STATISTICAL ANALYSIS IN ADVANCED SPECIAL AND ADVANCED LEVEL BIOLOGY IN ENGLAND

R. Tynan

Liverpool John Moores University (UNITED KINGDOM)

Abstract

The education white paper of 2010 [1] initiated reforms to curriculum design and assessment for schools in England. This re-established subject disciplines as the unit for school curriculum design, and written end-of-course examinations as the sole method for assessing substantive and discipline knowledge at General Certificate of Secondary Education (GCSE) and General Certificate of Education Advanced Special/Advanced (GCE AS/A) Level [1]. This reflected the importance that central policy makers in England placed upon learners experiencing different research methodologies and content areas and acknowledged the challenges previously experienced by teachers in schools and colleges if coursework and practical examinations contributed to external qualifications [1]. Due to the variability of living organisms, it can be argued that biology disciplinary knowledge is underpinned by statistical analysis [2]. Biological investigations often require randomised controlled trials (RCTs) and the use of statistical analysis [2] to establish new knowledge.

This paper reviews the Assessment Quality Association (AQA) GCE AS/A Level Biology specification for assessment from 2016 [3] as typical of the specifications compliant with the aims of the English government's white paper [1]. It considers the requirement on learners to be able to use statistical analysis and the substantive knowledge this supports. Learning, teaching, and assessment (LTA) approaches that have been used to deliver the examination specification [3] are discussed, drawing upon experiences with learners in secondary schools and colleges, and with pre-service science teachers on initial teacher education (ITE) programmes.

The statistical requirement [3] emphasises the importance of learners selecting and using appropriate statistical methods and the interpretation of findings to reach valid statistical conclusions. Assessment does not intend that learners remember formulae or methods. This paper considers the use of comparison of means, Chi-squared analysis, and Spearman's Rank Correlation to support the study of animal behaviour, genetics, evolution, ecology, and disease treatment.

Keywords: Statistical analysis, A Level, Biology, chi-squared, correlation, t-test, random sampling, substantive knowledge, disciplinary knowledge.

1 INTRODUCTION

Due to the high degree of variability amongst living organisms, biological epistemology is usually derived from data generated from experiments involving specific populations and repeated trials and statistical analysis to reach objective and valid conclusions [2]. Basic statistics texts describe the purpose, calculation, and interpretation of tests of significance [2] [5], whilst others explore more fully the relationship between sampling, tests of significance, the statistical power of tests, and effect sizes [6]. Statistical tests of significance aim to establish the probability of error when a null hypothesis (H^0) that there is no difference between two statistics is rejected. Fig. 1 illustrates the four possible outcomes of an experiment and identifies the rejection of H^0 as the only result where the probability of error can be quantified. Statistical power is a measure of the ability of statistical tests to detect patterns with low probabilities of being random, which much depends on sample sizes and methods [6]. Effect size metrics are concerned with degrees of separation or overlap of distributions regardless of the probabilities of error if these are assumed to be non-random [6]. Experimental designs vary but randomised controlled trials (RCTs) are common in biological research to reduce the effects of bias, and individual or sample error in the laboratory and field [2].

After the government white paper of 2010 [1], examination boards in England re-wrote their specifications to comply with its requirements. The white paper [1] heralded significant curriculum and assessment reforms, re-establishing subject disciplines as the unit of curriculum design in schools in England. It also ended the contribution of modular examinations, coursework and practical examinations to external qualifications offered at General Certificate of Secondary Education (GCSE) and General Certificate of Education Advanced Special and Advanced Levels (GCE AS/A-Level). GCSE level

ISBN: 978-84-09-62938-1

specifications emphasise data handling, interpretation and descriptive statistics, but do not require or assess statistical analysis [4]. For the purposes of this study, the Assessment Quality Association (AQA) GCE AS/A-Level Biology specification for first assessment in 2016 [3] was selected as typical of the biology specifications compliant with the aims of the English government's white paper [1] that require and assess aspects of statistical analysis.

		Null Hypothesis (H^0)	
		H^0 Valid	H^0 Invalid
Conclusion	Reject H^0	Type 1 Error (5% chance of error)	Correct
	Accept H^0	Correct	Type 2 Error (error unknown)

Figure 1 Statistical errors.

