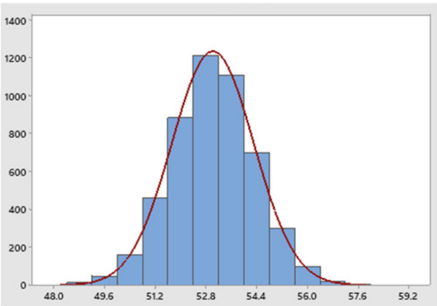
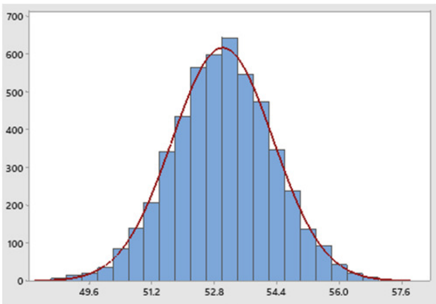
UPCSE Biology – Basic Statistic course

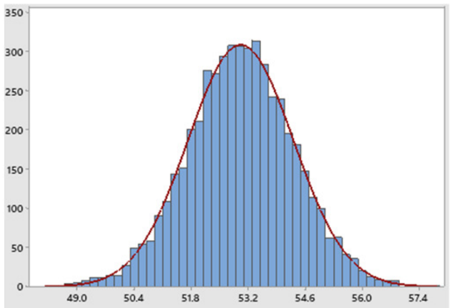
The z-test and the t-test

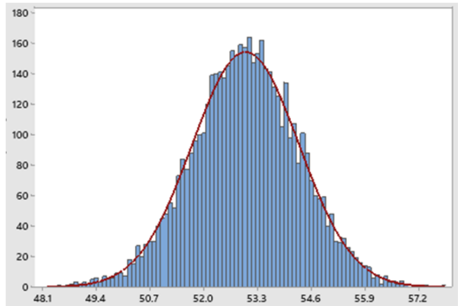
Revision

- In the last lesson we saw that the more data we collect for the distribution of heights of a population of people the more the distribution looks bell shaped.





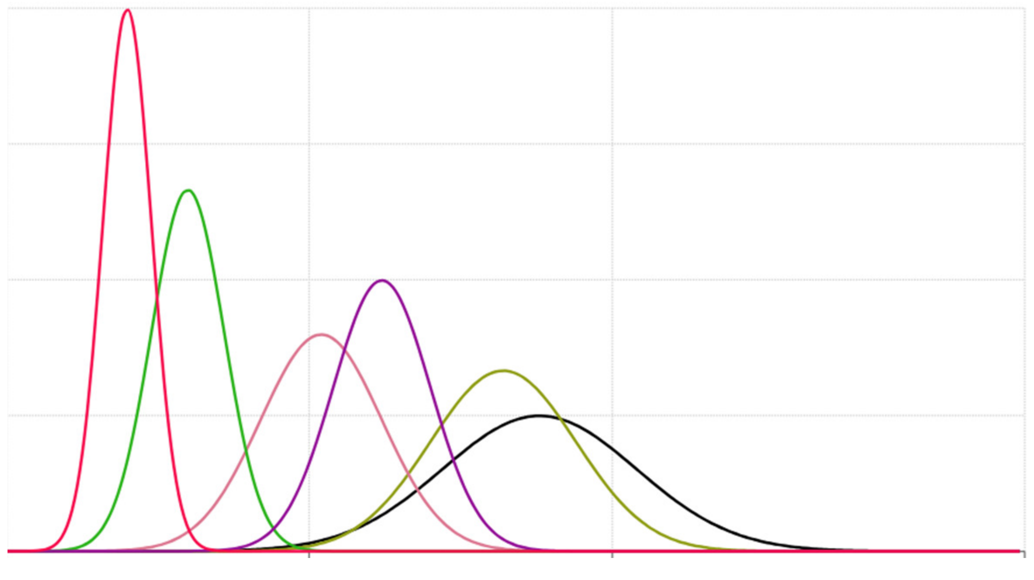




- Many things in life have this distribution such as a species of animal or plant.

Revision

- Since each species of animal or plant has it own bell shaped curve for the distribution of height (or any other property such as weight, volume, etc.)
- This means that there are thousand or tens of thousand of normal distribution curves.

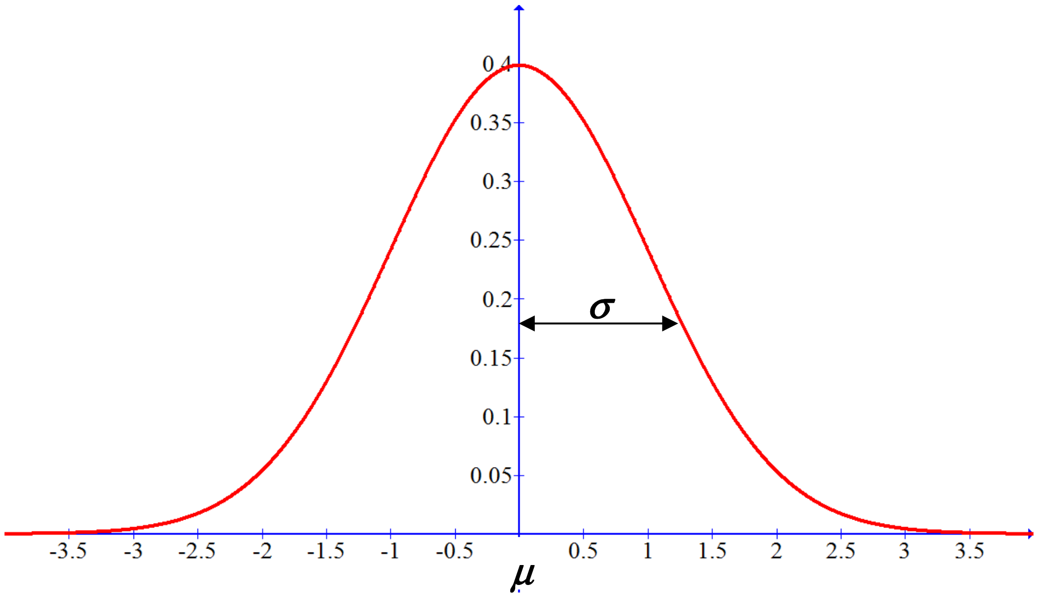


Revision

- We then saw that these thousands or tens of thousand of normal distribution curves can be reduced to one standard form via the z transformation:

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

(x represent the raw data value collected from experiments, μ is the population mean, standardised to 0, and σ is the standardised standard deviation)



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

Statistical testing

5) Consider a sample of people who choose to go on a diet:

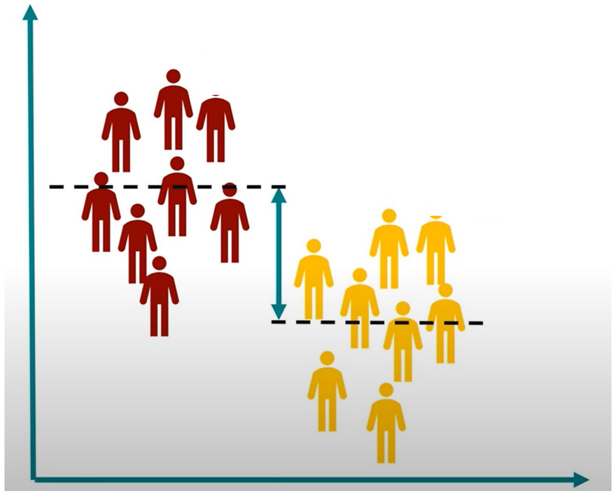
6) We want to know if their mean weight has changed significantly beyond what we usual expect an average person's mean weight to be.

A sample of people taken from a population

The group's mean weight is significantly different compared to average of the population

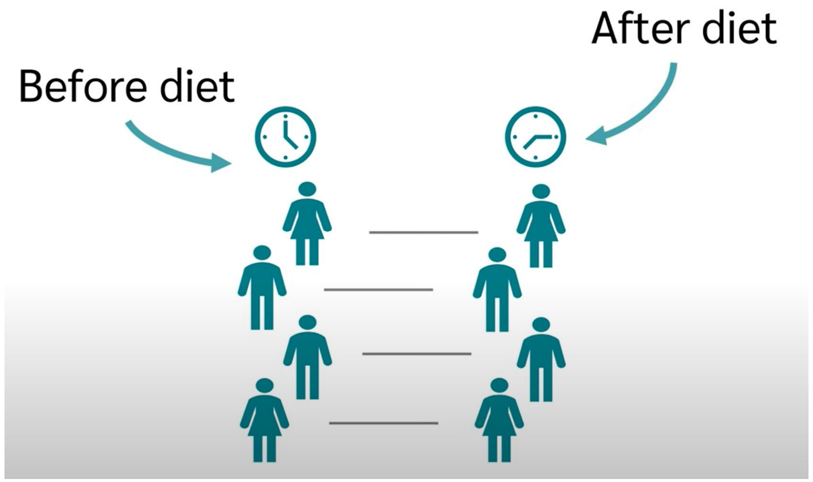
Statistical testing

7) Or we want to know if their mean weight has changed significantly beyond the mean weight of another group of people (control group)



The group's mean weight is significantly different to the control group

8) Or we want to know if their mean weight has changed significantly compared to before they went on a diet.



