LECTURE 8: FURTHER HYPOTHESIS TESTING

Goodness of fit test

Chris Robertson

Testing Goodness of Fit

Scientific Amer... / Vol. 199, No. 2... / CELESTIAL NAVIG...

JOURNAL ARTICLE

CELESTIAL NAVIGATION BY BIRDS

E. G. F. Sauer

Scientific American

Vol. 199, No. 2 (August 1958), pp. 42-47

Published by: Scientific American, a division of Nature America, Inc.

https://www.jstor.org/stable/24941079

Page Count: 6

This is an article from 1958 which demonstrated that a particular type of bird used the orientation of the stars in the night sky to determine which way to fly when they migrated. This paper motivates the example in this study session.

Testing Goodness of Fit

Let’s look at an example based upon this experiment

Birds come to a feeding station. They can set off in one of 4 directions – north, east, south, west. Are birds equally likely to set off in any direction?

Over the course of 3 days, 84 birds were observed and the following data collected

North: 20; East: 27; South: 18; West: 19

Hypotheses

If birds are equally likely to set off in one of 4 directions then

$P(\text{North}) = P(\text{East}) = P(\text{South}) = P(\text{West}) = \frac{1}{4}$

This is the Null Hypothesis (H_0).

The alternative is that the probabilities are not equal to ¼. We just need at least one, but if there is one not equal to ¼ there must be at least one other as the probabilities must all sum up to 1.

H_1: probabilities not as specified in H_0.

Hypotheses

This is a more complex hull hypothesis than we had with the one sample tests of a mean.

Here we are testing if a particular probability distribution is able to describe a set of data.

In the example the probability distribution is a discrete uniform distribution but this type of test can be used for more complex distributions

The principles of the hypothesis test are exactly the same as for the one sample tests of the mean.

Expected Values

If H_0 is true, what do we expect to happen?

We make things a little bit more general by introducing the notation $P(X_i) = p_i$ so you can see what to do in other experiments.

$P(X_i)$ refers to the probability that the direction X_i is chosen where i takes the values 1,2,3,4 with X_i = North, etc.

Also $\sum_i p_i = 1$

Expected Values

If H_0 is true, what do we expect to happen? We make things a little bit more general by introducing the notation $P(X_i) = p_i$ so you can see what to do in other experiments.

There are n=84 observations and each of the 4 outcomes has the same probability – $P(X_i) = ¼ = p_i$, i = 1, 2,3,4.

X_1 – North; X_2 – East; X_3 – South; X_4 – West;

If H_0 is true we would expect there to be

$E_i = n \times p_i$ observations corresponding to outcome i.

$E_i = 84 \times ¼ = 21$

E_i is the expected value

Observed Expected

To see if the data provide evidence against the null hypothesis compare the observed and expected values

	Observed	Expected
North	20	21
East	27	21
South	18	21
West	19	21
Total	84	84

If the observed are quite different from the expected then this will suggest that H_0 is not supported by the data

Observed Expected

A natural way to compare the observed and expected values is to look at their differences

These differences will always add up to zero

	Observed	Expected	O-E
North	20	21	-1
East	27	21	6
South	18	21	-3
West	19	21	-2
Total	84	84	0

Also, if the expected value is large then a small difference is not going to be very important

Observed Expected

This is what the observed and expected values might look like if these as a big difference

	Observed	Expected	O-E
North	10	21	-11
East	11	21	-10
South	25	21	4
West	38	21	17
Total	84	84	0

And a small difference

	Observed	Expected	O-E
North	20	21	-1
East	21	21	0
South	22	21	1
West	21	21	0
Total	84	84	0

Observed Expected

Also, if the expected value is large then a small difference is not going to be very important

What this means is that if the expected values were 210, rather than 21 (i.e we observed 840 birds rather than 84) then a difference of 6 from 216 to 210 would not be a very big relative difference and is relatively unimportant given that you were expecting 210 (6/210=0.029).

A difference of 27 to 21 is also 6 but is a bigger relative difference (6/21 = 2/7=0.286)

So if the expected values are large then small differences do not count for much

Observed Expected

Calculate $\frac{(O-E)^2}{E}$; this will always be positive and is equal to zero when the observed and expected values are exactly the same

	Observed	Expected	O-E	(O-E)^2/E
North	20	21	-1	0.047619
East	27	21	6	1.714286
South	18	21	-3	0.428571
West	19	21	-2	0.190476
Total	84	84	0	2.380952

Test Statistic
Chi-Squared Test

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Observed Expected

The reason for dividing the squared difference $(O - E)^2$ by E is to take into account the impact of the differences relative to the expected value.

