Making statistics accessible to *non-maths* students

Challenging conventions

Dr. Christine Pereira
Christine.Pereira2@brunel.ac.uk

Brunel University London
BEEC Brunel Educational Excellence Centre
ASK ACADEMIC SKILLS

Just imagine...

Just imagine...

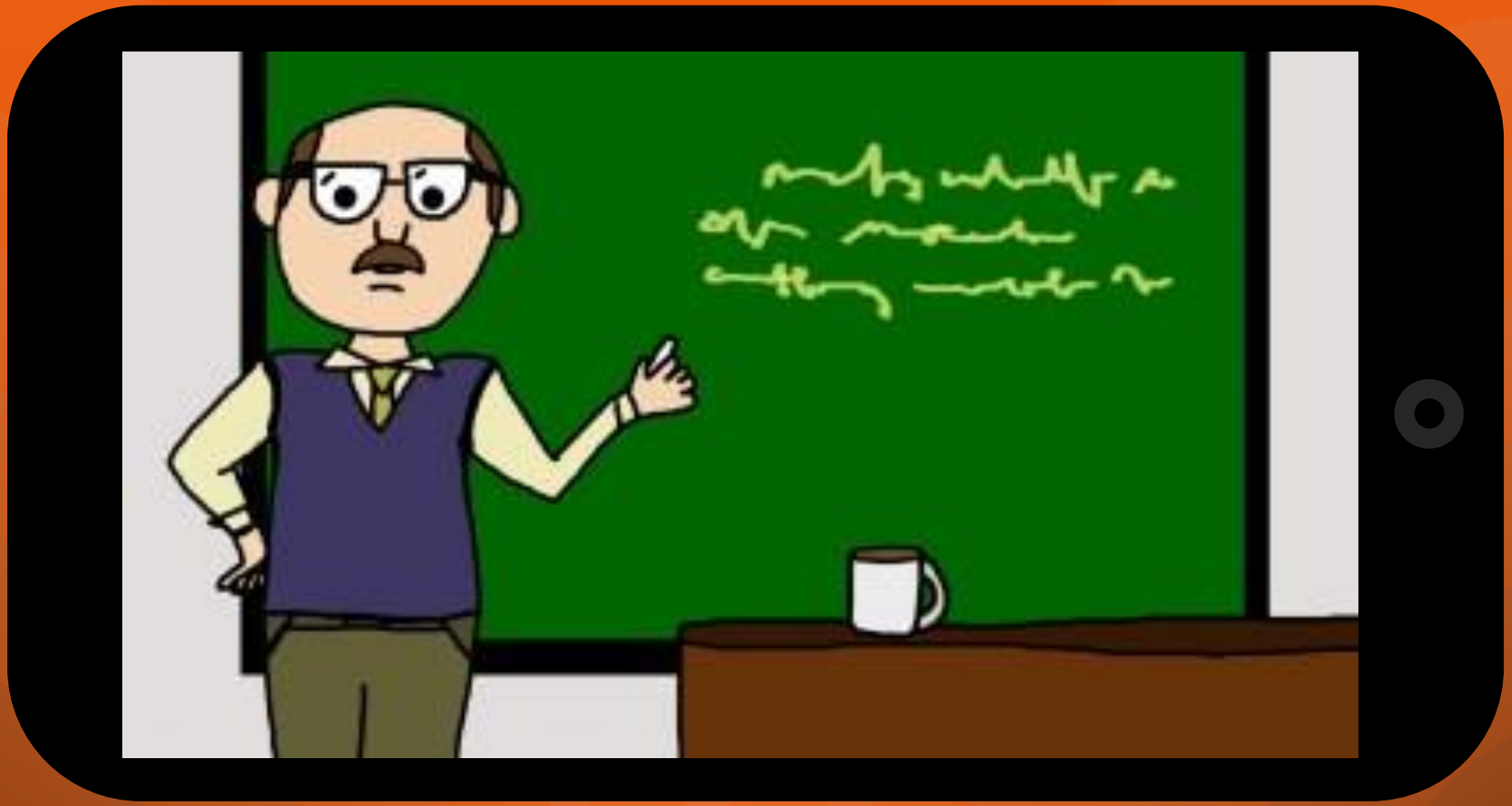
Course Guide

Modules
~~~~~~~
~~~~~~~
~~~~~~~

LFS1234

~~~~~~~
~~~~~~~
~~~~~~~

Just imagine...