The group's mean weight is significantly different compared to before the diet

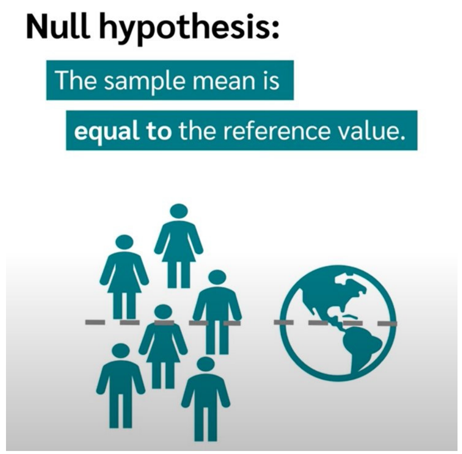
Statistical testing

So

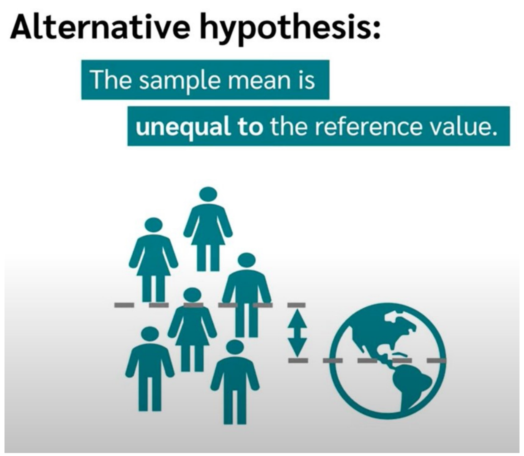
9) is the change in mean weight part of the usual change in weight which occurs naturally?

10) or is the change in mean weight highly unusual because some biological effect has actually occurred?

11) We frame these questions as statistical hypotheses: the null hypothesis H_0, and the alternative hypothesis H_1:



Difference is due to natural/sample variation



Difference is due to actual biological effect

Statistical testing

- *Another example:* Consider taking a sample from a population of birds. We conduct stats test to find out if a mean/average of that sample is what we would consider usual or unusual *in the statistical sense*.

- Specifically, if we are testing the mean weight of bird eggs for birds of species A we want to know if this mean weight has changed beyond what we usual expect the mean weight to be.

Statistical testing

- Specifically,
  - we take a sample of eggs from birds of type A,
  - we calculate the mean weight,
  - we then compare (by doing specific stats tests) if this mean weight is *statistically* different from the expected mean weight. Specifically this means
    a) is the change in mean weight part of the usual change in weight which occurs naturally for birds of type A?
    b) or is the change in mean weight highly unusual because some actual biological effect has occurred?

Asking questions about sample means

Example

- We then frame the last two questions as statistical hypotheses:
  - For a) above: There is no real difference in mean egg weight. The difference we have measured is due to sample variation. In other words, we picked a sample “over here” rather than “over there”.

    Our statistical hypothesis is then stated as

    Null hypothesis is H_0: The mean egg weight is statistically equal to the population mean.

Asking questions about sample means

Example

- There are only two reasons why such a difference could occur:
  - For b) above: There is a real difference in mean bird weight due to diet. In this case, further biological study would be needed.

  Our statistical hypothesis is then stated as

  Alternative hypothesis is H_1:

  The mean egg weight is statistically different from (greater and/or less than) the population mean.

Testing large samples: z-test

Example

- We want to study the mean weight of eggs from the red-tailed hawk.
- Information about the average weight and spread of weight for a population of 500 hawks is as follows:

$$\mu = 50\text{g} \qquad \sigma = 5\text{g}$$

Testing large samples: z-test

Example

- We suspect that the egg weight has changed due to a change in diet. We will test this by taking a sample of 10 hawks and record the following information:

$$\bar{x} = 59\text{g}$$

- Does this sample average suggests a significant change in mean weight or does this mean weight fall within the natural variation of egg weight?

Testing large samples: z-test

Example

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

Asking questions about sample means

Example

- One might think that all we have to do is find the difference in means: 59g – 50g = 9g.
- But how do we know if 9g is a small or large difference? We don't, which is why we need to calculate how much larger 9g is compare to the standard deviation/spread of 5g. Hence the z formula:

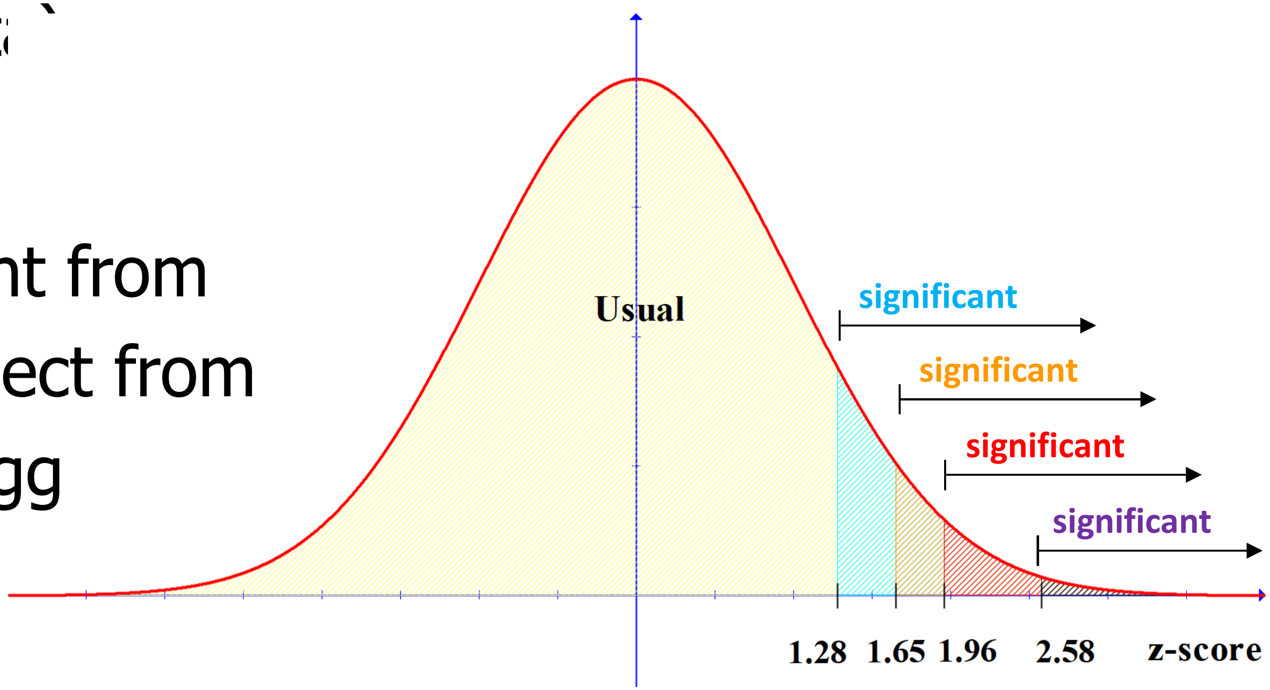
$$z = \frac{50 - 59}{5} = 1.8\text{g}$$

Asking questions about sample means

Example

We now compare this z-value against the standard normal distribution (here the z-values have replaced the x values of our original raw dat; `

Is a z value of 1.8g significantly different from what we would expect from usual variation in egg weight?