The GCE AS/A-Level Biology specification [3] lists and explains biology content, mathematical and practical requirements and the structure and weightings for their assessment. A key attribute of the new GCE AS/A-Level specifications [3] is that both substantive and disciplinary knowledge (theoretical content and investigation skills respectively) are assessed only on the final written examinations. In science subjects, to avoid the disappearance of practical work and investigations from the school curriculum, examination specifications prescribe a list of required practical investigations that teachers must accredit that learners have experienced [3]. However, the skills that these are intended to develop are not assessed unless they can be tested by written examination questions. Although investigation skills can be argued to contribute to all three assessment objectives stated by the AQA GCE AS/A-Level Biology specification [3], 15% of marks are linked directly to practical work on the two AS examination papers and 'at least 15%' (p.74, [3]) over the three A-Level papers.

The specification [3] designates four sections of content and their associated mathematical and practical requirements for assessment at AS-Level in a one-year programme. A-Level students study these and a further four sections of content over two years [3]. Statistical analysis is mentioned in several areas of the specification [3] linked to content and skills development activities. The mathematics requirements name the calculations and tests of significance that could be examined. The specification [3] requires learners to be able to select and use three tests of significance: Chi Squared (χ^2) test, Students' t-test, and correlation coefficients [3]. However, learners are also expected to be able to calculate and use standard deviations and demonstrate an understanding of the laws of probability in order to evaluate experimental designs and the strength of evidence [3]. Notably, the content, mathematical and practical assessment requirements [3] do not define the H^0 nor its use in reaching statistical conclusions. It does not describe the use of distribution probability tables nor the importance of an objective, agreed international standard for acceptable error when rejecting the H^0.

The named statistical tests [3] have very clear and different uses in biology. The student t-test [2] [5] compares means, starting with the H^0 that there is no difference between them or between them and the mean of the sampled population. It can be used to compare the effect of changing an independent variable on a measured outcome, or dependent variable, compared to a control group. It is a parametric test that assumes the variable follows a normal distribution (bell-shaped curve), to calculate a table of probabilities for obtaining different t-values with different sample sizes by random fluctuations in the data. The H^0, that this is case and there is no difference between two means, is rejected if the probability of error in doing so is 5% or less.

χ^2 analysis [2] [5] is only used with counts and frequencies (never measurements) where individuals belong to mutually exclusive categories. This is a non-parametric test that does not assume a normal distribution and is used to compare observed counts and frequencies with expected results. The expected results can be predicted by a theoretical model, assuming random distribution, or calculated using contingency tables. The H^0 states that there is no difference between observed and expected counts. It has several applications in biological investigations [2] discussed further below with respect to the AQA AS/A-Level Biology specification [3].

Correlation coefficients are used to investigate associations between paired data sets. The H^0 states that there is no correlation between the two sets of data. There are parametric and non-parametric versions [2][5] with teachers often choosing non-parametric correlation methods for smaller data sets that are unlikely to demonstrate normal distributions.

The aim of this paper is to discuss the opportunities for using learning teaching and assessment (LTA) activities involving the named statistical tests, utilizing primary or secondary data, that support and enhance the learning of content and delivery of required practical work [3].

2 METHODOLOGY

The activities and LTA experiences discussed here were conducted with 16–19-year-old learners during the delivery of their AS or A-level curriculum during the author's time as a secondary teacher (1979-2007) in schools with a sixth form. Whilst working as teacher educator and subject tutor on ITE programmes, they were shared with pre-service science teachers (2007- time of writing) and refined during subject knowledge and pedagogy sessions. This practitioner research constitutes a natural, ethnographic, pedagogical study with data collected from literature, document analysis, participant observation of LTA activities, and their evaluation by pre-service science teachers. No other primary data was collected from participants. Following British Educational Research Association (BERA) guidelines, there are no ethical issues in the anonymous reporting of the findings [7]. The literature used was encountered during LTA activities and should not be considered a systematic or comprehensive review. The AQA Biology Specifications for GCE AS/A-Level [3], referred to throughout this paper, were last downloaded from the AQA website on April 30th, 2023.