The reason for squaring the differences is to get positive differences.

Mathematically it is better to square the differences to get positive values rather than taking the absolute value.

The reason for this is because squaring the differences leads us to a probability distribution which we can use

You will not do the mathematics of this until 3rd year in Mathematics and Statistics.

Observed Expected

If the expected value is 10 and the difference 3 then
$$\frac{(O-E)^2}{E} = 0.9$$

If the expected value is 100 and the difference 3 then
$$\frac{(O-E)^2}{E} = 0.09$$

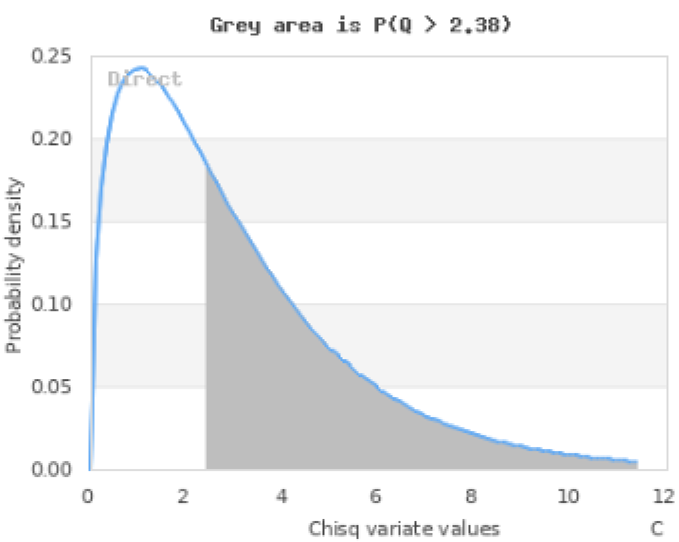
Small differences with large expected values do not have a big contribution to X^2;

Large differences with small expected values do have a big contribution.

Chi Squared Disribution

If H_0 is true then the test statistic follows a χ^2_{n-1}distribution - a “chi-squared” distribution on n-1 degrees of freedom; $E_i > 5$

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$



This is another probability distribution

It is a Skew distribution

The Minimum is zero

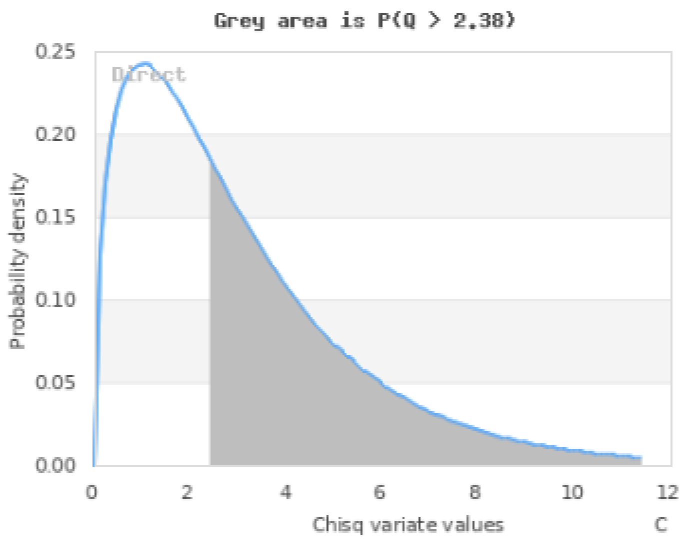
It has one parameter – the degrees of freedom

Same as t-distribtuion

Chi Squared Disribution

If H_0 is true then the test statistic follows a χ^2_{n-1} distribution - a “chi-squared” distribution on n-1 degrees of freedom; $E_i > 5$

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$



The degrees of freedom (df) controls the location of the distribution

The mean is the df

The higher the df the more the peak is moved to the right (high values)

Hypotheses and Chi Square

H_0: P(North) = P(East) = P(South) = P(West) = ¼

If H_0 is true exactly then $$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 0$$

H_1: probabilities not as specified in H_0.

If H_1 is the true hypothesis then the O_i values will be far away from the E_i values and this means that $\frac{(O_i - E_i)^2}{E_i}$ will be a lot bigger than zero.