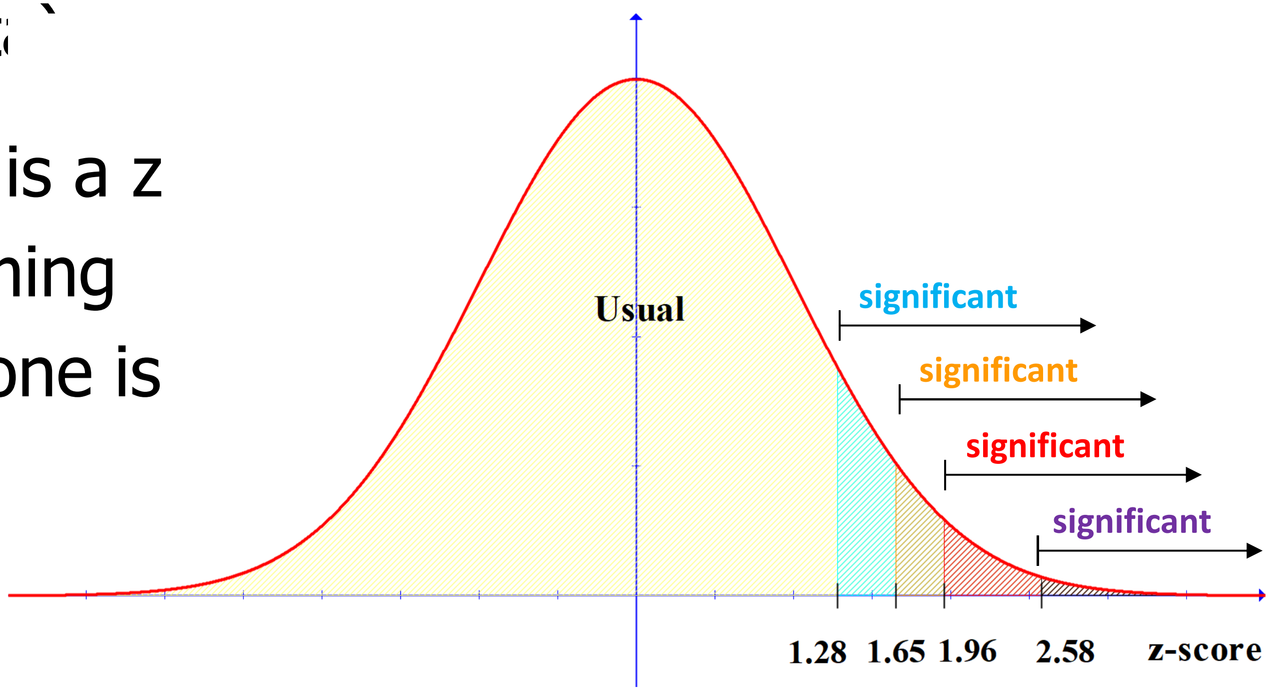


Asking questions about sample means

Example

We now compare this z-value against the standard normal distribution (here the z-values have replaced the x values of our original raw dat: `

Or: How surprising is a z value of 1.8g assuming chance variation alone is operating?



Asking questions about sample means

Example

- To answer this we have to look at the probability of getting $z = 1.8$. Probability values are found either in textbooks, online or via Minitab.
- We look at a table of probability values for z-scores, shown on the next slide.

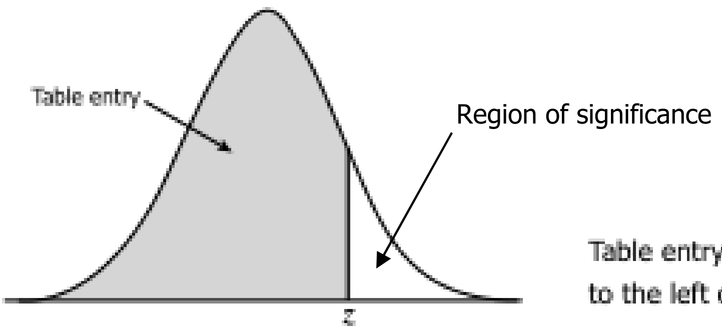


Table entry for z is the area under the standard normal curve to the left of z.

z = 1.8 – 1.80.

Look down the z column (left-most column) until we get to 1.8

Then look across to the **.00** column (for the second decimal place value).

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
……..					……..					……..
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

Then read off the number 0.9641.

Asking questions about sample means

- This means that there is a probability of 0.9641 that the mean sample weight of 59g occurred by chance.
- Or: there is a 96.41% chance of getting a result of 59g purely by random sampling.
- But in stats we don’t answer our hypotheses this way. Rather we answer it as the probability of our z-score being significant (see graph three slides up).
- In that case we have a probability of 0.0359 (=1 – 0.9641) that mean sample weight of 59g occurred not by chance but because of an effect.

P-values & significance levels

- The next question is, Is 0.0359 an acceptable probability? Do we want to accept of reject H_0 on the basis of this probability?
- And this is where things becomes problematic because the answer to this question is partly based on personal experience of the biological situation.
- The answer to this question relies on you, the biologist, choosing something called a level of significance, denoted α.

P-values & significance levels

- Levels of significance are reference probabilities which we use to compare our p-values against.
- These levels of significance are probability value we ourselves choose, with the stipulation that they act as critical values beyond which beyond which the p-value indicates a real biological effect (not due to chance).

P-values & significance levels

- To repeat: An α-level of significance is what we believe to be an acceptable probability, beyond which the p-value indicates a real biological effect.
- So you can choose $\alpha = 0.1$ (i.e. 10%), $\alpha = 0.05$ (i.e. 5%) or $\alpha = 0.01$ (i.e. 1%) as suits your purpose and the biology you are studying.
- Only an experienced biologist can determine the appropriate/suitable α-level given the biological phenomenon s/he is studying.

P-values & significance levels

- So, if we decide that we base our decision to accept or reject H_0 on a significance level of $\alpha = 0.05$ we are saying that …
  - any p-value smaller than 0.05 implies we reject H_0;
  - any p-value larger than 0.05 implies we accept H_0;

P-values & significance levels

- For example, if we have calculated z to be
  - z = 1.4 we get from tables that p=0.0808. this is a small probability but still 0.0808 > 0.05 so we accept H_0;
  - z = 2.3 we get from tables that p=0.0107. this is an even smaller probability. Now, since 0.0107 < 0.05 we do not accept H_0;

P-values & significance levels

- *On the other hand:*

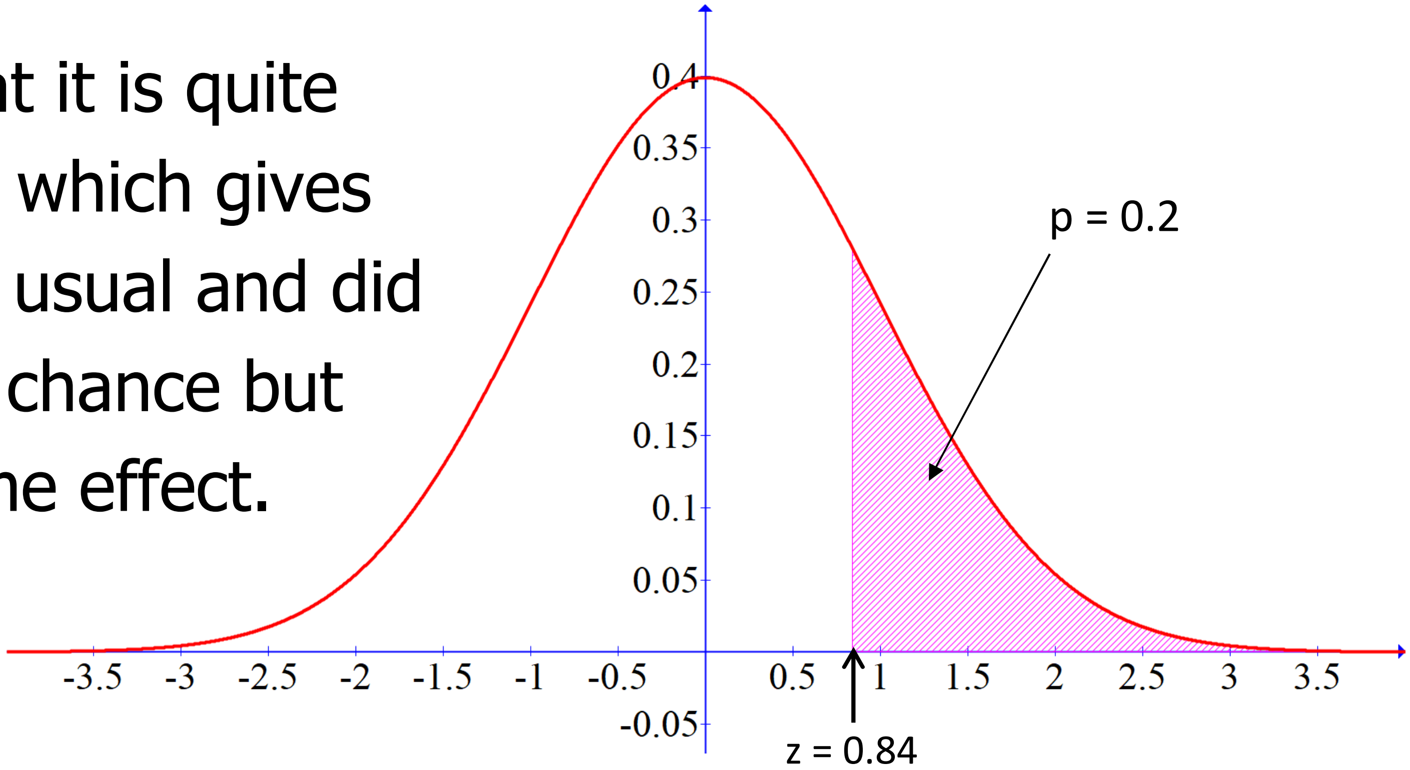
  If we choose a large value of α, say $\alpha = 0.3$, we could get away with a larger z value (in magnitude), say $z = 0.84$, resulting in a probability of 0.2 (from tables), and still end up rejecting H_0 when it is much more likely that we should have accepted H_0.

