Read This First:

This document summarises one possible order of topic coverage of the AH Statistics course.
It is NOT being presented as the only way, or even the best way to cover the course content.
In each section, the order of the bullet points is not necessarily the exact order of teaching.
Where any resources are listed, they are given roughly in order of their use for that topic.
The resources listed are those which are used during lessons and/or set for pupils to complete for homework.

Information that is NOT included here are:
- actual lesson content
- detailed lesson plan notes
- technical usage of TI-Nspire, etc

All information is provided here without warranty or guarantee of its accuracy.

Nevil Hopley

Version 1: April 2019
Version 2: June 2019
Version 3: July 2022
Version 4: September 2022

Reference Codes:
DAM Data Analysis and Modelling Skills
1.1 Applying skills to data collection, presentation and interpretation
1.2 Applying skills to probability theory
1.3 Applying skills to discrete random variables
1.4 Applying skills to particular probability distributions

SI Statistical Inference Skills
1.1 Applying skills to sampling and the central limit theorem
1.2 Applying skills to intervals and estimation
1.3 Applying skills to bivariate analysis

HTHypothesis Testing Skills
1.1 Applying skills to parametric tests
1.2 Applying skills to non-parametric tests
1.3 Applying skills to bivariate tests

Textbooks Referenced, in addition to CIMT eBooks:
ACCIALS "A Concise Course in Advanced Level Statistics", by Crawshaw & Chambers, ISBN: 978-0748717576

Maths in Action: AH Statistics 2 by Nelson Thornes, ISBN: 978-0174315452

General Background Reading Resources:
https://www.bbc.co.uk/programmes/p02nrss1/episodes/downloads [BBC More or Less Podcast]
https://www.significancemagazine.com/ [Royal Statistical Society's Significance Magazine]
https://www.bbc.co.uk/programmes/p08tw9pl [A Short History of Probability]

11 x 55 min lessons

- Appreciate the necessary conditions for, and use of, the addition and multiplication laws of probability. [DAM1.2]
- Using set notation for probability theory $P(E \cup F)$, $P(E \cap F)$, $P(E|F)$, and $P(\bar{E})$, and compounds of these, such as $P(\bar{E} \cap F)$, [DAM1.2]
- Calculate probabilities for events which are not mutually exclusive. [DAM1.2]
- Compare calculated theoretical probabilities with those obtained experimentally, or by simulation using appropriate technology. [DAM1.2]
- Include Combinations and Permutations for use in Binomial Distribution and Mann Whitney Hypothesis test
- Calculate simple conditional probabilities. [DAM1.2]
- Calculate conditional probabilities requiring the use of Bayes' theorem or equivalent methods. [DAM1.2]

Notes:
Tree diagrams have events at their vertices, and probabilities on their edges/lines

Resources Used:
CIMT Statistics p5 Ex 1A no. 1-6, 8
DAM Cuboctahedron Net.pdf
CIMT Statistics p6/7/8 read, omit Activity 4, p8 Ex 1B no. 1, 2, 4
CIMT Statistics p9-13 Section 1.4 Read and take notes
CIMT Statistics p13 Ex 1C no. 1-4
TI-Nspire: Tree Diagrams.tns
CIMT Statistics p15 Example
CIMT Statistics p16 Ex 1D no. 1-3 (tree diagrams)
CIMT Statistics p18 Example for Class 7C
CIMT Statistics p19 Ex 1E no. 1-4
CIMT Statistics p23 Ex 1F no. 4, 5, 7
CIMT Statistics p21 Example
CIMT Statistics P22 Example
CIMT Statistics p23 Ex 1F no. 2
CIMT Statistics p24 Ex 1.8 no. 1, 2, 4, 8, 9, 12, 13, 14, 18, 21, 25, 27
CIMT Statistics p24 Ex 1.8 Extension questions no. 15, 17, 19, 28, 29, 30
DAM Combinations.pdf
DAM Combinations - Solutions.pdf
DAM Combinations Probabilities.pdf
DAM Combinations Probabilities - Solutions.pdf

Background Information:
Bayes' Theorem with Lego: https://www.countbayesie.com/blog/2015/2/18/bayes-theorem-with-lego
How Juries are fooled by Statistics: https://www.ted.com/talks/peter_donnelly_shows_how_stats_fool_juries
Dating for Bayesians: https://www.businessinsider.com/dating-for-bayesians-heres-how-to-use-statistics-to-improve-your-love-life-2013-11?IR=T
Understand the difference between P(vaccincated|infected) and P(infected|vaccinated): https://twitter.com/howie_hua/status/1421502809862664197

2 x 55 min lessons

- Present and interpret sample data in an appropriate form using a table, dotplot, stem-and-leaf diagram and boxplot. [DAM1.1]
- Identify possible outliers and suggest possible action to be taken. [DAM1.1]
- Appreciate that there are different methods of data collection and the difference between discrete and continuous data. [DAM1.1]

Notes:
Include examples of data that are categorical, qualitative, ordinal as well as numerical, quantitative, etc and consider the different charts used to display data sets that involve one or more of them.