3 RESULTS AND DISCUSSION

3.1 Statistical analysis in the AS/A-Level examination specification

Selection and use of named statistical tests is one of eleven mathematical skills described in the data handling sub-section of the AQA AS/A-Level Biology specification [3]. In addition, there are four other subsections covering a wide range of mathematical skills. Many of the skills mutually support each other and examination questions involving statistical tests could also assess related mathematical skills. There is also a degree of overlap with practical skills to be assessed such as those involving descriptive statistics, evaluating results and drawing conclusions, with the practical skills section referring to processing and analysing data as exemplified in the mathematical skills section [3].

There is no prescribed teaching sequence for the AQA GCE AS/A-Level Biology content [3], but the suggested associated learning experiences indicate the intended mathematical progression for learners moving between AS and A-Level studies [3]. Content sections 1 to 4 are examined on two AS-Level papers, and the first and third A-Level papers [3]. The associated mathematical requirements focus on formulae, calculations, data handling, descriptive statistics, and evaluation [3]. Percentage change, means, standard deviations, sampling and errors are introduced first, with reference to tests of significance in an A-Level module to be examined on the second and third A-Level papers [3].

The specification [3] utilises much cross-referencing between sections. This avoids repeating information but complicates tracing themes and ensuring that all requirements are met in a planned curriculum. The mathematical appendix names three statistical tests of significance. It requires learners to correctly select and use the student t-test, χ^2, and correlation coefficients [3], MS 1.9, p.70. They are also required to calculate means and standard deviations and use these as a measure of dispersion around a central tendency [3], MS 1.10, p. 70. Standard deviations often form part of the calculation of parametric tests of significance including the t-test and Pearson's correlation coefficient [2]. Standard deviations can also be used to calculate 95% confidence limits of means and 5% errors for repeated measurements to develop an understanding of t-test calculations.

The statistical tests named in the AQA specification [3] are used to identify differences between data sets that have a low probability of being due to random variation. The student t-test compares means, χ^2 differences between observed and expected counts, and correlations coefficients the relationship between paired data sets. Introductory statistical readers describe the contexts and conditions for the appropriate use of each, with examples, formulae, instructions, and relevant probability tables [2] [5] [6]. There is no requirement for learners to remember formulae for statistical tests although calculating them with instructions could be included in a statistical analysis question to examine other mathematical skills [3]. The apparatus and techniques requirements [3] refer to using software to handle and analyse data. It follows that activities and learning experiences can involve templates with statistics formulae inserted or manually entered in spreadsheet cells. This allows learners to focus on the selection and use of statistical tests rather than their construction or calculation. Then learners need only collect and enter raw data and interpret the result. Previous proceedings papers, focusing on the impact of the 2010 education white paper [1] on biology fieldwork in England [8] [9], contain some examples of statistical analysis templates. Learning experiences linked to statistical tests could involve primary or secondary data but should provide the learner with the opportunity to select the appropriate test and make a valid statistical conclusion [2] [3]. Valid statistical conclusions involve random or systematic sampling [2] [5] [6].

Table 1 summarises opportunities associated with statistical analysis suggested by the AQA GCE AS/A-Level Biology specification [3] content, mathematical requirements, practical skills and required practical work. There are several precise references to the use of means and standard deviations but few references to the named statistical tests. However, the references to evaluation, errors, uncertainty, processing data using software and sampling would seem to imply opportunities to utilise statistical tests [3]. It could be that suggesting statistical analysis examples more often in the specification would give teachers the false impression that it had a higher assessment emphasis than other mathematical and practical skills inconsistent with it contributing only a fraction of the examination paper marks weightings of 10% and 15% for mathematical and practical skills respectively [3]. The specification leaves teachers free to deliver more than the minimum requirements for statistical approaches or not, as they deem appropriate for their learners and their own statistical experience and expertise. Table 1 includes content and suggested activities that are suitable for LTA activities involving statistical analysis not referred to by the specification [3].