  This is shown in the graphs below.

P-values & significance levels

p=0.2 (the pink area) takes up a lot of the region under the curve. In the context of $\alpha = 0.3$ this is the significance region.

This means that it is quite likely that data which gives z > 0.84 is not usual and did not happen by chance but because of some effect.



P-values & significance levels

- So, be careful not to set an α-level too high or too low. This may make you accept/reject H_0 when in fact you would have rejected/accepted H_0 using a more sensible cut-off value for α.
- **Crucial point:** The α-level you choose is a reflection of your personal decision about how strict you wish to be in knowing H_0 is valid. Any p-value beyond your α-level means you may unfairly reject H_0 when it would have been better to accept it.

P-values & significance levels

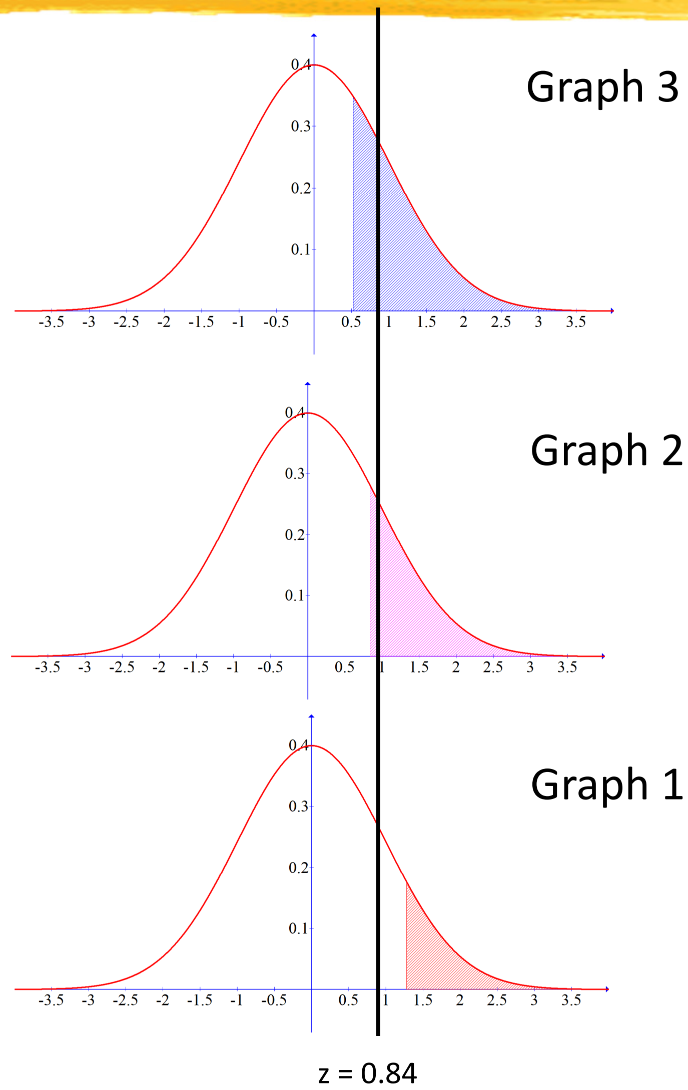
The black vertical line defines z = 0.84
and the pink area of graph 2 is p = 0.2

The orange region of graph 1 has area p = 0.1
and relates to a z = 1.28;

The blue region of graph 3 has area p = 0.3
and relates to a z = 0.525;

So, if we have actually calculated a value z = 0.84 and

- we have set a significance level of $\alpha = 0.3$, we reject H_0 because z lies in the blue region;
- we have set a significance level of $\alpha = 0.2$, we reject H_0 because z lies in the pink region (just);
- we have set a significance level of $\alpha = 0.1$, we accept H_0 because z does not lie in the orange region.



P-values & significance levels

- The question then becomes:

What is the most suitable α-level to choose
in order to minimise the risk of accepting or rejecting H_0
when in fact we should have rejected or accepted H_0?

- How do you overcome the problem of unfairly rejecting (or accepting) H_0?
- The answer is: Know your biology. You are the specialists in the specific biological situation you are studying. Never give up your biology experience to the calculation!

P-values & significance levels

- All of this can be summarised as follows:

The level of significance is the level you personally wish to set as the probability of obtaining your results just by chance.

So, how much chance (probability) are you prepared to accept as chance/luck before you accept that something is really happening?

- The smaller the p-value the more unusual your results. What α-level do you want to set (to compare against the p-value) before you consider your results unusual?

Asking questions about sample means

Returning to our example:

- z = 1.80 gives p = 0.0359
- Setting $\alpha = 0.1$ we see that p lies well into the significant region, so we reject H_0;
- Setting $\alpha = 0.05$ we see that p lies in the significant region, so we reject H_0;
- Setting $\alpha = 0.01$ we see that p lies in the "usual" region, so we accept H_0.

The t-test (for small samples)

- Any test involving the z-score formula only applies to large amounts of data. The z formula only applies for large sample size, n >= 30, but usually much bigger than this.
- But what if we don't have such an amount of data? What if we only have a sample of size n=5 or n=10 or n=15?
- Here we will look at the t-test, which is a variation on the standard normal distribution for small samples.

The t-test (for small samples)

- Then we use a different formula, given by

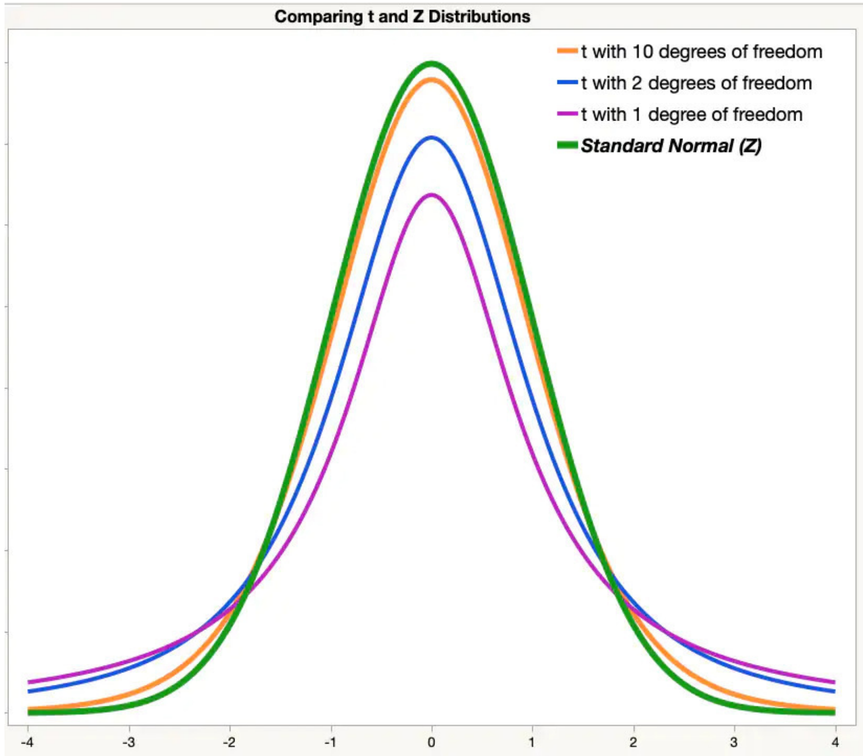
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

where $\bar{x}$ is the sample mean, μ is a known reference mean we are measuring against, s is the sample standard deviation, n = sample size.

- *Optional:* See separate Word document for where this formula comes from.

