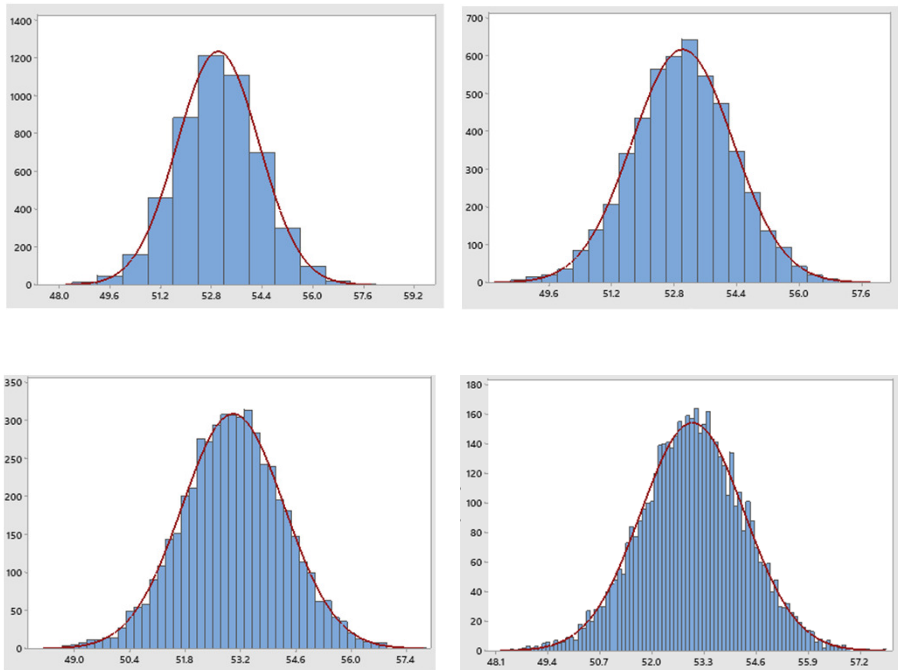
UPCSE Biology – Basic Statistic course

Testing for goodness-of-fit

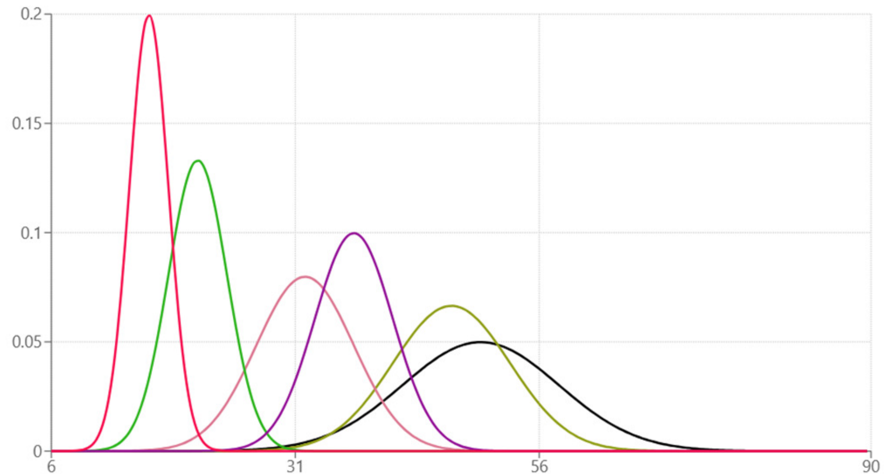
The χ^2 (chi-squared) test

Revision

1) More data leads to a better defined normal curve



2) Each species has its own normal curve leading to thousands of normal curves

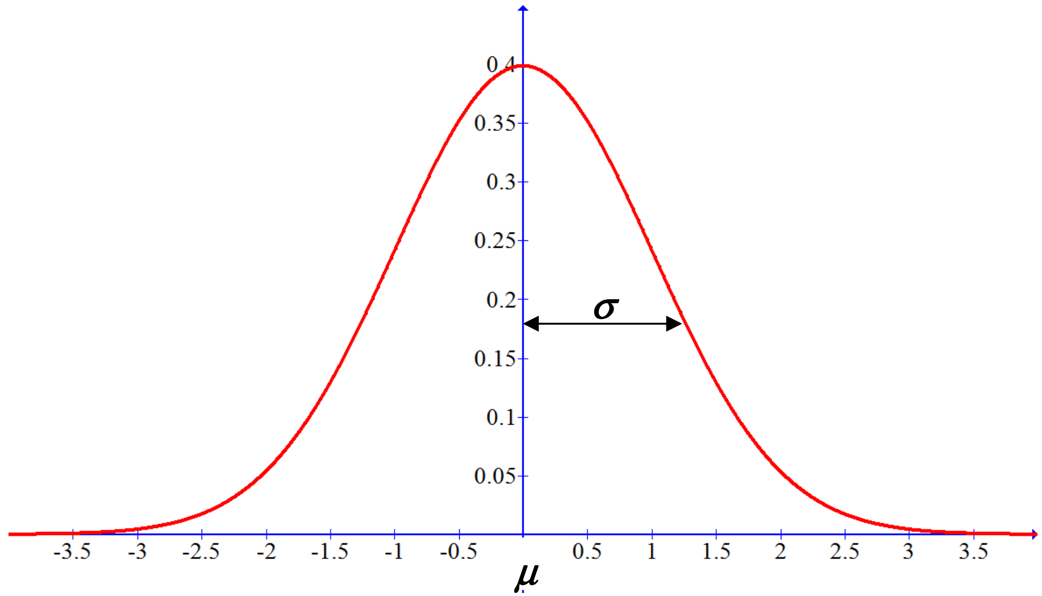


Revision

3) Standardise all normal curves via the z transformation:

$$z = \frac{x - \mu}{\sigma}$$

4) This leads to the standard normal distribution



The standard normal distribution curve is a plot of all values from the z formula

Revision

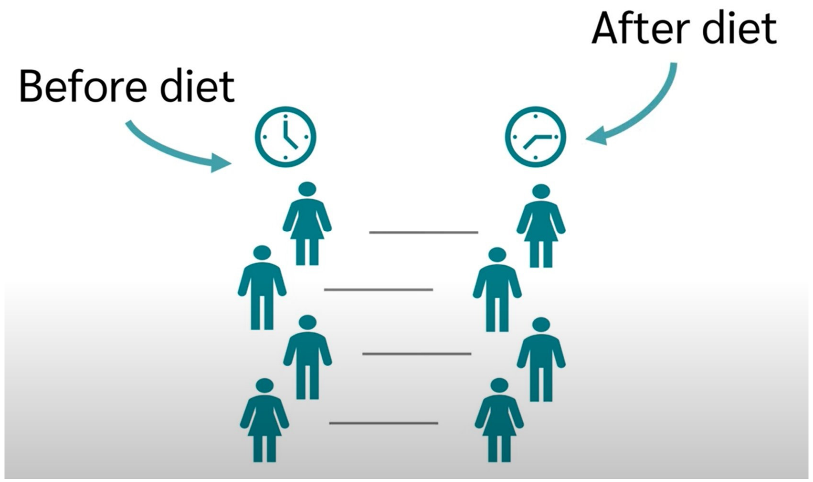
- We have seen in previous slides that
  - if we take *large* samples from a normally distributed population we can perform significance testing using the z-distribution,
    - the z-test is used to test hypotheses based on population data (here we know μ and σ)
  - if we take *small* samples from a normally distributed population we have to perform significance testing using the t-distribution.