Table 1 Opportunities for using standard deviation and statistical tests of significance in the AQA GCE AS/A-Level specification for assessment from 2016 [3]

Section	Reference	Opportunity for statistical analysis
Content		
Mass flow	3.3.4.1	AT h: Students could design and carry out an investigation into the effect of a named variable on human pulse rate or on the heart rate of an invertebrate, such as Daphnia. No statistical reference but suitable for comparison of means or correlation.
Species and taxonomy	3.4.5	No reference but opportunity to measure diagnostic features and use standard deviations and 95% confidence limits to identify hybrids and separate races or species.
Variation	3.4.7	AT k, MS 1.2, MS 1.10: Quantitative investigations of variation within a species involve: collecting data from random samples, calculating a mean value of the collected data and the standard deviation of that mean, interpreting mean values and their standard deviations. Students will not be required to calculate standard deviations in written papers.
Nutrient cycles	3.5.4	PS 1.1: Students could devise investigations into the effect of named minerals on plant growth. No reference to statistical analysis but suitable for comparison of means or correlation depending on the experimental design.
Control of heart rate	3.6.1.3	Students could design and carry out an investigation into the effect of a named variable on human pulse rate. No statistical reference but suitable for comparison of means or correlation.
Inheritance	3.7.1	Use of Chi squared specified in the content and opportunity for skills development: MS 1.9: Students could use the chi squared test to investigate the significance of differences between expected and observed phenotypic ratios.
Genotype and phenotypes	3.7.1	No statistical reference but the use of correlation coefficients to demonstrate phenotypic environmental tracking by plant leaves would be a suitable investigation or secondary data exercise.
Population genetics (Hardy-Weinberg equilibrium)	3.7.2	No statistical reference but suitable for chi squared test.
Populations in ecosystems	3.7.4	No statistical reference but the suggested investigations are suitable for correlation coefficients and chi squared association analysis.
Required practical work		
1: Investigation into the effect of a named variable on the rate of an enzyme-controlled reaction.	7.2	PS 3.3 Students could calculate the uncertainty of their measurements of the rate of reaction.
10: Investigation into the effect of an environmental variable on the movement of an animal using either a choice chamber or a maze.	8.2	No reference to statistical analysis but suitable for the chi squared test.
12: Investigation into the effect of a named environmental factor on the distribution of a given species.	8.2	No reference to statistical analysis but the investigation could involve correlations and chi squared association analysis.

***AT:** Apparatus and techniques, **MS:** Mathematical skills, **PS:** Practical skills*

3.2 Common features of statistical tests for A-Level Biology

A recurring theme in the mathematical and practical skill descriptors associated with statistical analysis in the AQA AS/A-Level specification [3] is calculating errors and evaluating confidence in data and conclusions. In AS modules early references to means and standard deviations introduce ideas about central tendencies, outlying results, and errors. Understanding means and standard deviations is necessary to understanding how t-tests and Pearson's correlation work and what they can demonstrate. This leads onto an understanding of when it would be more appropriate to use a non-parametric method such as χ^2 or Spearman's rank correlation.

Field Studies Council tutors [10] use statistical analysis to make sense of field work data with sixth form learners. They introduce statistical analysis as an objective method for ensuring that different investigators reach the same conclusion regarding data by using an internationally agreed standard for acceptable error when risk taking. This is achieved by the adoption of the H^0 that there is no difference between two sets of data or correlation between paired data sets. The H^0 can only be rejected if there is a probability of 0.05 (5%) or less that a difference or correlation could be explained by random variations in the data. It also assumes that sampling will be objective, either random or systematic. Error can only be quantified when the H^0 is rejected (Fig. 1).

3.3 Comparing means: from means and standard deviations to the student's t-test

Although the AQA AS/A-Level Biology specification [3] mentions the use of modes and medians as measures of central tendency when appropriate it does not elaborate on this. The student's t-test compares two means. This is a parametric method that assumes the means are calculated from variables that demonstrate a symmetrical normal distribution (Fig. 2). For normal distributions the mode, median and mean have the same numerical value. For example, this condition is satisfied if, in a class survey of height measurements (Table 1, Variation, 3.4.7), the largest number of participants measured 1.7 metres (mode), 50% of the others were less than this height and 50% more (median), and the calculated mean (sum of variables divided by the number of variables) is also 1.7 metres. This does not always occur but most parametric methods give valid conclusions even when distributions are skewed to a small extent (Fig. 3).