Resources Used:
https://www.mygreatlearning.com/blog/types-of-data/
TI-Nspire: Standard Deviation Concept.tns
TI-Nspire: MyWidgets > StemAndLeaf.tns
CIMT Statistics p50 Example Data - sort ascending [data is not in a TI-Nspire file, so type it in manually]
CIMT Statistics p72 Example for back-to-back Stem and Leaf, and upper/lower fences

Background Reading
Interactively explore global shipping: https://www.shipmap.org

6 x 55 min lessons

- Use a Mann-Whitney test to assess evidence about the medians of two populations using independent samples. [HT1.2] (n,m<9 only at this stage)
- Use a Wilcoxon Signed-Rank test to assess evidence about the population median from a simple random sample and about the population distributions from paired data. [HT1.2] (small sample ONLY at this stage)
- Select and justify the choice of an appropriate test, together with its underlying assumptions. [HT1.2]
- Selecting relevant data [SI2.2 & HT2.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

<u>Notes:</u>
In SQA exams, the assumptions behind all hypothesis tests must be stated contextually, eg.
"the test scores are assumed to be the same shape and spread" and not "the data is the same shape and spread" (for Mann-Whitney)
"the difference in weights is assumed to be symmetrically distributed" and not "the data is assumed to be symmetrical" (for Wilcoxon)

Students need to be precise about their conclusions to Hypothesis tests – be careful not to be too bold in them and make sure they are contextualised.
You either "Have evidence to reject H_0" or "Have no evidence to reject H_0"
If you have evidence to reject H_0, then you need to give a non-technical, contextual conclusion
No need to reference H_1 specifically, but rather put H_1 into words

When conducting a Wilcoxon Paired Rank Sum test, you are testing H_0: median difference = 0, and not H_0: median A = median B.
This is because the median of differences is not the same as the difference of medians (but it is true that the mean of differences is the same as the difference of means)
In addition, you require the assumption that the paired differences are symmetrically distributed.

<u>Resources Used:</u>
TI-Nspire: NonParametric Assumptions.tns
TI-Nspire: MyWidgets > MannWhitneyTest.tns
HT Mann Whitney Introduction Example.pdf
HT Mann-Whitney Test Practice Exercise.pdf
HT Mann Whitney Test Recipe.pdf
HT Mann Whitney Distributions CfE.pdf
HT Mann-Whitney Formula Justification.pdf
https://nhost.shinyapps.io/MannWhitneyWilcoxonDistributions/
CIMT Further Statistics p80 example
TI-Nspire: MyWidgets > Wilcoxon2Test.tns
CIMT Further Statistics p82 example (includes tied ranks)
CIMT Further Statistics p84 Ex 4F no. 1-6 (introducing 2 tail tests in no. 4-6)
TI-Nspire: MyWidgets > Wilcoxon1Test.tns
HT Wilcoxon Signed Rank Single Sample Test.pdf (updated front page to include notes on differences and medians)

4 x 55 min lessons

- Appreciate that there are different methods of data collection, and be able to generate a simple random sample from a population. [SI1.1]
- Describe and distinguish between simple random, systematic, stratified and cluster sampling. [SI1.1]
- Appreciate that non-random sampling methods such as quota or convenience sampling could lead to an unrepresentative sample and biased conclusions. [SI1.1]
- Selecting relevant data [SI2.2 & HT2.2]

<u>Resources Used:</u>

https://www.thoughtco.com/what-is-statistical-sampling-3126366 [What is Statistical Sampling]
https://www.thoughtco.com/types-of-samples-in-statistics-3126353 [Types of Samples in Statistics]
https://www.thoughtco.com/sampling-designs-used-in-sociology-3026562 [Sample Designs Used in Sociology]
https://www.thoughtco.com/random-sampling-3026729 [Random Sampling Methods: Simple Random Sampling]
https://www.thoughtco.com/cluster-sampling-3026725 [Random Sampling Methods: Cluster Sampling]
https://www.thoughtco.com/stratified-sampling-3026731 [Random Sampling Methods: Stratified Sampling]
TI-Nspire: Stratified Random Sampling.tns
https://www.thoughtco.com/systematic-sampling-3026732 [Random Sampling Methods: Systematic Sampling]
https://www.thoughtco.com/convenience-sampling-3026726 [Non-Random Sampling Methods: Convenience Sampling]
https://www.thoughtco.com/quota-sampling-3026728 [Non-Random Sampling Methods: Quota Sampling]
https://www.thoughtco.com/simple-vs-systematic-random-sampling-3126369 [Comparison: Simple Random Sampling vs Systematic Random Sampling]
https://www.scribbr.com/methodology/systematic-sampling/ (for explanation of why the order matters)
SI Sampling Methods - Fish and Trees.pdf (Access link to files on RStudio Cloud: https://rstudio.cloud/project/118975)
SI Sampling Methods - Characteristics.pdf
SI Sampling Methods - Strengths and Weaknesses.pdf
CIMT Statistics: p45 Ex2.4 qn 1-5 to checkup on knowledge gained
Quick Sampling Quizzes.pdf