Revision

- The t-test is used to test sample means which comes from a standard normally distributed population,
  - in a before-and-after situation;
  - where our sample size is too small to use the z-test;
  - when we don’t know the population standard deviation, nor the actual population mean (but a suitable reference μ can be used).
- Below is a general diagrammatic representation of this situation.

Revision

5) For our course we do the t-test to test means in a before-and-after situations.



The group's mean weight is or is not significantly different after the diet compared to before the diet

6) Now set up hypotheses:

- Null hypothesis H_0:
  The mean weight is statistically equal to the population mean.
- Alternative hypothesis H_1:
  The mean weight is statistically different from (greater and/or less than) the population mean.

Revision

7) Do the t-test using Minitab, or by hand with

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

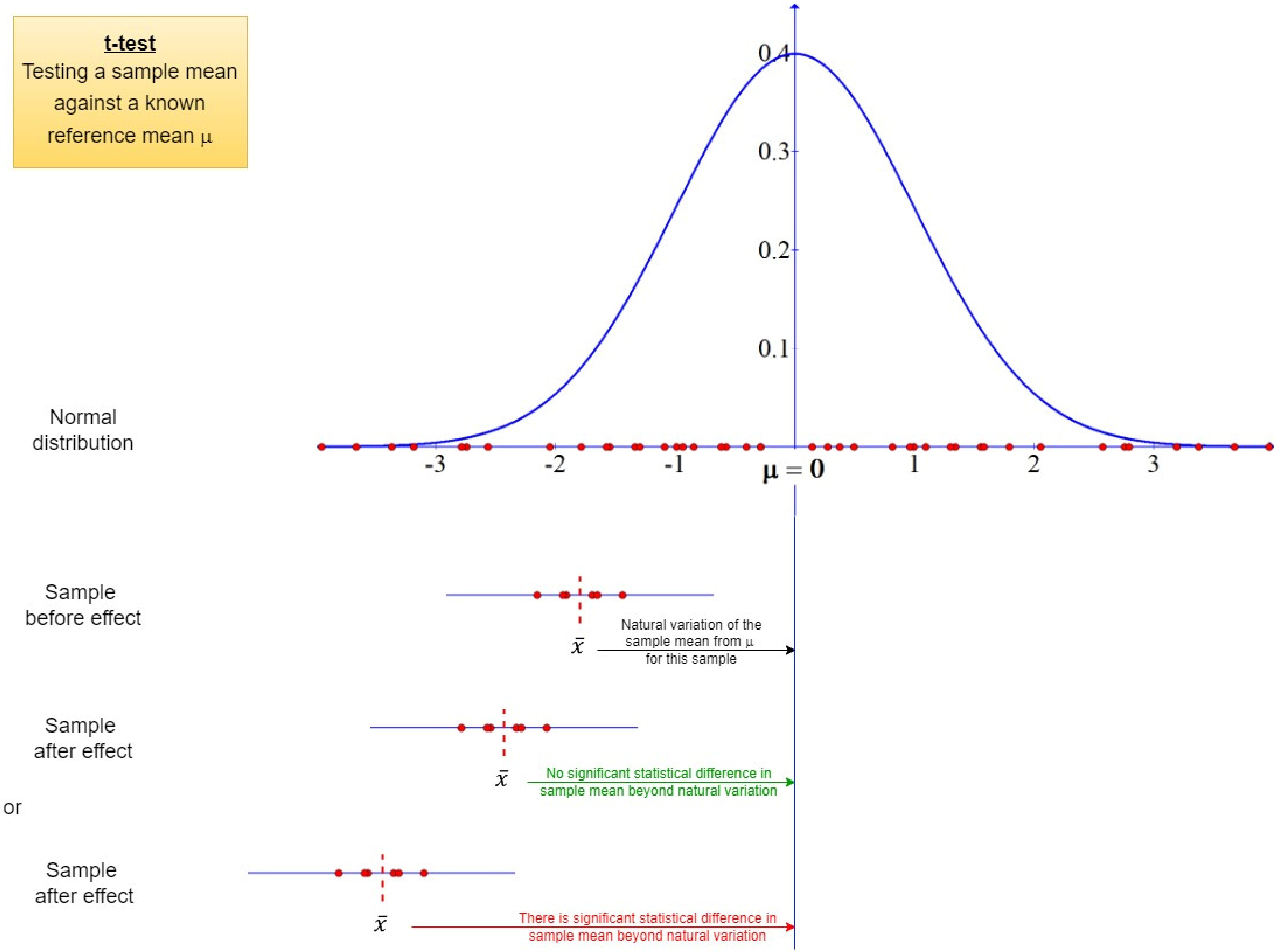
and appropriate degree of freedom.

8) Set an appropriate α-level, i.e. level of significance: 0.1, 0.05, or 0.01

9) Look up the t value you found in 7) (or calculated by Minitab) against the t value in tables for the degree of freedom and your chosen α-level.

10) Decide to accept or reject H_0. Either this is a statistically significant difference or not for you chosen α-level.

11) If there is a statistically significant difference look to the biology as to why.


t-test
Testing a sample mean against a known reference mean μ
0.4
0.3
0.2
0.1
Normal distribution
-3
-2
-1
μ = 0
1
2
3
Sample before effect
x̄
Natural variation of the sample mean from μ for this sample
Sample after effect
x̄
No significant statistical difference in sample mean beyond natural variation
or
Sample after effect
x̄
There is significant statistical difference in sample mean beyond natural variation

χ^2 (chi-squared) testing

- We will now study another statistical test called the χ^2 (pronounced chi-squared) test.
- This test will be used to see how close our experimental data is to the theoretical data we could have collected in an ideal world where we could have done the experiment millions of times.
- But the data we will be collecting will now come in the form of **categories** which are mutually exclusive.

χ^2 (chi-squared) testing

- For example
  - Red, pink, and white – coloured flowers;
  - The blood type of a person: A, B, AB or O;
  - Types of rocks: igneous, sedimentary, metamorphic
- "Mutually exclusive" means:
  - Red is red, not pink or white;
  - Someone has blood type O only, not O and AB
  - Rocks are either igneous, sedimentary, or metamorphic but not a combination of each.

χ^2 and categorical data

- "Categorical data represent the distribution of samples into several mutually exclusive categories, which usually involves counting how many objects are in each qualitative category."

From https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3021327/

- So, more precisely, given an object of study (flowers, blood type, etc.), the test is about seeing how close each category is to the respective category of the perfect theoretical data we could have collected in an ideal world.