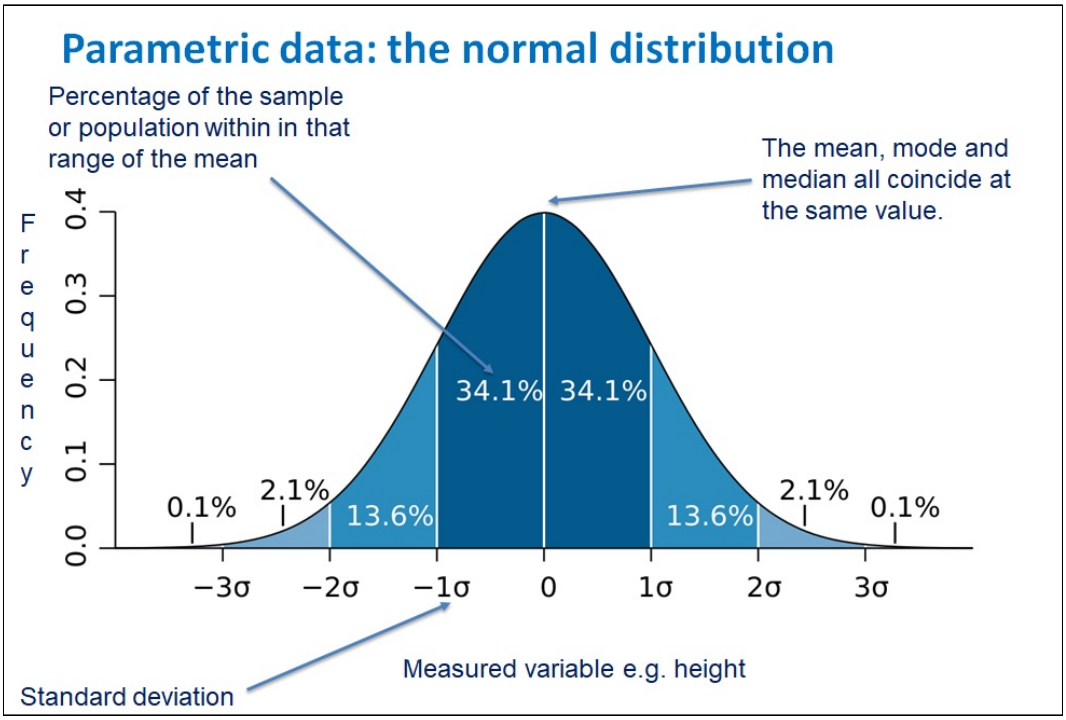

Figure 2 Teaching slide showing the characteristics of normal distributions.

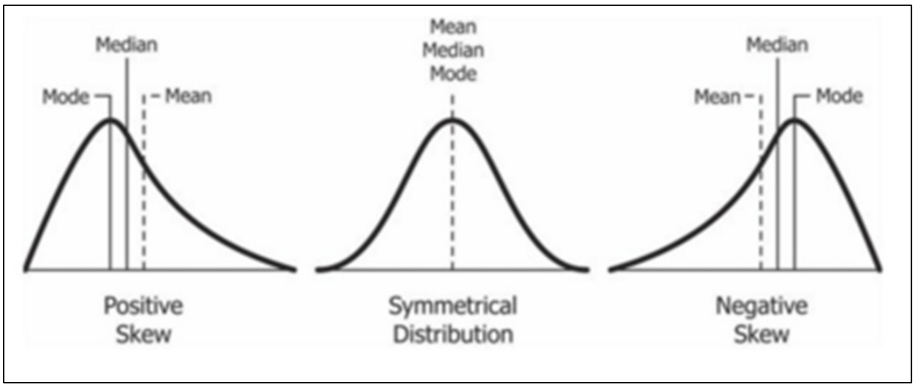

Figure 3 Teaching slide showing skewed distributions.