<u>Aspect to Emphasise</u>

Correct language is very important here.
The use of ‘accurate’ and ‘precise’ have clear meanings (see image here: https://sixsigmadsi.com/precision-and-accuracy/)
Encourage pupils to talk about whether or not the sample is ‘representative of the population’

<u>Background Information:</u>

Tracking Phones on London Underground: http://www.gizmodo.co.uk/2017/02/heres-what-tfl-learned-from-tracking-your-phone-on-the-tube/

3 x 55 min lessons

- Use a chi-squared test for association in a contingency table. [HT1.2]
- Deal with small expected frequencies. [HT1.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

Notes:
Include the context when stating H_0 and H_1, for example:
"H_0: there is no association between price of flights and month of the year" and not "H_0: there is no association", etc.

If there is evidence of an association, go further and conjecture what the type of association may be, in terms of the context, eg. "older people appear to use mobile phones less than younger people"

Resources Used:
TI-Nspire: Chi Squared Contingency Table.tns
HT Chi-Squared Contingency Table - Combining Categories - 2 examples.pdf
CIMT Statistics p210 Example
CIMT Statistics p212 Ex 11.3 no. 2
CIMT Statistics p212 Ex 11.3 no. 3, 4 (all contingency tables) [not 4b]
Chi-squared distribution: http://www.distributome.org/js/sim/ChiSquareSimulation.html
Analysis Decision Flowchart: https://statistics.nhost.uk/flowchart.html
Interactive Probability Distributions: https://www.geogebra.org/classic#probability

Background Information:
Simpson's Paradox: https://www.bbc.co.uk/programmes/p03sm8vw

5 x 55 min lessons

- Construct the probability distribution of a discrete random variable. [DAM1.3]
- Calculate the mean and standard deviation of a discrete random variable. [DAM1.3]
- Generate values of discrete data by simulation or experiment and compare their distribution to theoretical models. [DAM1.3]
- Include cumulative distributions in preparation for formal distributions.

<u>Resources Used:</u>
CIMT Statistics p87 read for discrete vs continuous
CIMT Statistics p90 Ex 4A no. 1, 2, 4, 5
TI-Nspire: Standard Deviation Divisor HH.tns
You Tube: https://youtu.be/-cozoTAwa54 [Standard Deviation and Variance Formulae (8m48s)]
You Tube: https://youtu.be/aPXyyYYA8U0 [Variance Formulae (6m15s)]
CIMT Statistics p92 Ex 4B no. 1, 3
CIMT Statistics p94 Ex 4C no. 1-4
CIMT Statistics p97 Ex 4.5 no. 1-4, 6-8 (omit qn 5)

3 x 55 min lessons

- Use the laws of expectation $E(aX+b) = aE(X)+b$; $E(X \pm Y)=E(X) \pm E(Y)$; $E(aX \pm bY)=aE(X) \pm bE(Y)$. [DAM1.3]
- Use the laws of variance $V(aX+b)=a^2V(X)$ and $V(X \pm Y)=V(X)+V(Y)$, where X and Y are independent. [DAM1.3]
- Calculate $SD(aX \pm bY)$ where X and Y are independent. [DAM1.3]

<u>Notes</u>
Emphasise the difference between $V(nX)$ and $V(X_1+X_2+\ldots+X_n)$, preferably by way of contextual examples.
In SQA exams, when justifying the use of $V(X \pm Y)=V(X)+V(Y)$, give the assumption of independence in context where possible, eg “the games are independent” and not “the random variables are independent”

<u>Interesting fact about the word 'Variance':</u>
"The 1 October 1918 was something of a watershed moment for statistics.
On that date, R A Fisher published a paper that introduced the statistical term ‘variance’ for the mean of the squares of the deviations of a measurement from its mean.
What followed was a whole new field of statistical analysis."