χ^2 (chi-squared) testing

- Another way of saying this is, does the distribution of our experimental data fit well enough the distribution of the theoretical data?
- Yet another way is, is the theoretical model which represents the theoretical data a good enough fit or model for our experimental data?
- The chi-squared test then asks, How well does our experimental data fit the theoretical data?

General aim of why we use χ^2 distribution and test

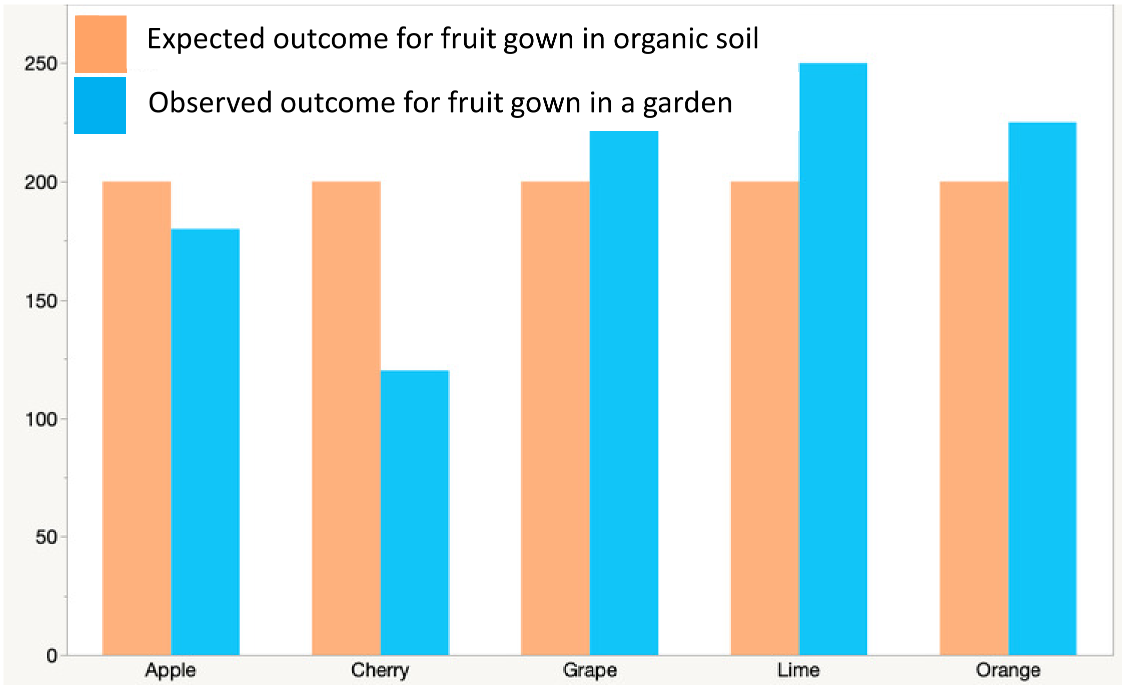
- In other words, Is their a significant difference between
  - the number of red flowers, pink flowers, and white flowers after a cross-pollination experiment,

  compared to
  - the number of red flowers, pink flowers and white flowers we would expect theoretically from genetic theory (i.e. Mendelian theory)?
- This is illustrated below using the example of growing fruit/

General aim of why we use χ^2 distribution and test

Example: Is there a significant difference between each different fruit grown in a garden and the same fruit grown in organic soil?

I.e., is there a significant difference in the height of the blue bars compared to the height of the orange bars?



Comparison of number of fruit grown

Chi-squared testing allows us to compare the effect of each category (observed against expected) on the total (all fruit)

General aim of why we use χ^2 distribution and test

- To find out we do a simple difference calculation between
  - the data collected from experiment, called *observed data* (the blue bars),

  and
  - the data we would expect to obtain in a theoretically perfect situation, called *expected data* (the orange bars) where, here, the theoretical expectation is that we should be able to have 200 of each fruit.
- It is for us to know the distribution of expected data. More on this later.

General aim of why we use χ^2 distribution and test

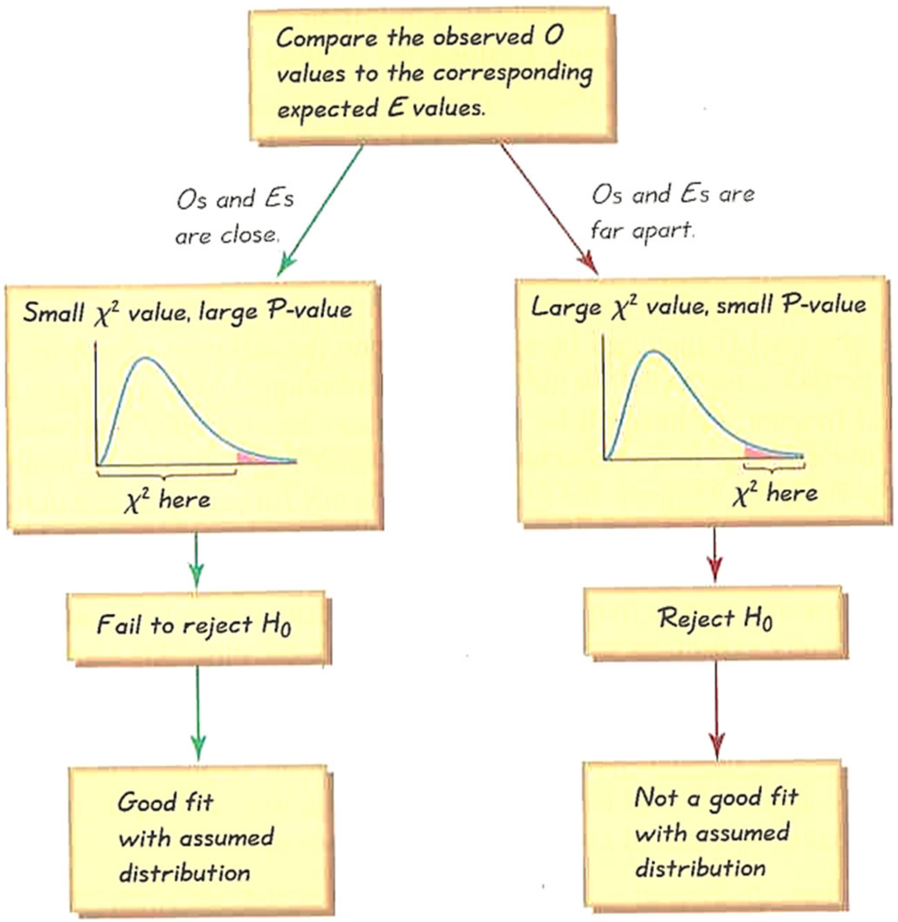


Figure 11-3 Relationships Among the χ^2 Test Statistic, P-Value, and Goodness-of-Fit

If the observed distribution (O) is close to the expected distribution (E) then our data is a good fit to the theoretical data/model.

Otherwise it is not a good fit to the theoretical data/model.