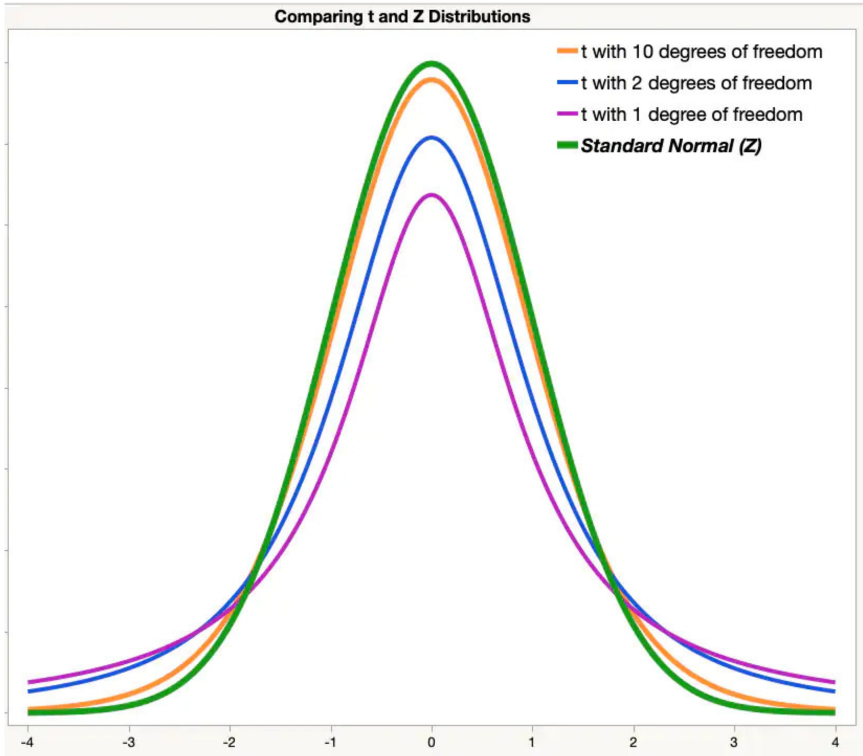
The t-distribution and degrees of freedom

- The standard normal distribution is a distribution of z-scores
- What does the distribution of these t-scores look like? It depends on the sample size. See the graph on the right.



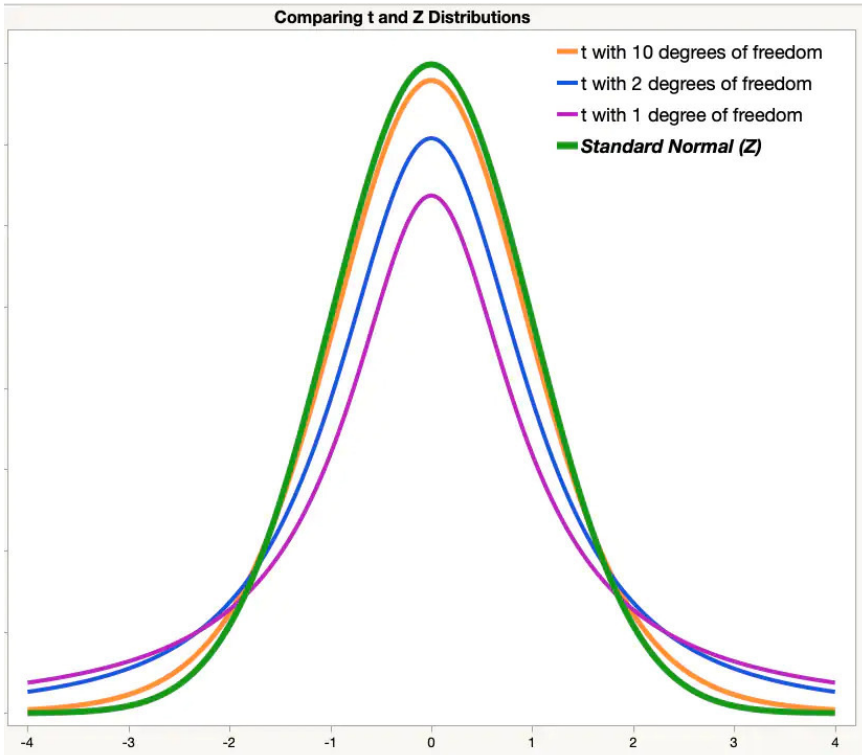
The t-distribution and degrees of freedom

- There is one t-distribution curve for n = 2, one t-distribution curve for n = 3, one t-distribution curve for n = 4, etc.
- For technical reasons these curves are based on a concept called *degrees of freedom* given by n – 1. See Word document for what degrees of freedom means.



The t-distribution and degrees of freedom

- Note that the larger the sample size the closer the t-distribution curve approaches the standard normal (z) curve.



Three basic types of t-tests

- There are then 3 types of t-test we can perform:
  - One-sample t-test: A hospital wants to know if the mean cholesterol level $\bar{x}$ of a sample of patients is different from the reference mean level μ;
  - Two-sample dependent t-test (i.e a paired t-test) in a before-and-after situation: An ecology group wants to test whether or not there is a difference in mean CO_2 levels in a river before and after a storm has occurred;

Three basic types of t-tests

- There are then 3 types of t-test we can perform:
  - Two-sample independent t-test: A hospital wants to measure the average effect of a new drug between two groups, a control group and a placebo group, to see if the drug is effective;
- On the UPCSE biology course you will only need to learn about the two-sample dependent t-test, otherwise known as the paired t-test.

Asking questions about sample means: Paired t-test

Example

- An ecology group wants to test whether or not there is a difference in mean CO_2 levels in a river before and after a storm has occurred.
- Two random samples of 10 readings each were collected from the same section of river, before and after the storm.
- Data is shown on the next slide.

Asking questions about sample means

Observation	CO2 level (Before)	CO2 level (After)
1	640	657.2
2	655	672.2
3	662	665.2
4	648	679.2
5	670	652.2
6	659	668.2
7	645	676.2
8	668	662.2
9	652	681.2
10	661	666.2

CO2 concentrations in μmol/kg

Asking questions about sample means: Paired t-test

Example

- We want to test for any increase in CO_2 levels after the storm compared to before.
- Since we are dealing with a small sample, and the same situation (same section of river) but simply a change in circumstance (before and after the storm), this qualifies as a paired t-test case.
- Then, to perform a two-sample paired t-test we find the difference between each before-and-after observation.

Asking questions about sample means

This gives us the following table, where CO_2 concentration is in μmol/kg.

Observation	CO2 level (Before)	CO2 level (After)	Difference (After – Before)
1	640	657.2	17.2
2	655	672.2	17.2
3	662	665.2	3.2
4	648	679.2	31.2
5	670	652.2	−17.8
6	659	668.2	9.2
7	645	676.2	31.2
8	668	662.2	−5.8
9	652	681.2	29.2
10	661	666.2	5.2

Asking questions about sample means

Example

- We now test the difference instead of each sample separately.

- The t-formula for testing a difference in sample means is a slight variation on the one shown previously:

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

where $\bar{d}$ is the difference between the sample means, i.e. $\bar{d} = \bar{x}_{after} - \bar{x}_{before}$ (instead of the difference between sample mean and population mean).

Hypothesis statement

- In this case our hypotheses will be
  - null hypothesis H_0: the mean CO_2 level is statistically the same as the standard CO_2 level
    - The difference in mean CO_2 concentration between our sample and the population value is due only to sample variation.
  - alternative hypothesis H_1: the mean CO_2 level is statistically greater than the standard CO_2 level.
    - The difference in mean CO_2 concentration between our sample and the population value is due to a real biological effect.

Hypothesis statement

- Note that there are three options to an alternative hypothesis. The alternative hypothesis H_1 can either be
  - the mean CO_2 level is statistically greater than the standard CO_2 level → one-sided t-test.
  - the mean CO_2 level is statistically less than the standard CO_2 level → one-sided t-test.
  - the mean CO_2 level is statistically different than (i.e. less than or greater than) the standard CO_2 level → two-sided t-test.

Asking questions about sample means

Example: continued

- For our problem we have
  - Sample size: $n = 10$;
  - Mean difference: $\bar{d} = 12$;
  - Standard deviation of differences: $s_d \approx 16.41$;
  - Change is possibly "greater than".

- **Question:** Is the mean difference large or small?
  Is this difference due to natural variation in the CO_2 concentration or choice of sample …

Asking questions about sample means

Example: continued

- **Question:** … or is this difference due to some real effect biological effect which is affecting the river's CO_2 concentration?
- **Answer:** Since n = 10 we perform a t-test (not a z-test).

Example: Continued

- So,

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{12.0}{5.19} \approx 2.31$$

with degrees of freedom: 10 – 1 = 9 (tells us which t-curve to use)

- This *t* value is a measure of how different the sample means are.
- The question then is, When does the difference become so large that it is not due to sampling variation, but to actual ecological effect?

P-values & significance levels

- Statistically speaking we are asking:

  "Is this *t* value significant?"

- In order to answer this question we now need to set a significance level α against which to compare the resulting p-value related to this t-value.
- This will help us decide if we accept or reject H_0.
- Then,
  a) for problems where H_1 states "there is a change in …", the *larger or smaller* the value of t compared to the critical value the less likely H_0 is true;

P-values & significance levels

- Then,

  b) for problems where H_1 states "there is an increase in …" the *larger* the value of t compared to the critical value the less likely H_0 is true;

  c) for problems where H_1 states "there is an decrease in …" the *smaller* the value of t compared to the critical value the less likely H_0 is true;

- We have situation b).

P-values & significance levels

- The question then becomes

What is the most appropriate
α-value to choose in order to
minimise the risk of accepting/rejecting H_0
when in fact we should have rejected/accepted H_0?