<u>Resources Used:</u>
TI-Nspire: Laws of Expectation and Variance.tns
You Tube: https://youtu.be/8wg-tyyTf7o [Proof of E(aX+b)=aE(X)+b and V(aX+b)=(a^2)V(X) (5m55s)]
You Tube: https://youtu.be/GY-kXk4UFrI [Proof of E(aX+bY)=aE(X)+bE(Y) (5m10s)]
You Tube: https://youtu.be/YhI0vHppbus [Proof of Var(aX+bY)=(a^2)Var(X)+(b^2)Var(Y), if X and Y are independent (7m29s)]
DAM Laws of Expectation and Variance Exercise.pdf (Highlight question 3 when 4X is not the same as $X_1+X_2+X_3+X_4$)
DAM Laws of Expectation and Variance Exercise Solutions.pdf

4 x 55 min lessons

- Calculate binomial probabilities. [DAM1.4]
- Use standard results for the mean and variance of this distribution. [DAM1.4]
- Simulate this distribution using appropriate technology and compare them to probability distribution models. [DAM1.4]

<u>Resources Used:</u>
TI-Nspire: Binomial Distn.tns
https://phet.colorado.edu/en/simulation/plinko-probability for Binomial Simulation
http://www.distributome.org/js/sim/BinomialSimulation.html for the Binomial Distribution Simulation
DAM Pascals Triangle.pdf
CIMT Statistics p104 Example
CIMT Statistics p105 Example
CIMT Statistics p107 Ex 5A no. 1-5
You Tube: https://youtu.be/anblNexgS9A [Proof of formulae for E(X) and Var(X) for Binomial Distribution (4m00s)]
CIMT Statistics p110 Example
CIMT Statistics p111 Ex 5B no. 1-5, 7-10
CIMT Statistics p112 Ex 5.3 no. 1-6
Diagram of Relationships between Distributions: https://www.johndcook.com/blog/distribution_chart/
Interactive Probability Distributions: https://www.geogebra.org/classic#probability

<u>Aspect to Emphasise</u>
Discuss/explore when you add $B(n_1,p_1)$ to $B(n_2,p_2)$ do you get another Binomial Distribution, or not?
This helps prepare pupils' thinking for when independent Normal distributions are combined.

---

2 x 55 min lessons

- Calculate uniform probabilities. [DAM1.4]
- Use standard results for the mean and variance of this distributions [DAM1.4]
- Simulate these distributions using appropriate technology and compare them to probability distribution models. [DAM1.4]

<u>Resources Used:</u>
TI-Nspire: Sum of RV Distributions.tns
You Tube: https://youtu.be/KA3dKdGryWc [Proof of formula for sum of first n integers (2m32s)]
You Tube: https://youtu.be/b4TqJaJVO6Q [Proof of formula for sum of first n square numbers (7m29s)]
You Tube: https://youtu.be/SSWEdTfy1so [Proof of formulae for E(X) and Var(X) for Uniform Distribution (3m56s)]

<u>Aspect to Emphasise</u>
Discuss/explore when you add $U(k_1)$ to $U(k_2)$ do you get another discrete Uniform Distribution?
This helps prepare pupils' thinking for when independent Normal distributions are combined.

4 x 55 min lessons

- Calculate Poisson probabilities. [DAM1.4]
- Use standard results for the mean and variance of this distribution. [DAM1.4]
- Simulate this distribution using appropriate technology and compare them to probability distribution models. [DAM1.4]

<u>Resources Used:</u>
TI-Nspire: Poisson Distn.tns
TI-Nspire: Poisson Simulation v3.tns
http://www.distributome.org/js/sim/PoissonSimulation.html for Poisson Distribution Simulation
CIMT Statistics p118/119 (for criteria)
CIMT Statistics p119 (both Examples manually)
You Tube: https://youtu.be/xbLUVHPYsIY [Proof of formulae for E(X) and Var(X) for Poisson Distribution (14m01s)]
www.johndcook.com/blog/distribution_chart/ [Diagram of Distribution Relationships]
CIMT Statistics p120 Ex 6A no. 1, 2, 3, 5
CIMT Statistics p122 Example - read and work through (For part c), discuss whether to use Y=10X or Y = $X_1+\ldots+X_{10}$)
CIMT Statistics p124 Example - read and work through
CIMT Statistics p125 Ex 6B no. 1-4
CIMT Statistics p129 Ex 6.4 no. 1, 3, 4, 7, 8, 13, 15
Interactive Probability Distributions: https://www.geogebra.org/classic#probability

<u>Aspect to Emphasise</u>
Discuss/explore when you add Po(λ_1) to Po(λ_2) do you get another Poisson Distribution?
This helps prepare pupils' thinking for when independent Normal distributions are combined.