The χ^2-squared test helps us determine this

(this picture is taken from p594 of Mario Triola's "Elementary statistics" (10th Edition))

General aim of why we use χ^2 distribution and test

- So, the χ^2 test is based on the differences between O and E, so
  - close agreement between O and E will give a small χ^2 value (a large p-value);
  - A large discrepancy between O and E will give a large χ^2 value (a small p-value);

On observed and expected outcomes

Example

- Consider an unbiased coin. The word “unbiased” is a technical term in stats meaning that the coin has not been altered in any way, and that there is an equal chance of getting a heads or a tails when tossing the coin.
- Now consider tossing this unbiased coin 10,000 times

On observed and expected outcomes

Example

- The distribution of heads and tails would be very close to 5000 heads and 5000 tails, say 4998:5002
- Here, 5000 is the *expected value* for getting a Heads as well as a Tails, and 5000:5000 is the expected distribution relating to this experiment an infinite number of times.

On observed and expected outcomes

Example

- If we repeat the experiment 100 times then our expected distribution will be 50:50.
- If we repeat the experiment 20 times then our expected distribution will be 10:10.
- If we repeat the experiment 8 times then our expected distribution will be 4:4.
- Etc.

On observed and expected outcomes

Example

- For the sake of practicality let us say that performing the experiment 1,000,000 times acts as an "infinite" number of times …
- … then our *expected values*, i.e. the theoretically perfect outcome, would be for us to have a distribution of

500,000 heads : 500,000 tails.

On observed and expected outcomes

- Now consider again tossing this unbiased coin 10 times.
- It is possible for us to actually get the result of 5:5, heads:tails on one throw.
- But it is also very likely that we will get another result.
- So, suppose we obtain 3 heads and 7 tails. These are the *observed* (experimental) *values*.

On observed and expected outcomes

Diag (a): Toss a coin 1,000,000 times and repeat n times

The distribution of H and T is always expected to be 500,000H : 500,000T

Diag (b): Toss a coin 10 times and repeat n times

The distribution of H and T is observed to be

5H : 5T

or 4H : 6T

or 3H : 7T or …

On observed and expected outcomes

- The χ^2 test involves comparing the observed values against the expected values to see how far off our experimental distribution of 4:6 or 3:7 (observed values) is from the theoretical distribution of 5:5 (expected values).

On observed and expected outcomes

- The question then becomes,
  - How significantly different is a distribution of 4:6 or 3:7 compared to 5:5?
  - I.e., how far away can our observed results of
    - 4H and 6T
    - or 3H and 7T

    be from the expected 5H and 5T to still consider the coin to be unbiased/fair?

On observed and expected outcomes

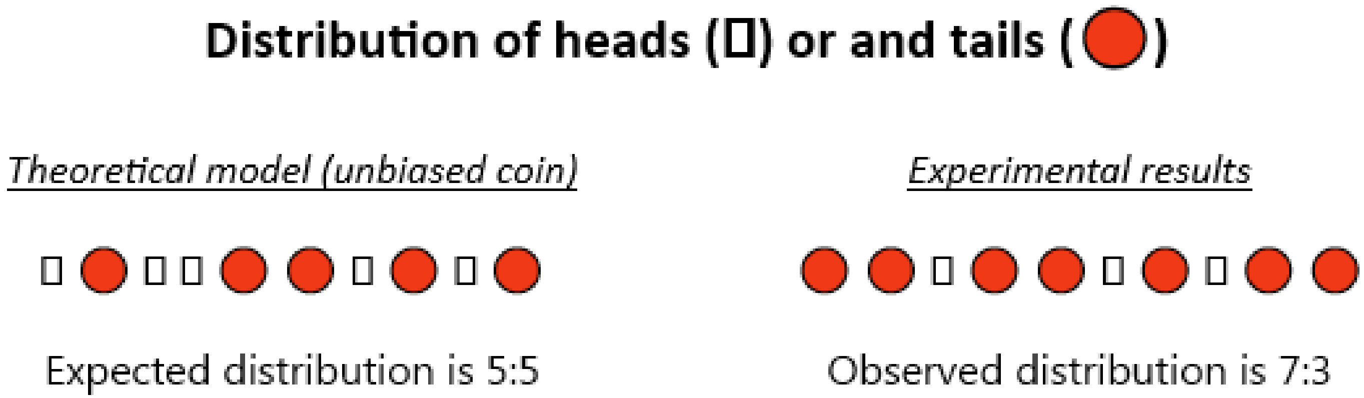
- Clearly, if we had obtained observed values of 5 heads and 5 tails then this would compare exactly with the expected values. So the coin would be unbiased/fair.
- We would also consider observed values of 4 heads and 6 tails (or 6 heads and 4 tails) to compare well against expected values, so the coin would still be unbiased/fair.

On observed and expected outcomes

- What about observed values of 3 heads and 7 tails, or 2 heads and 8 tails?
- How far off the expected values can we be for us to conclude that the coin is biased (i.e. it is a trick coin of some sort)?
- What amount of departure from the theoretical expected model counts as statistical significance?
- This is what the χ^2 test allows us to decide.

On observed and expected outcomes

- Below is a general diagram representation of this situation.



- *Question:* Is the distribution of our experimental results significantly different from that of the (expected) theoretical model?
- χ^2 testing for goodness-of-fit helps us answer this.

Expected vs observed values

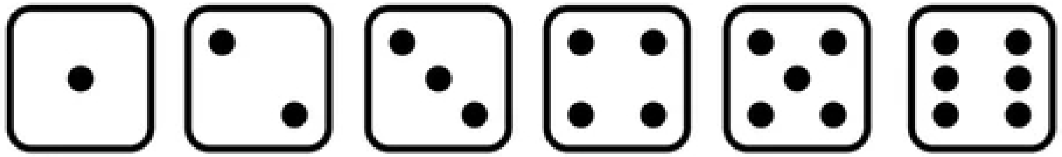
Another example

- As a more complex example, consider doing an experiment with a 6-sided die.

- Suppose we decide to throw it 120 times and record the number of 1s, 2s, 3s, 4s, 5, and 6s that occur. In theory we should **expect** to get each value one-sixth ($1/6^{th}$) of the time, i.e. 20 times each.

Expected vs observed values

Another example



We can expect each value to appear 1/6 = 0.1666666… times

When throwing the die 120 times we would expect each value to appear 120×1/6 = 20 times,

If we threw the die many more times, say 480 or 960 times, we would expect our frequencies for each score to lie more closely around 80 (480/6) or 160 (960/6).

Expected vs observed values

- This is the theoretical score, i.e. the score we would **expect** to get in a perfect world of
  - a perfect die,
  - no human or experimental errors,
  - Throwing the die perfectly randomly every time,
  - conducting the experiment an infinite number of times,

  etc.
- In that case we would expect to get exactly the same number of 1s, 2s, 3s, 4s, 5, and 6s from 120 throws.

Example

- We could therefore get a situation such as that illustrated by the table below where we throw the die 120 times.
- Our aim would be to see how well our observed values match the expect values.