Reconsidering what we teach

- Tackling the fear of statistics
- Needs based content – is it all really necessary?
- Concept based rather than formula based

As lecturers...

Assumptions?

Confidence?

Motivating?

Responsibility?

Tackling Fear

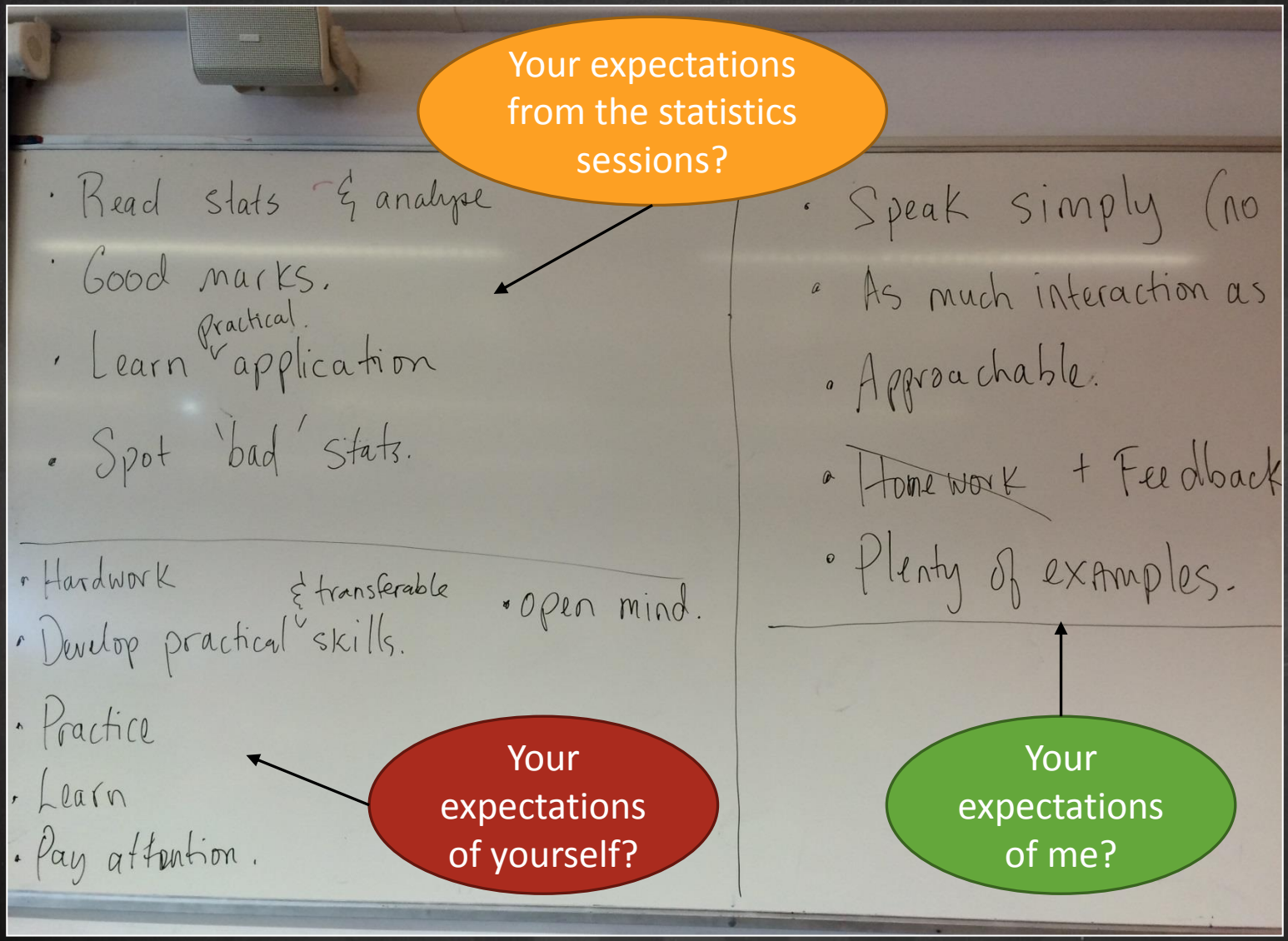
Talk to students about...