3 x 55 min lessons

- Use a chi-squared test for goodness-of-fit to a discrete distribution. [HT1.2]
- Deal with small expected frequencies. [HT1.2]
- Selecting relevant data [SI2.2 & HT2.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

<u>Resources Used:</u>
CIMT Further Statistics p94 Ex 5A no. 2, 3, 5, 6
CIMT Further Statistics p95 example (Do using p=0.5, then re-do estimating p from observed frequencies to give p=0.55)
CIMT Further Statistics p102 Ex 5B no. 1, 3
CIMT Further Statistics p98 example (Run through first without combining categories, then re-do with correct combining)
CIMT Further Statistics p102 Ex 5B no. 2, 4
CIMT Further Statistics p110 Ex 5.4 no. 1, 2, 4, 6, 8, 9, 10
Analysis Decision Flowchart: https://statistics.nhost.uk/flowchart.html
Interactive Probability Distributions: https://www.geogebra.org/classic#probability

1 x 55 min lessons

- Calculate rectangular (continuous uniform) probabilities [DAM1.4]
- Use standard results for the mean and variance of this distribution. [DAM1.4]

Resources Used:
ACCIALS pp 349-353 (Uniform Distribution) [where ACCIALs = “A Concise Course in Advanced Level Statistics”, by Crawshaw & Chambers, ISBN: 978-0748717576]
DAM ACCIALS p349-353 Ex 6F - Continuous Uniform Distribution - Solutions.pdf
CIMT Statistics p146-148
Extension into general Continuous Distributions: CIMT Stats p148 Ex7.7

---

4 x 55 min lessons

- Calculate normal probabilities. [DAM1.4]
- Use standard results for the mean and variance of this distribution. [DAM1.4]
- Calculate probabilities in problems involving the sum or difference of two independent normal random variables. [DAM1.4]
- Include 'backwards' questions to determine parameter(s) from probability information.

Resources Used:
TI-Nspire: Normal Distn.tns
TI-Nspire: Skewed Normal Distribution v2NH.tns
TI-Nspire: Normal Sample Boxplots.tns
http://www.distributome.org/js/sim/NormalSimulation.html for Normal Distribution Simulation
Interactive Probability Distributions: https://www.geogebra.org/classic#probability
CIMT Statistics p157 Ex8B no. 1-10
CIMT Statistics p159 Ex8C no. 1, 2, 3, 4, 6, 8
CIMT Statistics p160 Example (inverse Normal)
CIMT Statistics p161 Example (find μ and σ)
CIMT Statistics p162 Ex8D no. 1-6 (talk about Ex 8D Q1+4+5+6 where you need to assume normality)
CIMT Further Statistics p11 Example (with a focus on part (b))
CIMT Further Statistics p14 Ex1B no. 1, 2, 3
CIMT Further Statistics p13 Example. Lots of random variables here!
For part (i), talk about whether $X_1+X_2+...+X_{25}$ is same as 25X
For part (ii), talk about $B_1+...+B_{250}$ or $T_1+...+T_{10}$
For part (iii), $P(M_{25}>M_{26}) = P(M_{25}-M_{26}>0)$
CIMT Further Statistics p18 Ex1.11 no. 3, 4, 5
Extension: CIMT Further Statistics p18 Ex1.11 no. 6, 7

Aspect to Emphasise
Discuss/explore when you add $N(\mu_1,\sigma_1^2)$ to $N(\mu_2,\sigma_2^2)$ do you get another Normal Distribution?
Highlight how this question has been asked about several distributions already met in this course.

2 x 55 min lessons

- Demonstrate an understanding of appropriate conditions for a normal approximation to a binomial or Poisson distribution, together with the parameters of the approximate distribution. [DAM1.4]
- Demonstrate the use of a continuity correction when applying a normal approximation to the binomial and Poisson distributions. [DAM1.4]

Resources Used:
CIMT Statistics p165 example
CIMT Statistics p168 Ex 8E no. 1-5
CIMT Statistics p170 Ex 8.7 no. 1, 2, 3, 4, 5, 9

---

2 x 55 min lessons

- Use a Mann-Whitney test to assess evidence about the medians of two populations using independent samples. [HT1.2] (revisited, for n,m>8)
- Use a Wilcoxon Signed-Rank test to assess evidence about the population median from a simple random sample and about the population distributions from paired data. [HT1.2] (revisited, for large samples)
- Use a normal approximation, when required in any calculation of a test statistic or p-value. [HT1.2] (also consider how to deal with non-integer values of M-W test statistics, and continuity correction)
- The use of a continuity correction is expected if a normal approximation is employed. [HT 1.2]
- Select and justify the choice of an appropriate test, together with its underlying assumptions. [HT1.2]
- Selecting relevant data [SI2.2 & HT2.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