Score, x	1	2	3	4	5	6
Observed values (experiment)	15	22	23	19	23	18
Expected values (theory)	20	20	20	20	20	20

Problem statement

- Remember the steps we need to go through in doing significance testing. In the case of χ^2 testing we have
  - Problem statement,
  - Hypothesis statement,
  - Observed and expected values,
  - The χ^2 test and critical value
  - Significance-levels (addressed in previous lesson)
  - Degrees of freedom (addressed in previous lesson).

A Complete Example

- In a genetics experiment pure red-flowered plants are crossed with pure white-flowered plants (this is the F_0 generation).
- The next generation (F_1) all produced pink-flowered plants.
- The F_1 generation were allowed to inter-pollinate randomly to produce generation F_2.

A Complete Example

- The F_2 generation then consisted of 145 red-flowered plants, 289 pink-flowered plants, and 138 white-flowered plants.
- **Question:** Does the ratio 145 : 289 : 138 of red : pink : white-flowered plants agree with the expected Mendelian genetic theory ratio if 1 : 2 : 1?
  - **Note** that 145, 289, and 138 are frequencies or counts. The ratio 145 : 289 : 138 represent the observed distribution of these counts. We want to compare this against the expected distribution of 1 : 2 : 1.

Setting up hypotheses

Terminology & Notation

- The statistical variable we are testing here is “flower colour”.
- This variable has three categories: red, pink, white.
  - “Null hypothesis, H_0” is the term and notation used for the hypothesis that no effect or change has occurred beyond the way the biology naturally varies.
  - H_0 is expressed as an equation.

Setting up hypotheses

Terminology & Notation

- The statistical variable we are testing here is “flower colour”.
- This variable has three categories: red, pink, white.
  - “Alternative hypothesis, H_1” is the term and notation used for the hypothesis that an effect or change has occurred beyond the way the biology naturally varies. So there is a significant (other) biological effect.
  - H_1 is expressed as an inequality.

Setting up hypotheses

- **General examples of hypotheses**
  - ① H_0: "The number of geranium seeds germinating in composts A, B and C are the same"

    H_1: "There is a difference in the number of geranium seeds that germinate in composts A, B and C"
  - ② H_0: "The pH of a medium in which enzyme activity is measured does not affect the rate of reaction of enzyme catalysis."

    H_1: "The rate of enzyme catalysed reaction is affected by the pH of the medium in which enzyme activity takes place."

Setting up hypotheses

- **General examples of hypotheses**

  ③ H_0: "There is no bias in a lab technician's ability to read a scientific instrument."

or H_0: "The lab technician is able to read the scientific instrument correctly."

or H_0: "The lab technician's readings are accurately recorded."

How would you state H_1?

Complete Example continued – Our hypotheses

- So, our hypotheses are therefore
  - H_0: The ratio 145 : 289 : 138 of red : pink : white **is** in agreement with Mendelian ratio 1 : 2 : 1
  - H_1: The ratio 145 : 289 : 138 of red : pink : white is **not** in agreement with Mendelian ratio 1 : 2 : 1
- H_0 is called the *null hypothesis*, and H_1 is called the *alternative hypothesis*.

Complete Example continued – Observed and expected values

- So we have:

	Ratio (number of observations per category of red, pink, white)	**Total amount across all categories**
Observed (experimental) distribution	Red : Pink : White 145 : 289: 138	572
Expected (theoretical) distribution against which we are comparing	Red : Pink : White 1 : 2 : 1	4

- But 572 is 143 times bigger than 4 so we should have

Expected (theoretical) distribution based on our observed distribution	143 : 286 : 143	572

Complete Example continued – Observed and expected values

- Therefore we can set up our results as such:

	O (observed)	E (expected)
Red	145	143
Pink	289	286
White	138	143
Totals	572	572

Calculating the χ^2 test value

- Now that we have our table of observed and expected value we need to calculate something called an χ^2 test value.
- This test value will be some kind of measure of the size of the difference between our O and E values.
- This is how we do it. The χ^2 formula is given by

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

where O are observed values, E are expected values.

Calculating the χ^2 test value

- This formula gives a measure of the deviation of experimental data away from the theoretically expected results.
- Dividing by E has the effect of standardising each $(O–E)^2$ value. This then measures the relative size of $(O–E)^2$ compared to E.
- In this way, a small difference will be more important when E is small than when E is large:
  - When E is small, $(O – E)^2$ will be relatively large, potentially implying a significant difference;
  - When E is large, $(O – E)^2$ will be relatively small, potentially implying *no* significant difference;

Calculating the χ^2 test value

- For example, say you are measuring the wing span of two types of bats, and you collected the following data:

	Bat type 1	Bat type 2
Observed length	51 cm (O_1)	6 cm (O_2)
Expected length	50 cm (E_1)	5 cm (E_2)

- Then
  - $O_1 - E_1 = 1$ cm, and $O_2 - E_2 = 1$ cm.
  - But 1cm is quite small compared to 50cm, …
  - … whereas 1cm is quite large compared to 5cm.

Complete Example continued – Our χ^2 test value

- For our red-pink-white flower example we have:

	O	E	O – E	$(O – E)^2$	$(O – E)^2/E$
Red	145	143	2	4	0.028
Pink	289	286	3	9	0.031
White	138	143	–5	25	0.175
	$\chi^2 = \Sigma[(O - E)^2/E] =$				0.234

Significance of the χ^2 value

- The χ^2 value is a measure of the deviation away from the expected value.
- The question then is, when does the deviation become so large that it is not due to normal variation in each sample, but to actual biological/genetic effect?
- Statistically speaking we are asking:

"Is this χ^2 value significant?"

Significance of the χ^2 value

- The word “significant” is a key statistical term.
- The term “significant” is such an important term to understand in stats that I will repeat myself: What it means is, does our value of χ^2
  - imply that the difference between O and E values is due to a real biological/genetic effect?

  or
  - does it imply that the difference between O and E values is simply due to sample variation (i.e. to us having picked one sample instead of another)?

Significance of the χ^2 value

- Again, significance means that

"the observed difference appears to be real and is not due only to random chance in the sample(s). Whether or not a result that is statistically significant is also biologically significant is another question.

Moreover, the term significant is not an ideal one, but because of long-standing convention, we are stuck with it. Statistically *plausible* or statistically *supported* may in fact be better terms."

(taken from "**A biologist's guide to statistical thinking and analysis**", David S. Fay and Ken Gerow)

The χ^2 distribution

- The significance of our χ^2 value can only be decided by comparing it against the relevant theoretical model for χ^2.

- Remember that every real life situation has a theoretical model against which we can compare it:
  – heights of people or weights of birds eggs or length of shoots of plants or O_2 concentration in a stream, etc., against normal distribution.