So large values of X^2 indicate that there is evidence to reject H_0. This is always a one sided test.

Hypotheses and Chi Square

H_0: P(North) = P(East) = P(South) = P(West) = ¼

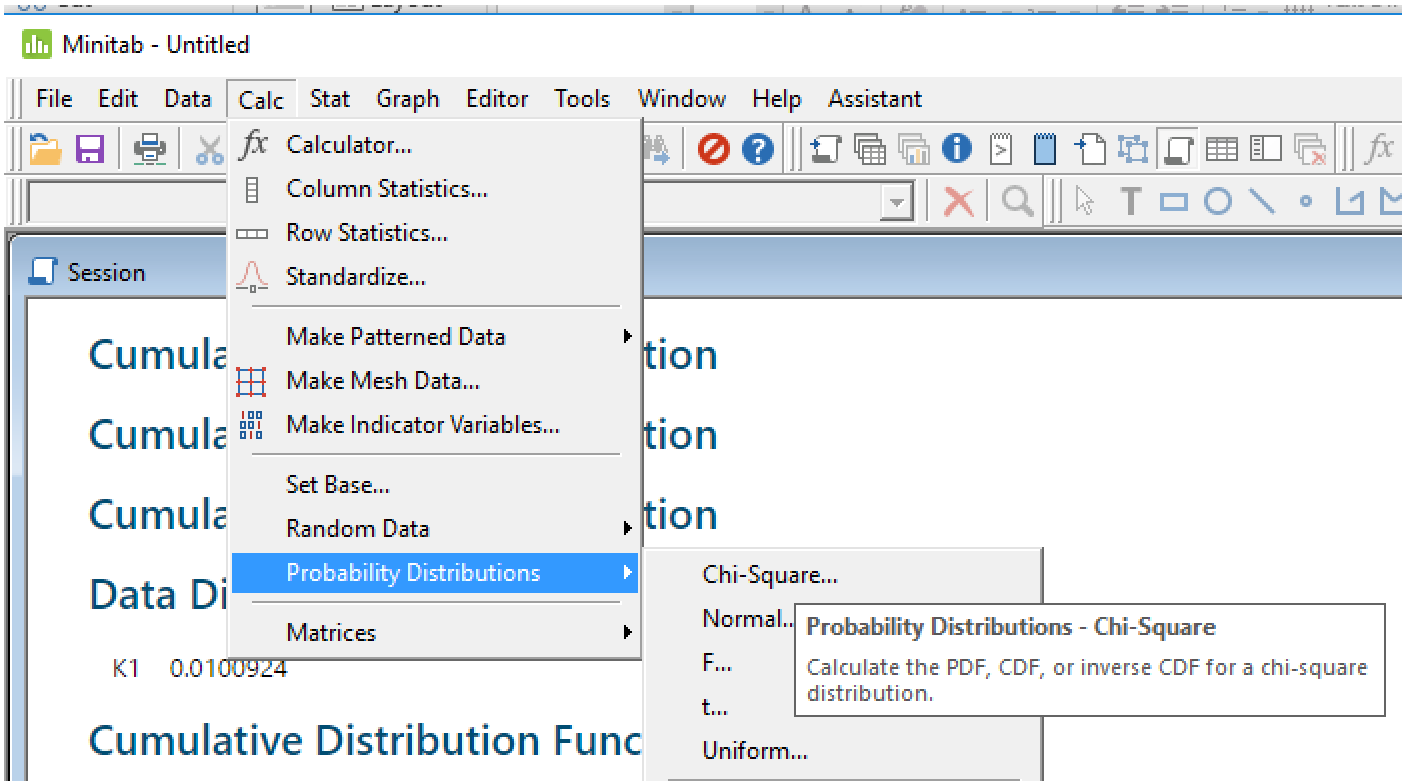
H_1: probabilities not as specified in H_0.