Intentions

Expectations

Past experiences

Expectations?



Expectations?

- Effort (unfortunately it will take some).
- Try the exercises.
- Give statistics a chance!
- Don't give up.
- Ask questions - I can't answer your questions if you don't ask them.
- Don't suffer in silence - if you don't get it ask someone. You can email me or post a question on HowCloud.

My expectations of you?

Needs based content

Essential vs Extra

Application

Statistical skills

Necessary?

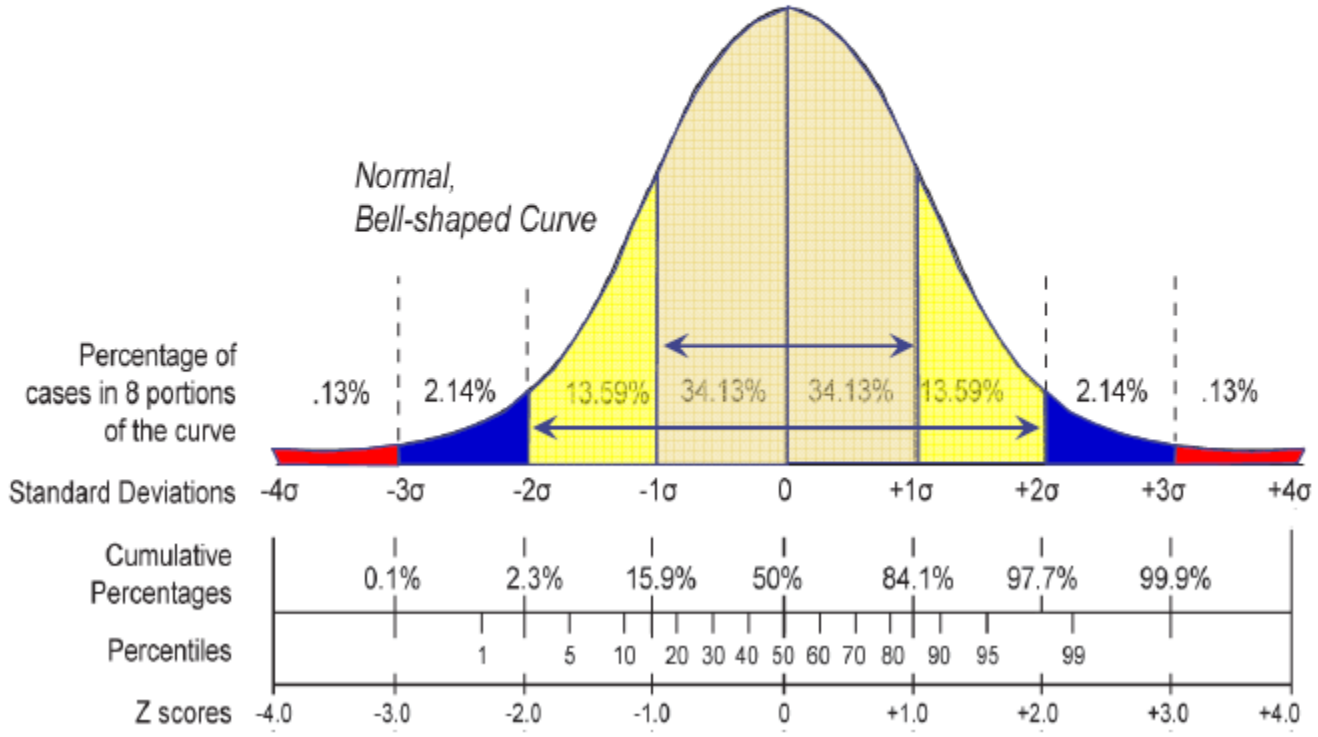
If X is a normal random variable with a mean μ and standard deviation σ, then the equation of the normal curve is

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Necessary?