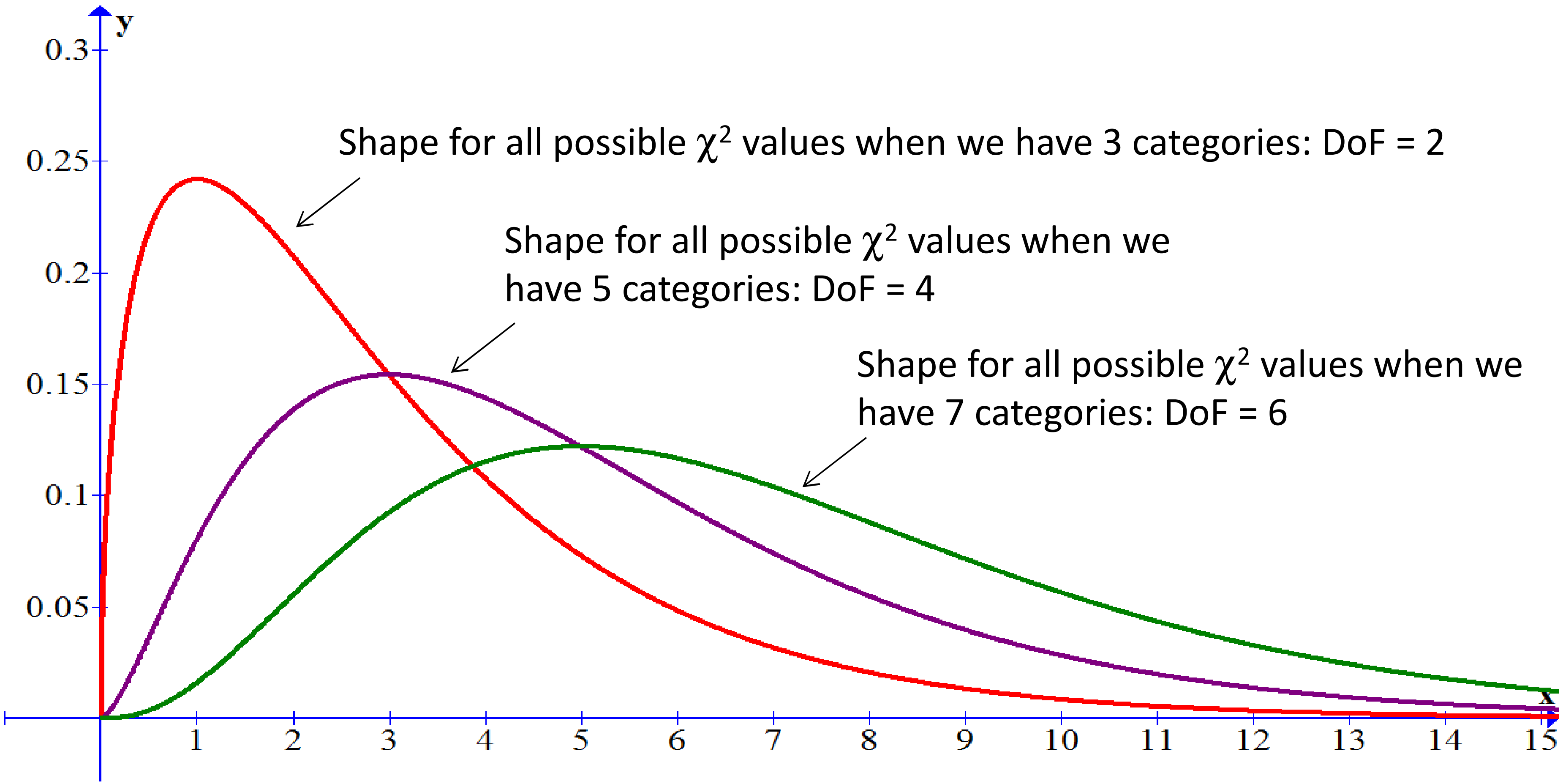
χ^2 (chi-squared) distribution

- The theoretical distribution we are going to use here is not the z- or t-distribution but is now what is known as the χ^2 distribution.
- Just like the standard normal distribution (i.e. the z-distribution) and the t-distribution have particular shapes so does the χ^2 distribution.
- And just as the t-distribution required the use of degrees of freedom (see previous lesson) so does the χ^2 distribution.

Degrees of freedom for χ^2

- This again relates to the idea of degrees of freedom which is designed to take account of variability produced by the number of categories we have.
- The following slide are three χ^2 distribution curves. Later slides will describe where these curves comes from.

χ^2 (chi-squared) distribution

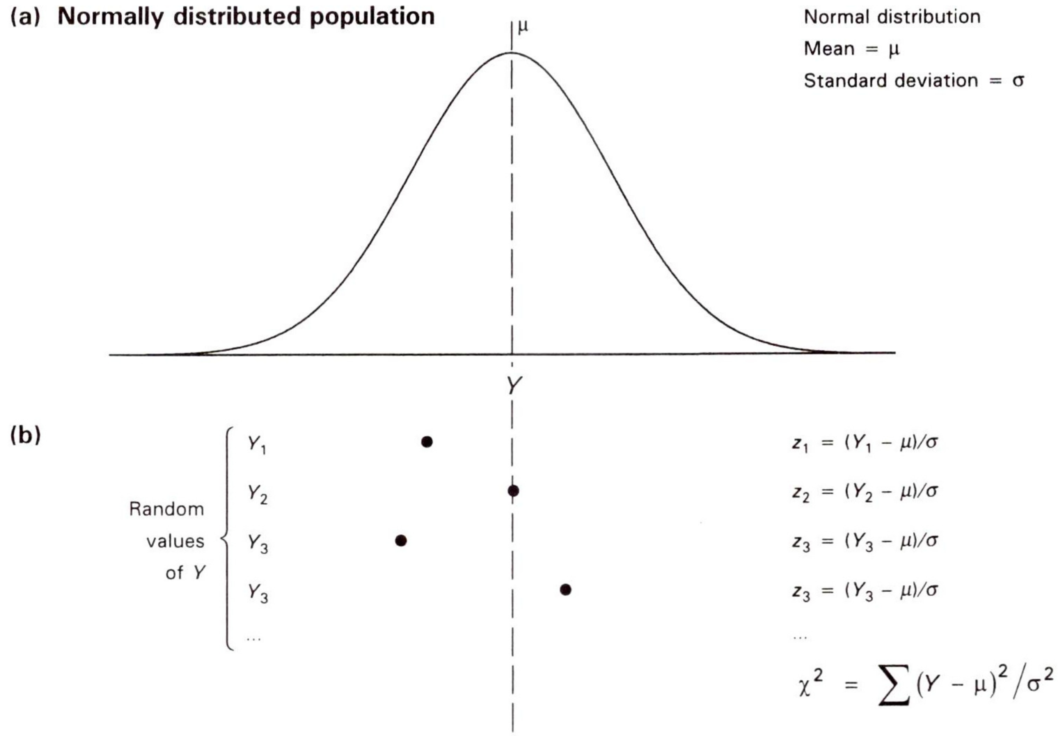


χ^2 (chi-squared) distribution

- These χ^2 distribution curves come from plotting all possible χ^2 values (from the formula on slide 43) for a given degree of freedom …
- … and a specific χ^2 value tells us the extent to which the distribution of our experimental data fits, or doesn't fit, the theoretical distribution.
- We will now see what this means.