Resources Used:
HT Normal Approx to MW and Wilcoxon Tests.pdf
TI-Nspire: Normal Approx to MW and Wilcoxon.tns [the data sets to accompany the above pdf file]

8 x 55 min lessons

- Use the sample mean as a best estimate of the population mean. [SI1.1]
- Use the sample variance as an estimate of the population variance. [SI1.1]
- Demonstrate an understanding that the sampling distribution of the sample mean from a parent population that is normal is itself normal. [SI1.1]
- Describe the sampling distribution of the sample mean. [SI1.1]
- Demonstrate an understanding that the sampling distribution of the sample mean from a parent population which is not normal is approximately normal, by invoking the Central Limit Theorem when the sample is large enough. [SI1.1]
- Perform a specified test for the population mean, for the cases: [HT1.1]
  (a) σ^2 known (z-test)
  (b) σ^2 unknown but a large sample (z-test)
  (c) σ^2 unknown with a small sample (t-test)
- Describe the sampling distribution of a sample proportion. [SI1.1]
- Perform a z-test for the population proportion. [HT1.1]
- Select and justify the choice of an appropriate test, together with its underlying assumptions. [HT1.1]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

Notes:
In SQA exams, both the assumptions and final conclusions for each hypothesis test must be stated contextually, eg.
"the weights of rabbits are assumed to be the normally distributed" and not "the data is assumed to be normally distributed"
"we have evidence at the 5% level for the mean weight of rabbits to be more than 2kg" and not "we reject H_0"

Emphasise that the CLT only justifies an approximate normal distribution for a sample mean for sufficiently large samples, but it does not justify approximating the parent population variance with a sample variance.

Resources Used:
TI-Nspire: Samples - Normal.tns
TI-Nspire: Sum of Normal RV.tns
CIMT Further Statistics p18 Ex1.11 no. 3, 4, 7
CIMT Statistics p195 Example
CIMT Statistics p196 Example (do by both p-values and critical regions)
CIMT Statistics p201 Ex10.5 no. 1, 2, 3
TI-Nspire: Samples – Uniform.tns (for introducing CLT)
TI-Nspire: Samples - Binomial.tns

CIMT Statistics p198 Example
CIMT Statistics p201 Ex 10.5 no. 4, 5
CIMT Further Statistics p38 Example (but disregard σ=3, and refer to solution on p40)
CIMT Statistics p198 example with t_{63}

CIMT Further Statistics p41 Ex 3A no. 1-6
CIMT Further Statistics p47 Example
CIMT Further Statistics p48 Ex 3C no. 1-6
CIMT Further Statistics p54 Ex 3.6 no. 3, 10, 22, 19, 21a, 13, 24, 15, 18, 19, 12, 4, 6
Analysis Decision Flowchart: https://statistics.nhost.uk/flowchart.html
Interactive Probability Distributions: https://www.geogebra.org/classic#probability

Background Information:
Whilst the first version of the CLT was published in 1733, it didn't get its current name until 1920 .... 187 years later!
How the CLT began: https://www.johndcook.com/blog/2010/01/05/how-the-central-limit-theorem-began/
The Flaw of Averages: https://www.thestar.com/news/insight/2016/01/16/when-us-air-force-discovered-the-flaw-of-averages.html

Summary of Test Assumptions

$X \sim N(\mu, \sigma^2)$, or assumed μ is known, or assumed under H_0 σ^2 is known any sample size	X is **not** normally distributed μ is known, or assumed under H_0 σ^2 is known n>20	$X \sim N(\mu, \sigma^2)$, or assumed μ is known, or assumed under H_0 σ^2 is **not** known Estimate σ with s_{n-1}	$X \sim B(n, p)$ $X \approx N(np, npq)$
Then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ $\Leftrightarrow \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1^2)$	Then $\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$ by CLT $\Leftrightarrow \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \approx N(0, 1^2)$	Then $\frac{\bar{X}-\mu}{\sqrt{\frac{s_{n-1}^2}{n}}} \sim t_{n-1}$ When using SQA data booklet and $n \geq 42$, then $t_{n-1} \approx Z$	Then $\frac{X}{n} \approx N(p, \frac{pq}{n})$ $\Leftrightarrow \frac{\frac{X}{n}-p}{\sqrt{\frac{pq}{n}}} \approx N(0, 1^2)$
Test statistic, $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$	Test statistic, $z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$	Test statistic, $t = \frac{\bar{x}-\mu}{\sqrt{\frac{s_{n-1}^2}{n}}}$	Test statistic, $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$