Empirical rule: For a bell shaped distribution

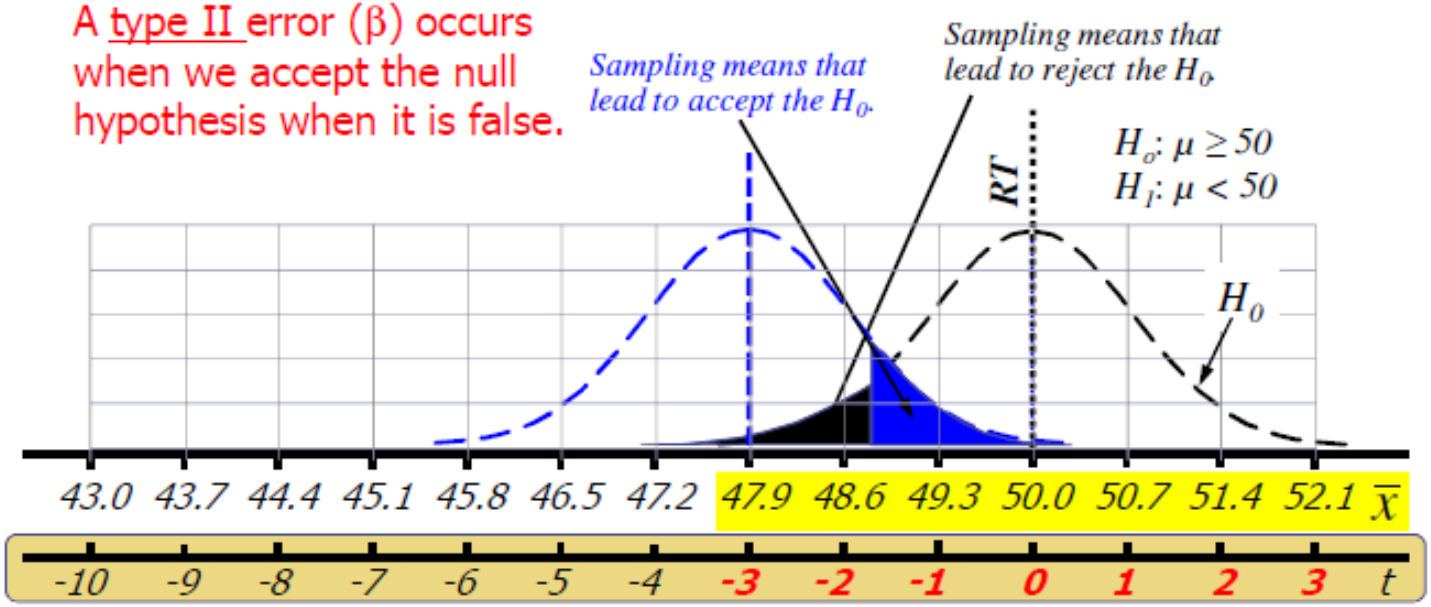
- 68.3% of the data fall between $\mu \pm 1\sigma$
- 95.4% of the data fall between $\mu \pm 2\sigma$
- 99.7% of the data fall between $\mu \pm 3\sigma$



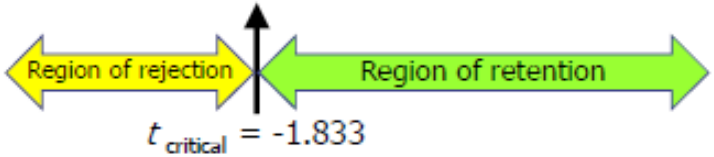
Necessary?

Type II Error - *Back to the example*

A type II error (β) occurs when we accept the null hypothesis when it is false.



Suppose our calculated sample mean is 47.9 (t = -3.0), the null hypothesis is still false but values > 48.8 would lead us to accept the null hypothesis (now we can make a type II error; P(β > 0); P(1-β < 100%).