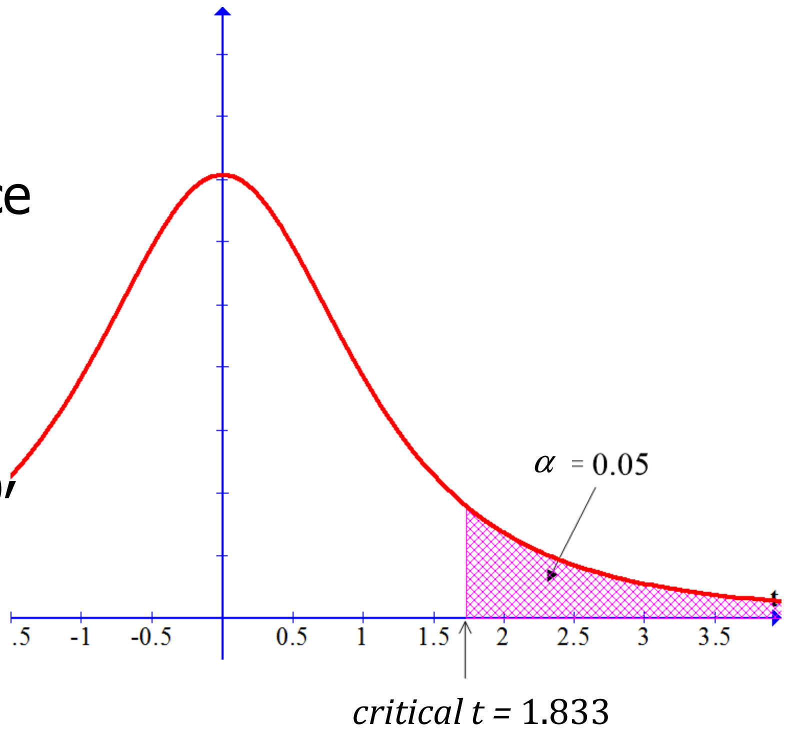
P-values & significance levels

For a significance level of 5%

- If I want to be confident that my data lies within the 5% range of data values in the right hand tail
  - Because I believe that data in this range will imply a biological significance

  then my t value …
  - which represents the cut-off point between accepting H_0 and rejecting H_0,

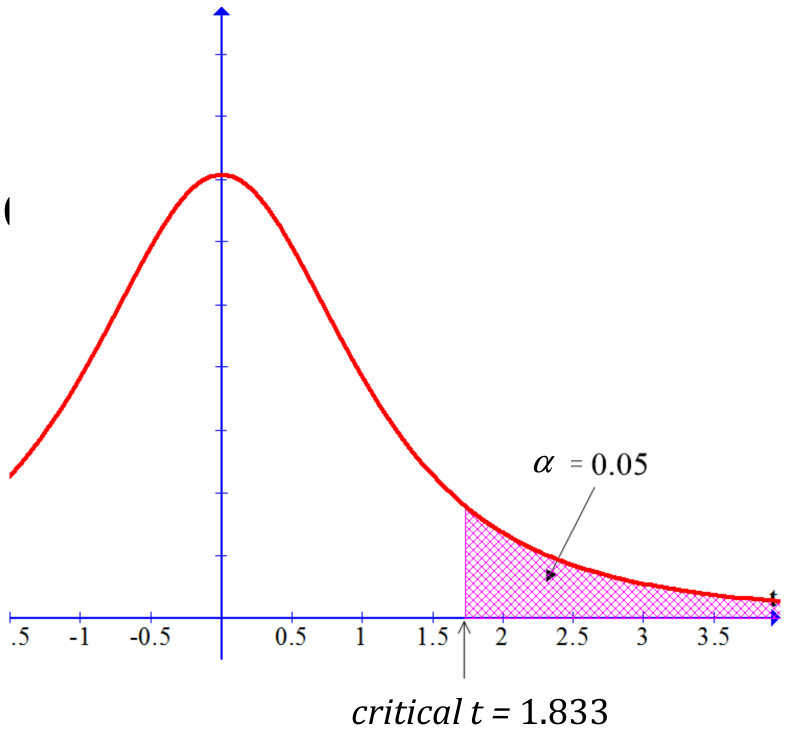
  … will have to be at least 1.833



P-values & significance levels

For a significance level of 5%

- Data lying in this region would then be considered a rare occurrence if it were to occur naturally (i.e. as relating to the H_0 hypothesis that such data is just natural variation and therefore statistically insignificant)
- It would therefore seem reasonable to interpret such data as being statistically significant, therefore confirming H_1.



Example 2: Continued

- So this is where we are at in our problem:

	Level of significance		
	10% (0.1)	**5% (0.05)**	**1% (0.01)**
t **= 2.31** **DoF = 10–1=9**	0.1 implies critical t = ??	0.05 implies critical t = ??	0.01 implies critical t = ??
Accept/reject H_0	?	?	?

Table of *t* -scores

<u>For our data</u>

t = 2.31

DoF = 9

α = 0.1, 0.05, 0.01

Table A-3 *t* Distribution: Critical *t* Values

	p values				
Degrees of Freedom	0.005	0.01	0.025	0.05	0.10
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325

Example: Continued

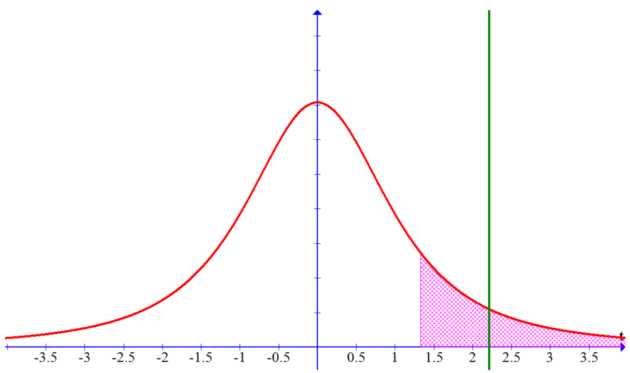
- Our *t* value is *t* = 2.217
- **Q:** Does our *t* value imply a (statistically) significant difference between the means before and after the storm?

	Level of significance		
	10% (0.1)	**5% (0.05)**	**1% (0.01)**
***t* = 2.31** **DoF = 9**	0.1 implies t = 1.383	0.05 implies t = 1.833	0.01 implies t = 2.821
Accept/reject H_0	Reject H_0	Reject H_0	Accept H_0

Example: Continued

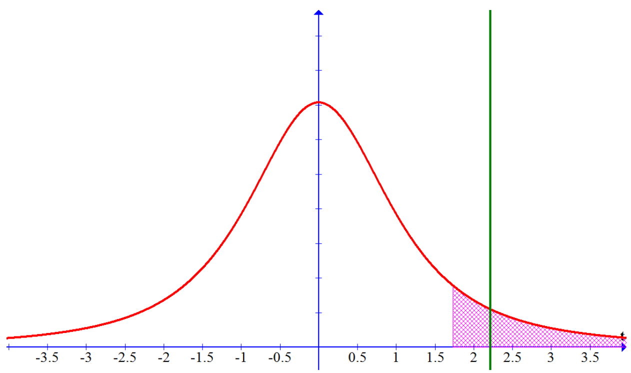
- Our t value is $t = 2.31$ (vertical green line in graphs below)
- **Q:** Does our t value imply a (statistically) significant difference in the means before and after the storm?

A: Yes



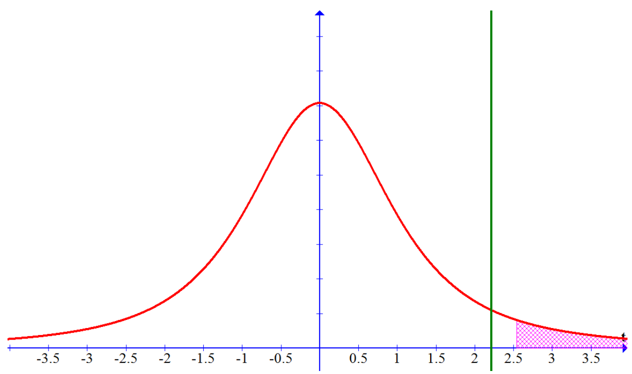
10% level

A: Yes



5% level

A: No

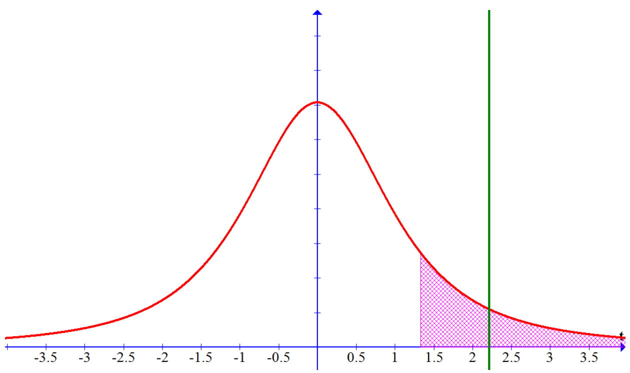


1% level

Example: Continued

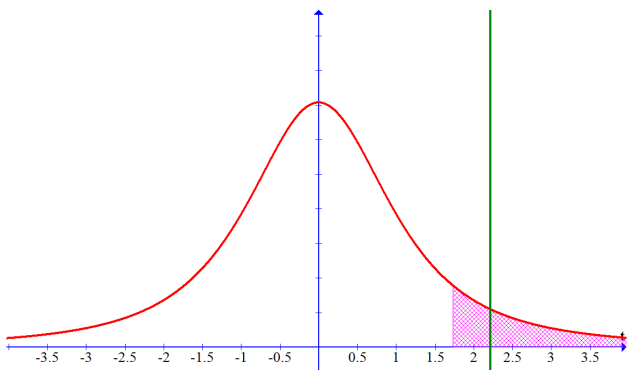
- So depending upon which significance level we use we end up either accepting or rejecting that there is a significant difference in CO_2 concentration in that section of river.