Notation

	Sample	Population
mean	$\bar{x}$	μ
standard deviation	s_{n-1}	σ
proportion	$\hat{p}$	p
Intercept	a	α
Slope	b	β
correlation coefficient	r	ρ
median	‘median’	η

6 x 55 min lessons

- Use a t-test to assess evidence about the population mean difference in a paired data experiment. [HT1.1]
- Test the hypothesis that two populations have the same mean, for cases where population variances are: [HT1.1]
  (a) known (z-test)
  (b) unknown but samples are large (z-test)
  (c) unknown and samples are small (t-test)
- Test the hypothesis that two populations have the same proportion, for only the case where both samples are large. [HT1.1]
- Select and justify the choice of an appropriate test, together with its underlying assumptions. [HT1.1]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

<u>Resources Used:</u>
http://www.distributome.org/js/sim/StudentSimulation.html for Students t Distribution Simulation
CIMT Further Statistics p74 Example
CIMT Further Statistics p75 Ex 4D (any 4 of the 6 questions)
CIMT Further Statistics p65 Example
CIMT Further Statistics p67 Ex 4B no. 2, 3, 5, 6
CIMT Further Statistics p70 Example
CIMT Further Statistics p70 Justification of Pooled s
CIMT Further Statistics p72 Ex 4C no. 1, 3, 4, 5
HT Two Proportions Introduction and Practice.pdf
Analysis Decision Flowchart: https://statistics.nhost.uk/flowchart.html
Interactive Probability Distributions: https://www.geogebra.org/classic#probability

<u>Background Information:</u>
'There is only one test!': http://allendowney.blogspot.com/2011/05/there-is-only-one-test.html
Heart Surgery Survival: https://www.bbc.co.uk/news/health-41763958

5 x 55 min lessons

- Calculate a z-interval for the population mean. [SI1.2]
- Appreciate the need to use Student's t-distribution when the population variance is unknown. [SI1.2]
- Calculate an approximate confidence interval for the population proportion. [SI1.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

<u>Note:</u>
In SQA exams, when it states "assess the evidence", it means that either confidence intervals or hypothesis tests are acceptable.

<u>Resources Used:</u>
Start on Which Test Quizzes - one per lesson, for 8 lessons.
TI-Nspire: Confidence Intervals v2.tns
CIMT Statistics p182 Example (z-interval)
CIMT Statistics p185 Ex 9.6 no. 1, 2, 3, 5, 6
CIMT Further Statistics p24 Example (t-interval)
CIMT Further Statistics p26 Ex 2A no. 1, 2, 3
CIMT Further Statistics p31 Example (proportion interval)
CIMT Further Statistics p32 Ex 2C no. 1, 2, 3
CIMT Further Statistics p34 Ex 2.5 no. 1, 3, 5, 8 (mixture)
SI Confidence Intervals Mixed Exercise.pdf
Analysis Decision Flowchart: https://statistics.nhost.uk/flowchart.html

<u>Background Information:</u>
Normal approximation for t distribution isn't good enough: https://www.johndcook.com/blog/2008/11/12/normal-approximation-for-student-t-distribution-isnt-good-enough/

A Cautionary Tale (from Jim Kay, Open University Lecturer)

Suppose that in the land of Hypotica, a sage randomly generates a sample of 16 observations from the $N(10, 5^2)$ distribution. The sage knows that the true mean is equal to 10, and that the variance is known to be 25. The sample mean is 9.67.

The sage gives a copy of the data to Robertus and Alicia and tells them that the sample has been generated from a Normal distribution which has known variance 25 (but she does not reveal true value of the mean, μ).

Robertus then conducts a test of the null hypothesis that $\mu = 11$ against the alternative that $\mu \neq 11$. He obtains the value of the test statistic as $z = (9.67 - 11)/(5/\sqrt{16}) = -1.064$. Since the true variance is known, the null distribution of Z is N(0, 1). Robertus then computes the p value of this two-sided test as $p = 2 \times P(Z < -1.064 \mid H_0) = 2 \times 0.1437 = 0.29$.

Seeing that there is little evidence against the null hypothesis, Robertus decides that the true value of μ is equal to 11.

Alicia conducts a test of the null hypothesis that $\mu = 9$ against the alternative that $\mu \neq 9$. She obtains the value of the test statistic as $z =(9.67 - 9)/(5/\sqrt{16}) = 0.536$. Since the true variance is known, the null distribution of Z is N(0, 1). Alicia then computes the p value of this two-sided test as

$p = 2 \times P(Z < -0.536 \mid H_0) = 2 \times 0.2960 = 0.59$.