Type II error for testing μ ≥ 50.0 against μ < 50.0

Essential -> Application -> Skills

1. Estimating percentages
   - Using a sample
   - Using sample statistics and the Normal Distribution
2. What happens to sample statistics when more data is collected (e.g. how does central tendency change?)
3. Meeting regulatory thresholds using percentiles
4. Confidence intervals
   - For observations
   - For the mean
5. The effect of N and SD on SE and confidence intervals
6. One-sample Z test

Teaching Approach

Concepts or Formulas?

Formula approach

$$\text{Z score} = Z = \frac{X - \bar{X}}{S}$$

Finding Z for the sampling distribution of the mean

$$Z = \frac{\left(\overline{X} - \mu_{\overline{X}}\right)}{\sigma_{\overline{X}}} = \frac{\left(\overline{X} - \mu\right)}{\frac{\sigma}{\sqrt{n}}}$$

Finding Z for the sampling distribution of the proportion

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

Formula approach

Z test for the mean (σ known)

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

t test for the mean (σ unknown)

$$t_{STAT} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

Z test for the proportion

$$Z_{STAT} = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

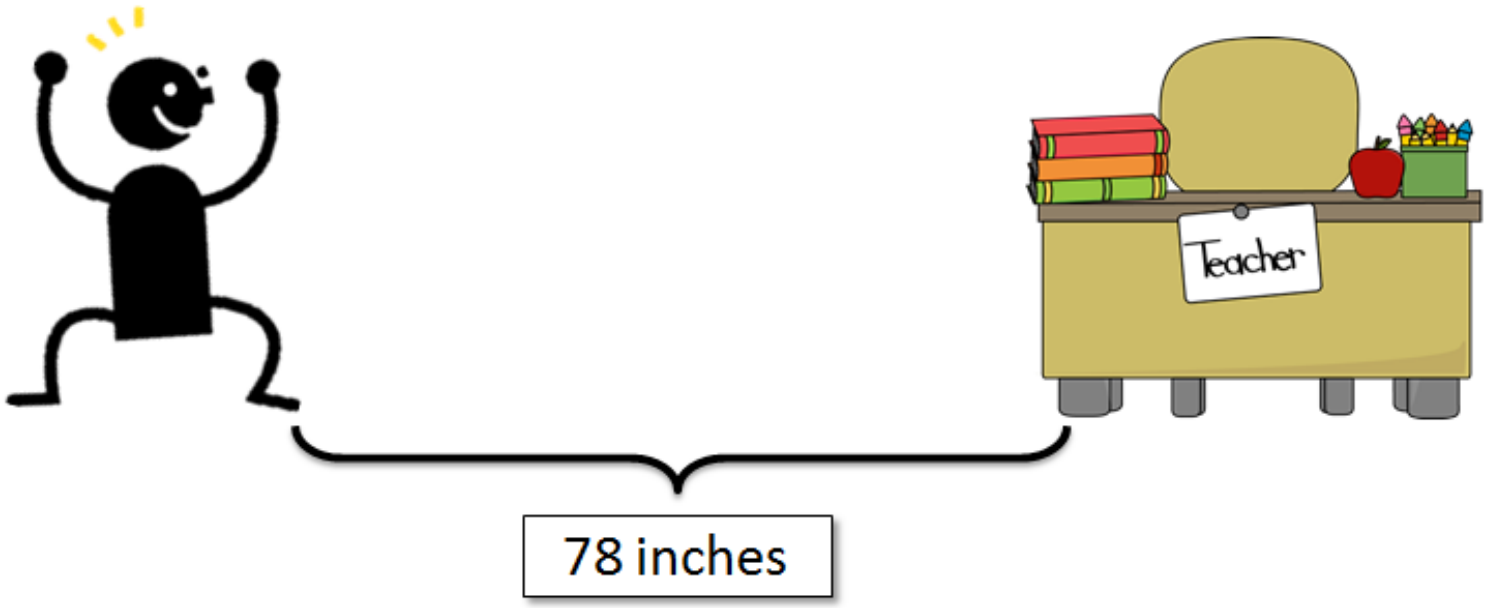
Concept approach

$$z = \frac{Observed\ - Expected}{Variation}$$

- Z is a distance and we count in SDs or SEs (***Variation***)
- Before getting to Z, we discover why…
  - We always subtract ***Exp*** from ***Obs***
  - We have to divide by ***Variation***

