the Goodness-of-Fit test is m-1,

or the number of observed count

Help

categories minus one.

Now, with 3 degrees of freedom, the critical value corresponding

to our alpha level of 0.05 is 7.81.

Now for the fourth step of the hypothesis test: the conclusion.

Since our chi-square statistic is greater than the critical chi-square,

we Reject the Null Hypothesis - the observed distribution

of the categorical variable of behavior does not follow a random distribution.

The pattern of behavior is NOT random.

We can also run the chisq.test() function in R to find the actual

p-value for our chi-square statistic - and we see that it is in fact BELOW

our alpha level of 0.05, indicating for us to Reject the Null Hypothesis.

Now to visualize this result, the easiest and simplest way to show it

is with the Table of Counts with percentages added

or to use the bar plot.

I've even seen some publications use a line to denote the expected count,

so that the readers can easily see just how far off the observed counts are.

And that's how we run a Hypothesis test on categorical data with just ONE

categorical variable.

But wait a second!

What if we didn't have a RANDOM distribution to our Null Hypothesis -

Help

could we test a specific expected distribution pattern?

Sure.

We'll tackle that scenario with our next video.

---

1. Choose the statement that is NOT true about the Chi-square test.

(1 point possible)

- Values for chi square can be negative, positive or zero. ✔
- The chi square test should only be performed on categorical data.
- The shape of the chi square distribution depends on degrees of freedom.
- The equation for calculating chi square is the same for the test of goodness of fit and the test of independence.

Help

Hide Answer

2. A major snack food company claims that its chips are "America's favorite." A statistics class tests this claim by asking a sample of 90 random students on campus to select their favorite chip from the company's (Brand A) and two other brands (Brand B and Brand C). Below are the results of how many students selected each brand in their taste test.

Brand A	**Brand B**	**Brand C**
38	28	24

(1 point possible)

2a. What is the distribution model for the null hypothesis?

- 50% Brand A, 25% Brand B, 25% Brand C
- 0% for Brand A
- 33.3% for all three categories. ✔
- Cannot be determined; it is not specified.

Help

Hide Answer

2b. Input or identify the following values necessary for this chi-square test. Assume a confidence level of 0.05.

(4 points possible)

Expected value of Brand A: *(Report as a whole number.)*

Answer: 30

Expected value of Brand B: *(Report as a whole number.)*

Answer: 30

Expected value of Brand C: *(Report as a whole number.)*

Help

Answer: 30

Degrees of freedom: *(Report as a whole number.)*

Answer: 2

Hide Answer

(1 point possible)

Chi-square statistic: *(Rounded to 2 decimal places.)*

- -2.54
- 6.02
- 3.47 ✔
- 2.24

Help

Hide Answer

(1 point possible)

Chi-square critical value: *(Rounded to 2 decimal places.)*

- 4.78
- 2.25
- 6.56
- 5.99 ✔

Hide Answer

(1 point possible)

2c. What conclusion should be drawn about the popularity of Brand A?

- There were more people that liked Brand A in our sample, so Brand A is the favorite of Americans.
- There is no evidence to suggest that Americans prefer Brand A to the other brands tested. ✔
- The proportion of people that liked Brand A best was significantly higher than the proportion that liked Brand B or Brand C.
- We violated the assumption of minimum cell size, so the results of this test are not reportable.

Help

Hide Answer