A: Yes



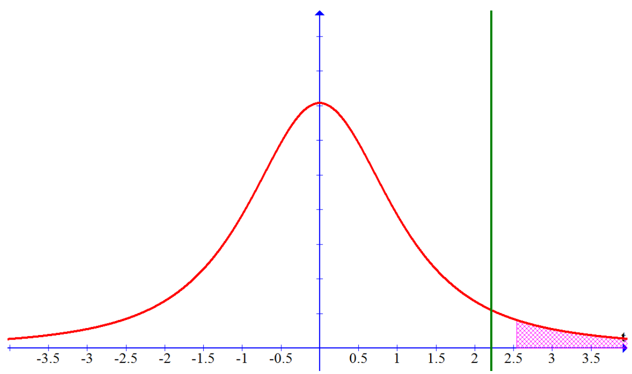
10% level

A: Yes



5% level

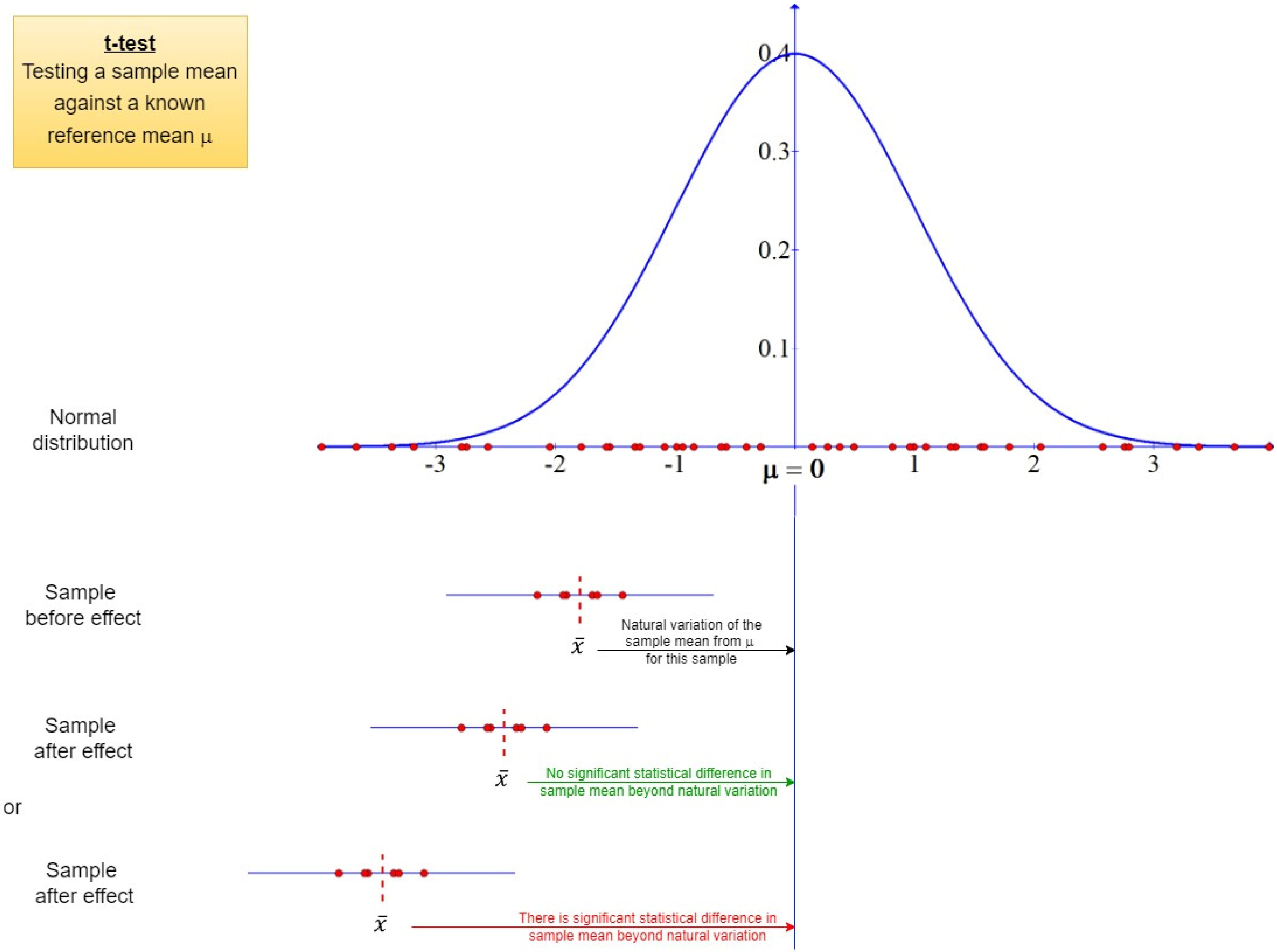
A: No



1% level

Summary

- 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.


t-test
Testing a sample mean against a known reference mean μ
Normal distribution
0.4
0.3
0.2
0.1
-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

Comment

Question:

- Given that we accept H_0 at the 1% levels of significance,
  - are we being too strict in setting a *p*-value of 0.01 for rejecting H_0?
  - are we accepting H_0 unfairly if we use the 1% level of significance?

Comment

- So, the level of significance, α, is a measure of how incredible it is for H_0 to occur naturally or just by chance.
- The lower the α-level the smaller the significance region under the t-curve (or any other curve used to model your situation), therefore more incredible it is for you to have got a result which lies in this small region.

Comment

- In other words, the more incredible it is for H_0 to have occurred by natural variation alone (although this is still possible).

- Adopting an α-level allows us to be justified in rejecting H0 as if it were false "because it has low […] credibility. Hypotheses with low […] credibility are, presumably, rarely true".

(p214, Stephen Spielman, "The Logic of Tests of Significance", *Philosophy of Science*, Sep., 1974, Vol. 41, No. 3, pp. 211-226).

Comment

- But low probability (which is what low credibility implies) does not mean 0 probability. It is still probable that an extreme t-value could occur by chance or natural variation.
- So if we set an that α-level of 0.05 we are saying that any t-value which lies in the region less than 0.05 implies that the chance of us getting that t-value is at most 1 in 20 experiments.

Comment

- So that, 1 in every 20 samples you collect will give a t-value which lies in the non-significance (H_0) region, and therefore that you should have accepted H_0 when in fact, you rejected it.
- Are you prepared to accept this mistake in deciding upon H_0? Are you prepared to accept making a false-positive claim 1 in every 20 tests you do?
- If so, use $\alpha = 0.05$. If not use $\alpha = 0.01$ (or even less). Then you will only make a false-positive in 1-in-a-100 tests.

Know your biology

- So, in order to be confident in you accepting or rejecting H_0 you need (amongst other things) to look at the biology of the situation, and not simply rely on the statistics.

- It is your experience as biologists about the phenomenon, situation, ecology, physiology, etc., of the situation which will help you determine an appropriate significance level, and whether or not it makes sense to accept or reject H_0.

Know your biology

- For example, the physiology of a rabbit may be very different to that of a human being.
- Would a significance level of 0.05 be acceptable for identifying a significant difference in heart rates before-and-after a stressful event,
  - for a human being?
  - for a rabbit?