The χ^2 distribution

- In more detail, consider a normal distribution of data: diag (a).
- Pick one data value (n=1) at random, say Y_1; find out how much Y_1 is distributed away from mean μ. Call this z_1
- Repeat lots of times;

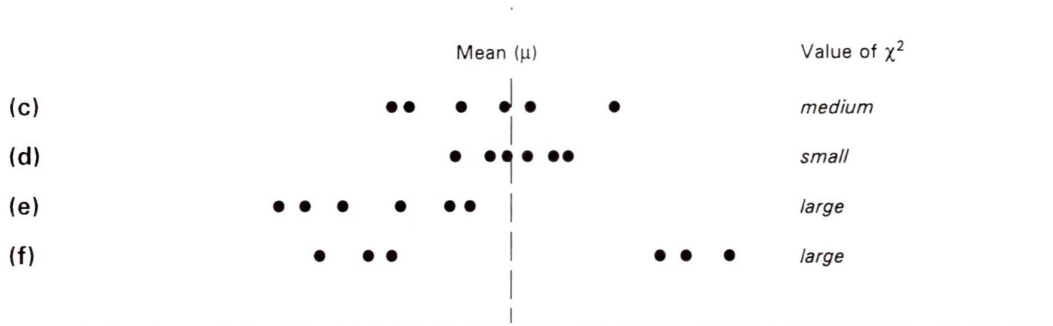


(Picture from p134 David Heath "An introduction to experimental design and statistics for Biology")

Final χ^2 is got from adding up all the z values.

The χ^2 distribution

- The χ^2 value is the sum of the variation of each sample data away from the mean μ.
- Diag (c), (d), (e), (f) below shows a few examples of small, medium, large variations from the mean.



(Picture from p134 David Heath "An introduction to experimental design and statistics for Biology")

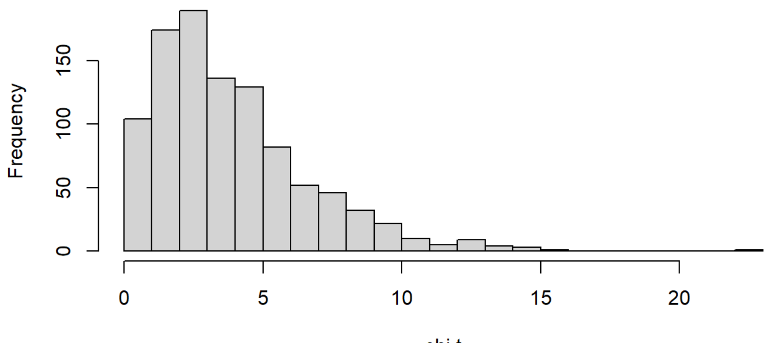
- Each of (c), (d), (e), (f) gives its own χ^2 value. We have five values so far.

The χ^2 distribution

- The χ^2 distribution is the distribution of all possible χ^2 values (of a specific degree of freedom. See later), which comes from performing the experiment an "infinite" number of times and plotting the resulting χ^2 values as a curve.
- This is the theoretical model we compare our calculated χ^2 values against;
- For χ^2 analysis we don't talk of data point X or Y and mean μ. Instead we talk of observed data and expected data.

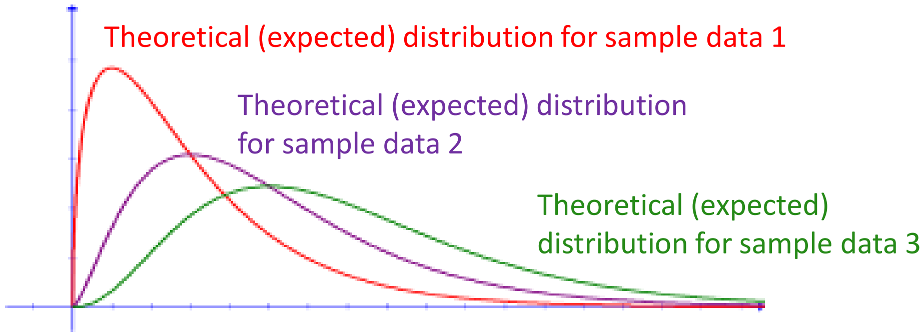
Where does χ^2 come from?

Experimental distribution, for one experiment, of the variation between observed and expected frequencies
(Y and μ in the previous slides)



This experimental distribution would need to be compared with the appropriate χ^2 (theoretical) distribution

Three theoretical distributions (red, purple, green) of total variation of data for each of three sample sizes (specifically, for three different DoF)



Three different distributions χ^2 (theoretical) distributions (red, purple, green)

Where does χ^2 come from?

- **The χ^2 distribution is our theoretical model against which we will compare our χ^2 test value.**
- The area under each curve classifies as a probability value.
- Areas to the right of the critical χ^2 value indicate (possibly) statistically significant results / biological effects because these areas represent very low probability of something happening by chance or natural variation.
- Probability values of the χ^2 distribution are given in tables at the back of textbooks, or via software.

Degrees of freedom - DoF

- We need to be able to find the relevant critical χ^2 value listed in tables so that we can compare this value with our calculated χ^2 value.
- To do this we need to choose the correct χ^2 distribution. This is done by considering the degrees of freedom (DoF) of our distribution.
- For our particular type of χ^2 testing we calculate this as k – 1 where k is the number of categories (this is similar to doing n – 1 for a sample size n, but for χ^2 we don't base DoF on sample size but on number of categories. However, the principle behind DoF here is the same as before. See Word documents of the lesson on t-test).

Degrees of freedom - DoF

- A Chi-squared distribution is based on the number of categories, not the total sample size (n).
- The "shape" of a Chi-squared distribution is determined only by what is called its Degrees of Freedom (DoF or df).
- For a standard goodness-of-fit test, the DoF is calculated as the number of categories minus one $k-1$.

Degrees of freedom - DoF

- *What degrees of freedom are about*

- The Chi-squared test focuses on the *variability between categories*.
- The *DoF* represents how many pieces of information in your category-counts are free to vary independently.
- If you have 4 categories and you know the total number of observations (n), once you know the counts for 3 of those categories, the 4^{th} category is automatically fixed.
- Therefore, you only have $4 - 1 = 3$ degrees of freedom. This is why the distribution is based on categories.

Degrees of freedom - DoF

Example

- In our example on coloured flowers we have 572 observations across 3 categories (red, pink, white flowers).
- If the expected ratio red:pink:white is 1:2:1 then we need to split 572 across the ratio 1:2:1.
- For the red flowers we have 143 flowers out of 572.
- For the pink flowers we have 286 flowers out of 572.

Degrees of freedom - DoF

- *Example*

- What does this mean for the white flowers? Well, we are forced to have 143 white flowers.

- It is like solving $143 + 286 + x = 572$, or more generally $a + b + x = 572$. We are free to change *a* and *b*, but once these are fixed we are not free to choose x since the total is always 572.