The examination specification [3] suggests the use of standard deviations to identify outlying measurements. This is a measure of the dispersion or spread of variables calculated as an average deviation of variables from their mean. A large standard deviation indicates a broad bell-shaped curve and a small standard deviation a narrow curve. Both means and standard deviations are formulae available on spread sheets for insertion into cells. Learners are not required to remember the formulae for the standard deviation (Fig. 4) nor calculate it from first principles on written papers, but this exercise would satisfy other mathematical skill requirements.

The areas under the curve in Fig. 2 also indicate the probabilities of a measurement being within the number of standard deviations shown. Around 95% of normally distributed variables are found within two standard deviations of the mean. So, a general approximation for biologists would be that measurements more than two standard deviations from the mean could be anomalous. To compare means suspected of being significantly different the 95% confidence limits of the means can be calculated. This would involve their standard deviations, sample sizes and the associated t-value for that sample. However, the progression from AS to A-Level described in the AQA specification moves straight onto the use of the t-test and does not include this approach [3].

$$\sigma = \sqrt{\frac{\sum(X-\mu)^2}{n}}$$

where,

σ = population standard deviation
$\sum$ = sum of...
μ = population mean
n = number of scores in sample.

Figure 4 Teaching slide showing a formula for calculating the standard deviation.

The student t-test does not necessarily require large sample sizes and can be used in a range of contexts involving measurements or counts comparing two treatments or contexts. This involves a simple experimental design such as comparing a control group and a treatment group, or an established and new treatment, or a *placebo* and a treatment. Multi-factor analysis is a slow process using t-tests and usually only used after analysis of variance (ANOVA) establishes it would be worthwhile.

When introducing the t-test it has been found useful to use an exemplar with clear cut results. Pedunculate Oak (*Quercus robur*) and Sessile Oak (*Quercus petraea*) are named after a diagnostic recognition feature. Pedunculate Oaks have a petiole or stalk at the point of attachment of their acorns that Sessile Oaks lack. However, Pedunculate Oaks also have distinctively shorter leaf petioles (1-5 millimetres) than Sessile Oaks (18-25 millimetres) [11]. In numerous locations, such as Coed Hafod (Summer Wood) in North Wales, both species of Oak can be found together in the same wood. Field Studies Council [10] tutors using this site developed a simple activity to help develop an understanding of Species and Taxonomy ([3], 3.4.5). Representative sampling of Oak trees of both species present in a wood or arboretum to measure a random sample of leaf petioles or stalks from each tree will demonstrate mean measurements within each species' leaf petiole range. The student t-test on the samples from the two Oak species will usually result in the conclusion that the difference between the means is significant with a 5% or less chance of error. However, if hybridisation has occurred at a site, hybrid trees demonstrate a mean leaf stalk intermediate between the other two means and the differences may or may not prove to be significant.

Learners only need use a spreadsheet to calculate the mean (m) and standard deviation (SD) for each set of measurements and the number of variables (n). Then the t value is calculated using the formula: (m1-m2) / Square root ((SD1/n1) + (SD2/n2)), which can be entered manually into a spreadsheet cell. Learners then consult a t-test distribution probability table (Fig. 5) indicating the probability of a given t-value arising due to random variations. Table 1 shows there are numerous opportunities to calculate means and standard deviations in the AS modules and any data from AS module activities and required practicals can be used later in an A-Level course to practise selecting and using t-tests (Table 1).