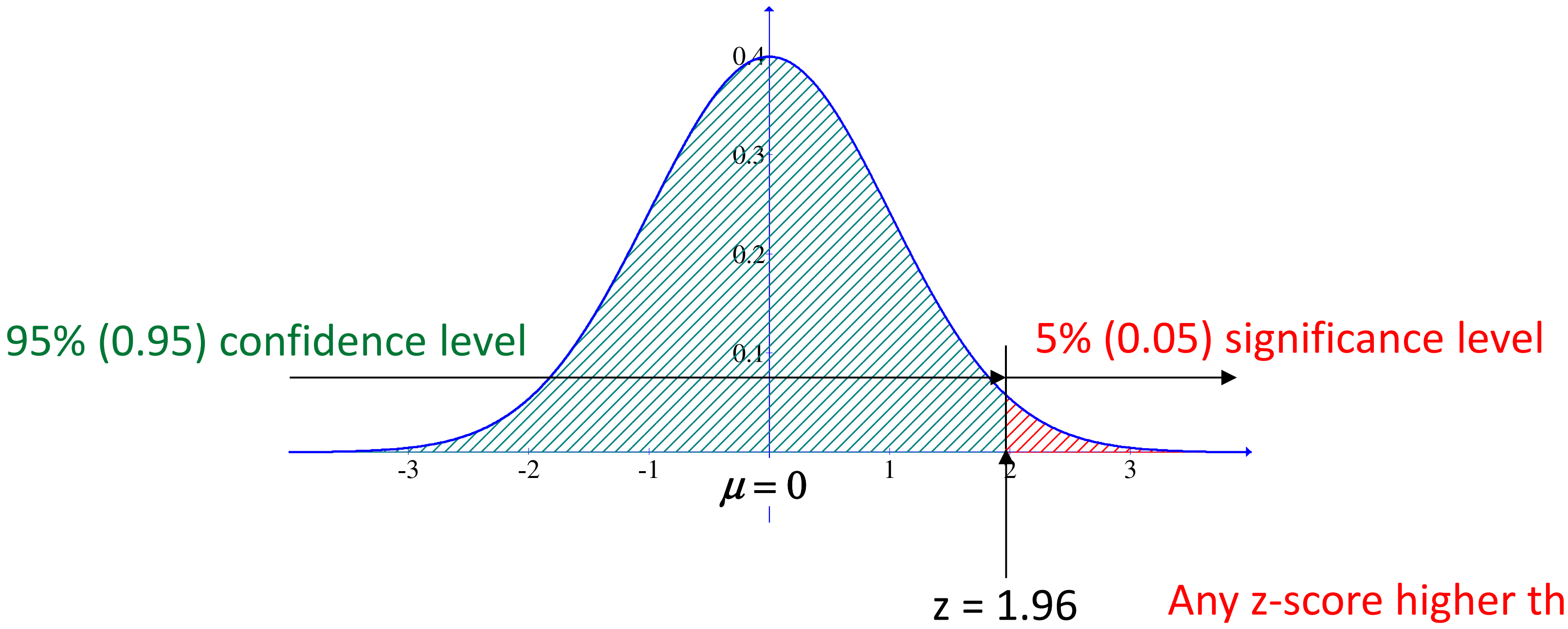
Know your biology

- Maybe, maybe not, since rabbits have a resting heart rate of 200-300bpm whilst for humans it is 60-100bpm.
- So if a human being had a heart rate of 200bpm then something has happened!
- If a rabbit has a heart rate of 200bpm then nothing has happened!

Distinction between significance levels and confidence levels

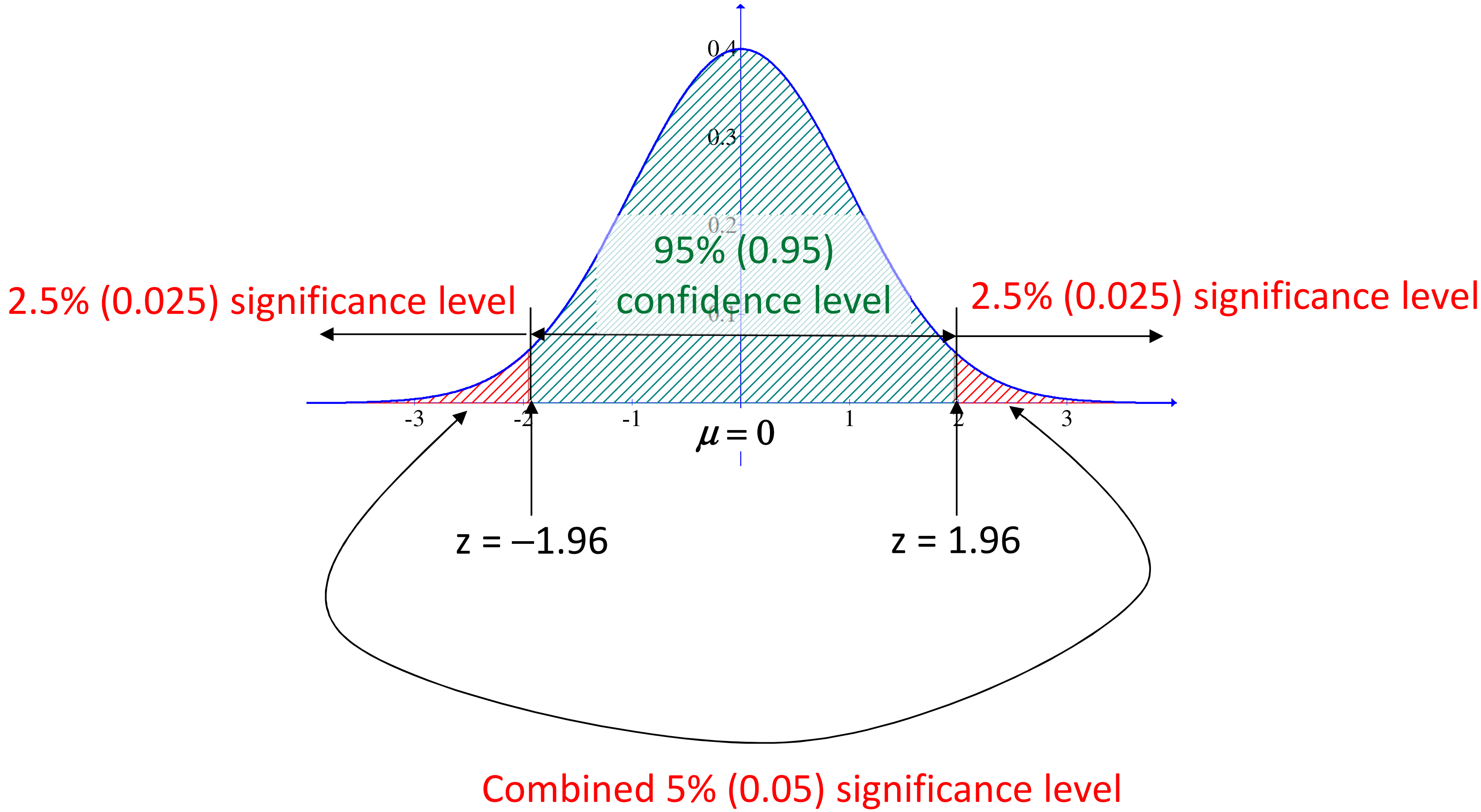
- Note that some books don't talk of significance levels but of confidence levels.
- The distinction between these two concepts is that a 0.05 (5%) level of significance is equal to a 0.95 (95%) confidence level. This is illustrated in the next slide.
- **Note:** Minitab use the idea of confidence levels not significance levels.

Distinction between significance levels and confidence levels



Any z-score higher than 1.96 from the z-formula (i.e. any p value lower than 0.05) suggests statistical significance, hence do not accept H_0

Distinction between significance levels and confidence levels



On justifying the acceptance or rejection of H_0

This, and subsequent slides, is a comment about how valid it is for you to accept or reject H_0.

- Consider throwing a die four times.
- The probability of getting fours 6s is 1/1296 = 0.000772.

On justifying the acceptance or rejection of H_0

- In this experiment there are no other effects which could impact the result of obtaining or not obtaining four 6s in a row.
- The biasedness of the die is the only effect which affects the result, and since we always make sure to set up non-biased experiments (as best we can) in statistics we can factor this out of our consideration when it comes to looking for effects which influence the outcome of our experiment (note that lack of bias does not mean lack of othr confounding effects).

On justifying the acceptance or rejection of H_0

- The only possible effect which could influence (in a biased sense) the outcome of getting four 6s in a row might be wind resistance?
- But even then, this only takes effect at high speeds, and throwing a die with your hand nowhere near produces such speeds.
- Another effect which could cause biasedness is wear-and-tear on the die.

On justifying the acceptance or rejection of H_0

- But if the die is sturdy and the surface you are throwing it on is soft and firm, then wear will hardly occur, if at all.
- However, in biology there are so many different effects which may influence the outcome of your experiment (apart from the effect you yourself are actually studying) that you don’t know what influence these other effects will have on the individual property you are studying.

On justifying the acceptance or rejection of H_0

- For example, in the example of CO2 concentrations in a river, other effects which might cause concentration to change significantly might be
  - the effect of the tides of the sea, and how this feeds trough to the river,
  - change in temperature of the river,
  - gas exchange between the river and the atmosphere,
  - change in mineral properties of the river,
  - change in acidity or alkalinity, etc.

On justifying the acceptance or rejection of H_0

- So the question is, Can you really separate out the different effects in order to analyse the one effect you are studying?
- Are the effect so interlinked that they cannot be separated out?
- This is for you to know as subject specialists.