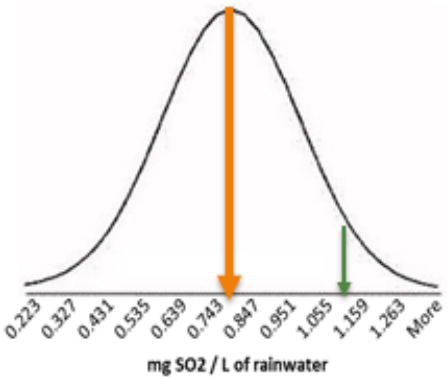
Concept approach

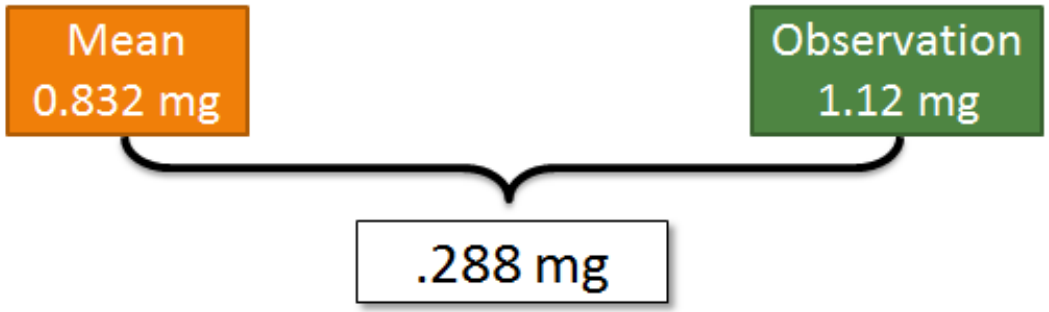
- How many feet am I standing from the table?
  - The distance is 78 inches
  - 1 foot = 12 inches



Concept approach

- How many standard deviations is the observation from the mean?
  - The distance is .288 mg
  - 1 SD = .171 mg





Formula approach

Confidence interval for the mean (σ known)

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Confidence interval for the mean (σ unknown)

$$\bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$$

Confidence interval for the proportion

$$p \pm Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

Concept approach

- A CI is an interval centred around the mean
  - i.e. the mean is always in the middle and you go the same distance either side
- Estimates for the mean ALWAYS use SE not SD
  - Test statistics, z/t-value, CI or probabilities
- $SE = \frac{SD}{\sqrt{n}}$
- Use correct distribution (Z/t) based on what is known
  - Z if SD is known, t if SD is not known
- If we're using Z:

$$Mean \pm Z(CI) * SE$$

Concept approach

- A CI is an interval centred around the mean
  - i.e. the mean is always in the middle and you go the same distance either side
- Use correct distribution (Z/t) based on what is known
  - Z if SD is known, t if SD is not known
- If we're using Z:

$$Mean \pm Z(CI) * Variation$$

- Estimates for the mean ALWAYS use SE not SD
  - Test statistics, z/t-value, CI or probabilities
- $SE = \frac{SD}{\sqrt{n}}$

Case study

MSc stats exam results...

Mark	No. Students	% Students
70+	17	74%
60-69	2	9%
50-59	3	13%
Below 50	1	4%

Case study

Student feedback...

"I was thinking that the statistics sessions would going to be again just for the sake of exam. But my prejudgement went wrong when I took several sessions with Christine. I found her teaching more practical as they were based on several real life examples; therefore easy to follow. The home assignments were enough to practice and improve the skills finally building confidence for exam. I am really impressed by her teaching method with full of listening and compassion."

"…I am a masters student…Previously I was educated at a respected state school in south bucks and…Cambridge University. However neither could conquer my inability in maths…I scrapped through previous exams but my lack of skill has always put me off the subject.

"When I saw stats on the Brunel curriculum I was therefore quite terrified. However Christine made it not only understandable, but enjoyable…I really felt I had finally gasped what had eluded me for so long and was rewarded with a score of 100% in the exam…"

Questions?