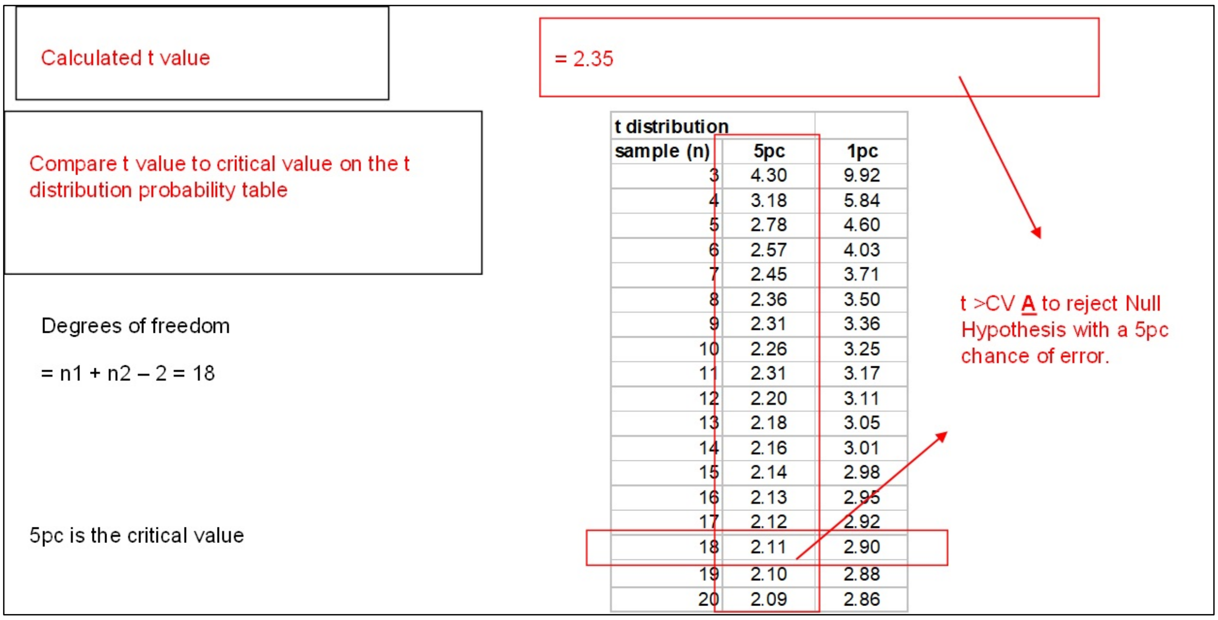

Figure 5 Teaching slide demonstrating a t-test conclusion.

3.4 Using the Chi squared (χ^2) test

χ^2 is a non-parametric test of significance suitable for smaller samples where the data is not normally distributed. The key issue for learners selecting and using χ^2 analysis appropriately is that it is only appropriate for comparing observed and expected counts and frequencies. The opportunities for using χ^2 in Table 1 demonstrate three contexts involving different ways to calculate the expected counts.

The only direct reference to a named statistical test in the content and suggestions for skills development is the use of χ^2in the A-Level genetics module [3]. Mendelian genetics provides a theoretical model for predicting genotype and phenotype ratios for monohybrid and dihybrid crosses involving fully dominant and recessive alleles or codominant alleles. These can be applied to the total count for the observed results to calculate expected results. For example, if 100 seeds germinate after a Mendelian monohybrid heterozygous cross the expected phenotype ratio is 3 plants with a dominant characteristic to every 1 plant with a recessive characteristic. The expected counts would be 75 and 25 respectively. If the observed counts were 69 dominant and 31 recessive plants, χ^2 can be used test if the difference between observed and expected counts is sufficient to reject the H^0.

The AQA specification [3] describes the Hardy-Weinberg equilibrium. Similarly, this predicts that allele frequencies for a pair of mendelian alleles will remain constant over generations given a set of assumptions, such as there is no natural selection occurring. Expected allele and phenotype frequencies can be calculated by the mathematical model $p^2 + 2pq + q^2 = 1$ and $p + q = 1$, where p and q are the probability (frequency) of alleles A and a, where A is dominant [3]. χ^2 can be used to look for significant variations in allele frequencies and phenotype frequencies to those predicted. They should remain the same over a period of several generations if the Hardy Weinberg assumptions are valid for a given population.

However, when an investigation involving a choice chamber changes one or more environmental factors and measures the distribution of an organism, the H^0 assumes that the factors have no effect on the organism. If so, it will be distributed randomly and found in equal numbers in all chambers at any point in time. The expected count for each chamber is then the number of organisms divided by the number of chambers. Fig. 6 shows a choice chamber spreadsheet template.

Chi2	**Choice Chambers**					
Conditions	**O**	**E**	**O-E=D**	**D2**	**D2/E**	
Dry Dark	20	25	-5	25	1	
Dry Light	10	25	-15	225	9	
Humid Dark	40	25	15	225	9	
Humid Light	30	25	5	25	1	
				Sum	20	
degrees of freedom= chambers - 1						
Chi2						
deg of free	5pc	1pc				
1	3.84	6.64		20>7.81	20>11.34	
2	5.99	9.21		Sig	less than 1%	
3	7.81	11.34			chance of error	
4	9.49	13.28				
5	11.07	15.09				
6	12.59	16.81				

Figure 6 Chi squared choice chamber spreadsheet template.