Seeing that there is little evidence against the null hypothesis, Alicia decides that the true value of μ is equal to 9.

They both tell the sage, and each other, their results, and the sage says that the true value of μ is 10. Robertus and Alicia are very confused by the contradictory results. The sage explains that the contradiction has arisen because they had both decided that their null hypothesis is true – an impossibility.

The sage, who prefers to consider interval estimates for unknown parameters where possible, then told them that the 95% confidence interval for μ is $9.67 \pm 1.96 \times 5/\sqrt{16} = (7.22, 12.12)$.

The most plausible value for μ given the data is 9.67, and Robertus' value of 11 and Alicia's value of 9 both lie within the confidence interval, and so they are both plausible values of μ, but not the true value.

4 x 55 min lessons

- Construct and interpret a Shewhart control chart for the sample mean or proportion. [SI1.2]
- Use the Western Electric rules to recognise when a process may be out of statistical control and influenced by common or special causes. [SI1.2]
- Estimate chart parameter(s) from a sample when population data is unavailable. [SI1.2]
- Selecting relevant data [SI2.2 & HT2.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

Notes:
When doing a p-chart and the Control Limits are <0 or >1, it's often due to the B(n,p) not being well approximated by the Normal, as either $np<5$ or $nq<5$

Resources Used:
http://www.learner.org/courses/againstallodds/unitpages/unit23.html
https://en.wikipedia.org/wiki/Western_Electric_rules
Nelson Thornes: Maths in Action: Statistics 2 pages 1-13
CIMT Further Statistics pages 147-166
CIMT Further Statistics p150 Activity 2 & 3
CIMT Further Statistics p163 Example
CIMT Further Statistics p165 Ex 8.7 no. 2, 5, 6

7-10 x 55 min lessons
- Interpret a scatterplot, observing whether or not a linear model is appropriate. [SI1.3]
- Calculate and interpret the product moment correlation coefficient. [SI1.3]
- Comment on given simple transformations to obtain improved models. [SI1.3]

- Test the hypothesis that the population correlation coefficient is zero. [HT1.3]
- Calculate the least squares regression line of y on x. [SI1.3]
- Calculate a fitted value and its residual. [SI1.3]
- Test the hypothesis that the slope parameter in a linear model is zero. [HT1.3]
- Interpret a residual plot. [SI1.3]
- Calculate and interpret the coefficient of determination. [SI1.3]
- Assess the reliability of prediction based on fitted values, considering the effects of correlation, interpolation and extrapolation. [SI1.3]
- Calculate the appropriate statistics required for bivariate intervals. [SI1.3]

- Construct a prediction interval for an individual response. [SI1.3]
- Construct a confidence interval for a mean response. [SI1.3]
- Appreciate the difference between regressing y on x and x on y. [SI1.3]
- Communicate appropriate assumptions. [HT1.3]
- Selecting relevant data [SI2.2 & HT2.2]
- Presenting and analysing the data [SI2.3 & HT2.3]
- Communicating the conclusion [SI2.4 & HT2.4]

Resources Used:
Nelson Thornes: Maths in Action: Statistics 2 pages 81-103
TI-Nspire: Bivariate - Sampling.tns with accompanying video: https://www.youtube.com/watch?v=faN0a1fg6yw
Interactive demo of Bivariate Sampling: https://nhost.shinyapps.io/Sampling/ with accompanying video: https://youtu.be/tbDj5j3zZRs
Significance of r: CIMT Statistics p221
Limitations of Correlation Coefficient: CIMT Statistics p223
Analysis Decision Flowchart: https://statistics.nhost.uk/flowchart.html

TI-Nspire: Bivariate - Sxx Syy Sxy r.tns with accompanying video: https://www.youtube.com/watch?v=KU8C-uEQXEo
TI-Nspire: Bivariate - Limitations.tns with accompanying video: https://www.youtube.com/watch?v=7L1fhxGtITs
http://guessthecorrelation.com/
Hypothesis Test on rho: CIMT Statistics p222 Example & p223 Ex12A no. 2, 4, 5
Equation of a Regression Line: CIMT Statistics p230
TI-Nspire: Bivariate - Regression Line.tns with accompanying video: https://www.youtube.com/watch?v=wTEuYR5z2ds
TI-Nspire: Bivariate - Residual Plots 1.tns with accompanying video: https://www.youtube.com/watch?v=nQz83tQ06hE
TI-Nspire: Bivariate - Residual Plots 2.tns with accompanying video: https://www.youtube.com/watch?v=x4io8FvUbH0
TI-Nspire: Bivariate - P and C Intervals.tns with accompanying video: https://www.youtube.com/watch?v=7W29mWHtb1g