- Hence here we have 2 degrees of freedom based on 3 categories.

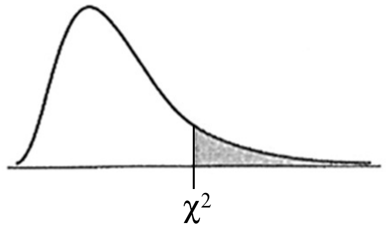
Complete Example continued – Towards the finish line

- Recall that we have calculated $\chi^2 = 0.234$.
- Since all E values are greater than 5 (the minimum needed to make χ^2 testing valid) we can say that $\chi^2 = 0.234$ is valid to use.
- We now find the χ^2 critical value. We do this by choosing an α-level of significance, and finding DoF.
  - Usual α-levels to take are 0.1 or the 10% level of significance, 0.05 or the 5% level of significance, or 0.01 or the 1% level of significance. See Karen/Alice for the usual significance level or confidence level to take.

Complete Example continued – Towards the finish line

- So for $\alpha = 0.05$ and DoF = 3–1 = 2 we have the χ^2 critical value to be as shown in the table below:

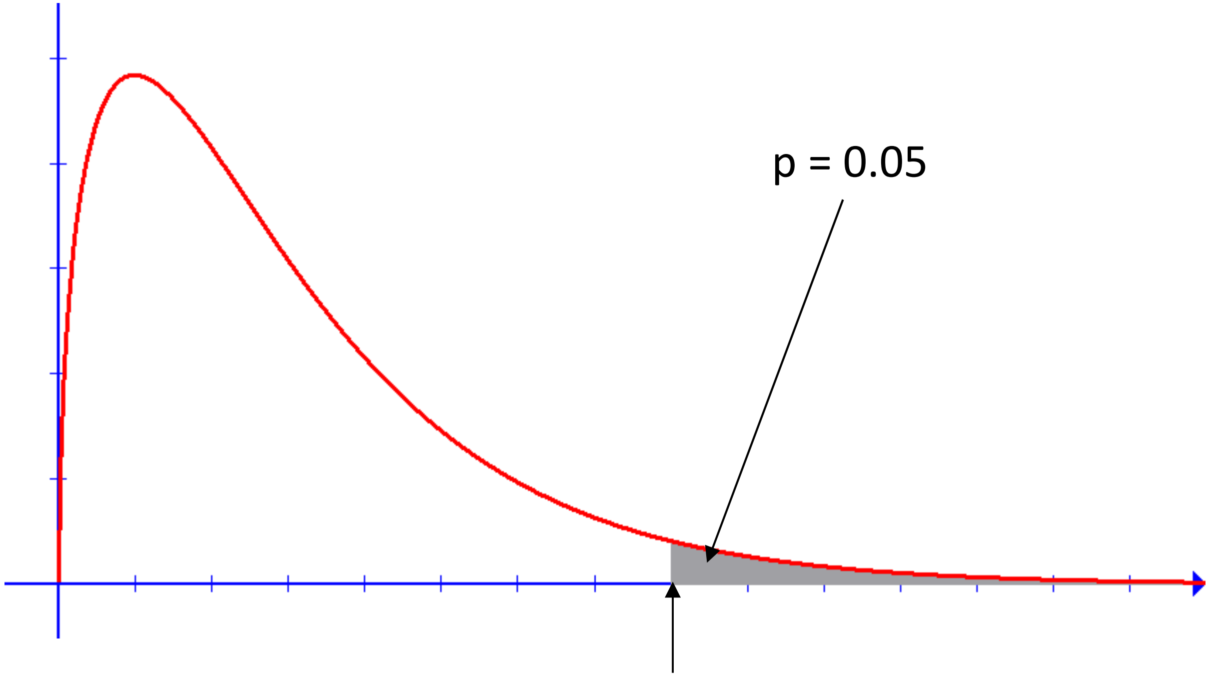
Table AIV.4 Chi-Square Probabilities



df\p	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	4×10^{-5}	16×10^{-5}	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	11.070	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

Complete Example continued – Towards the finish line

- So for $\alpha = 0.05$ and DoF = 2 (three categories of red, pink, white, hence DoF = 3–1 = 2) the χ^2 critical value is 5.991



Critical χ^2 = 5.991 for DoF = 2

at the 5% level of significance (= 95% confidence level)

Complete Example continued – Towards the finish line

- So our $\chi^2 = 0.234 < 5.991 =$ critical χ^2 value, hence we accept H_0 at the 5% level.
- So it seems that our experimental data agrees with the theoretical or expected ratio of 1:2:1 to within statistical limits, i.e. to within the natural variation exhibited by this biological process.
- What about the 10% level and the 1% level? See next slide.

Complete Example continued – Towards the finish line

Significance level	$\chi^2_{critical}$	Our χ^2 value	H_0
0.1 (10% level)	4.605	0.234	**Accept**
0.05 (5% level)	5.991	0.234	**Accept**
0.01 (1% level)	9.210	0.234	**Accept**

Complete Example continued – Towards the finish line

- So, even at the 0.1 level of significance (a very strict level for accepting/rejecting H_0) we still accept H_0.
- This implies that our ratio of 145 : 289 : 138 is well within acceptable variation of ratio for red : pink : white-flowered plants, …
- … and that the Mendelian genetic model of 1:2:1 is a good fit for this case.

Summary

- Calculate the Chi-squared value from your experimental data, and calculate the DoF.
- Look at the table of Chi-squared values (the Chi-squared distribution (i.e. the theoretical curve)) to see how common or rare your specific value is.
- If your value is extremely high (far out in the tail of the distribution), it is considered statistically significant, meaning your experimental results are unlikely to have happened by chance.

Constraints on using χ^2

- The χ^2-test is also used only when we have following type of data:
  - The data is not (and does not have to be) normally distributed;
  - Expected values are greater than 5. More on expected values later;
  - The data collected is frequency data, in other words data we count, i.e. "we have 11 of these and 15 of those".

Constraints on using χ^2

- *Frequency data* is a raw/absolute count of the number of objects in each category not a ratio, proportion or percentage.
- Ratios, proportions or percentages loose all the information about sample sizes:
  - Experimental data/frequencies of 6:4, 60:40, and 600:400 all have the same ratio or proportion of 0.6:0.4 and percentages 60%:40% but one is of sample size 10, the other 100, and the other 1000.

Constraints on using χ^2

- *Frequency data* is a raw/absolute count of the number of objects in each category not a ratio, proportion or percentage.
- Ratios, proportions or percentages loose all the information about sample sizes:
  - So if you are testing these proportions against an theoretical ratio of 1:1, the χ^2 test would give you the same value for all three sets of data.
  - But in terms of absolute size, 6:4 may or may not be more significant than 600:400.

Minitab

- *Minitab comment:*
  - The experimental data itself cannot be a ratio or proportion …
  - … but we can use χ^2 to test against a ratio or proportion (i.e. we can test against a theoretical distribution which has the ratio a:b:c)
  - We will see this in the Minitab part of the lesson.