X^2= 2.38, df=4-1 = 3

P-value is 0.497 > 0.05

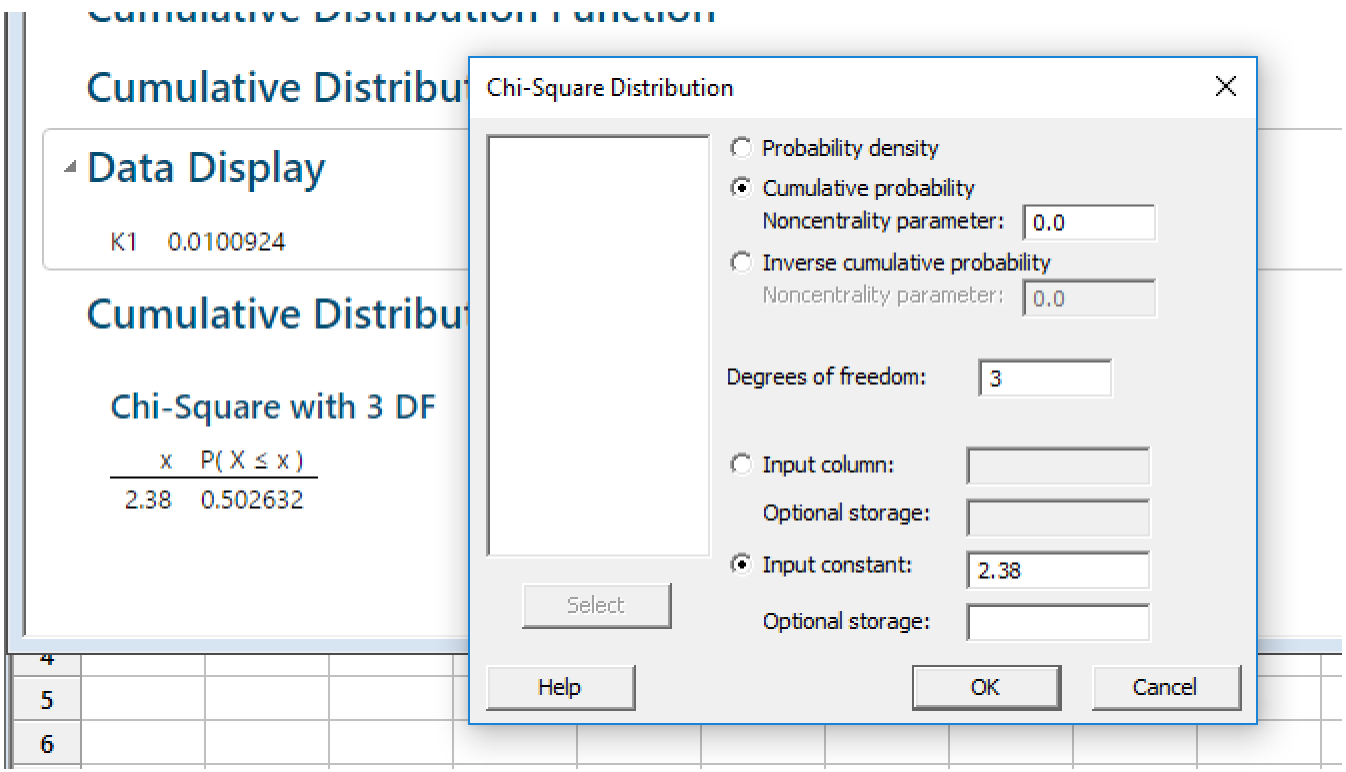
At the 5% significance level there is
No evidence to reject H_0
Conclude that birds move away
in the 4 directions at random

Chi Square p Values Minitab



Select Calc, then Probability Distributions, then Chi-Square

Chi Square p Values Minitab



Enter the degrees of freedom and the calculated value

This give P(X^2 < value) so you need to subtract the probability away from 1 to get the p value = 1-0.502632 = 0.497368

Key Points

Statistical testing and estimation (through confidence intervals) are all based upon probability models for the variation in the data

T-test and z-test test the value of a mean within an assumption of a normal distribution to describe the variation in the characteristic measured in the population

X^2 goodness of fit test is used to test if a probability model is valid.

Key Points

There are hundreds of statistical tests for different situations

Principles of each test are the same

H_0 represents your belief before collecting the data
Test statistic has a known distribution if H_0 is true.
Extreme values of the test statistic (high/low) indicate that H_0 is unlikely to be true

You can only reject H_0 or conclude that there is no evidence to reject H_0

It is wrong to say you accept H_0

Key Points

True hypothesis	Decide no evidence to reject H_0	Conclude evidence to reject H_0
H_0	Correct	Type 1 Error – significance level
H_1	Type 2 Error	Correct

In any test two types of errors can be made

Type 1 Error has a small probability defined by the significance level

Type 2 Error is controlled by the sample size and can have a large probability in small samples