Finally, some χ^2 experimental designs involve contingency tables where the expected counts are calculated from row and column totals (Fig. 7). The use of χ^2 for association analysis in this way together with comparison of means and correlations to satisfy the ecology required practical (Table 1) has been discussed previously in proceedings papers [8] [9]. Fig. 7 illustrates a template for a small-scale treatment trial.

Enter the Observed Results in the table.											
Observed:		**Treatment**									
		New	Old	Placebo	Totals						
Outcome	Pain relieved	26	17	8	51						
	Pain continued	3	9	7	19						
	Totals	29	26	15	70						
						Calculating expected counts:					
Expected:		**Treatment**				Success rate 50/70			0.73		
		New	Old	Placebo		Failure rate 20/70			0.27		
Outcome	Pain relieved	21	19	11		New treatment rate		29/70	0.41		
	Pain continued	8	7	4		Old treatment rate		26/70	0.37		
						Placebo rate		15/70	0.21		
						Expected numbers					
	o-e	d2	d2/e			New treatment and success			(50/70) x (29/70) x 70		= 21
	5	23.73	1.12			New treatment and failure			(20/70) x (29/70) x 70		= 8
	-5	23.73	3.01			etc…					
	-2	3.77	0.20								
	2	3.77	0.53								
	-3	8.58	0.78								
	3	8.58	2.11								
Total	0.00		7.76	significant	5% chance of error						
	degrees of freedom	(rows -1) x (columns -1)			= 2						
		critical values									
		5%	5.99	7.76>5.99							
		1%	9.21	7.76<9.21							

Figure 7 Chi squared contingency table spreadsheet template.

3.5 Using correlation coefficients

Correlation coefficients have the H^0 that there is no correlation between two sets of paired data. The common learner misconception is that rejecting H^0 implies a causal relationship between the paired variables, when establishing a correlation gives no information about causal relationships or their direction. Establishing a correlation is just the prompt to experiment in search of a causal relationship.

Pearson's correlation is readily available in spreadsheets and quick to use. However, as a parametric method, it assumes a normal distribution for the paired sets of data, which is more likely if the sample size is larger. Using secondary data to demonstrate Pearson's r is possible and useful. However, pragmatically, non-parametric methods are more appropriate for the limited amount of primary data that can be collected given the current time constraints on practical work in schools and colleges in England. Fig. 8 shows a spreadsheet template for Spearman's Rank correlation requiring relatively few pairs of data. The example activity is a colorimeter investigation of the starch and chlorophyll in leaves. It is reasonable for AS/A-Level learners to argue that starch, the plant's storage carbohydrate synthesised from glucose, might be positively correlated to chlorophyll levels in leaves and therefore, their capacity for photosynthesis.

Another useful exercise compares the surface area and thickness of leaves in Privet (*Ligustrum ovalifolium*) bushes taken from different heights or depth in the bush. This can demonstrate phenotypic environmental tracking. It investigates the hypothesis that outer or higher leaves have larger surface area but are thinner than the inner, lower leaves on a bush. Theoretically, this would fully utilise the light incident on the bush. The higher outer leaves present a large surface area to the incident light to maximise photosynthesis. However, if the remaining light that passes through the upper leaves to reach the lower and inner parts of the bush strikes a leaf, it has a longer path to travel through the spongey mesophyll, maximising the probability it will be absorbed by chlorophyll.

% light transmission					
chlorophyll	starch	rank c	rank s	d	d2
40	50	1	1	0	0
18.5	21.5	3	2	1	1
18	21	4	3	1	1
17.5	20.5	5	4	1	1
17	20	6	5.5	0.5	0.25
1.5	20	11	5.5	5.5	30.25
16.5	19.5	7	7	0	0
16	19	8	8	0	0
20	18.5	2	9	-7	49
15	17	9	10	-1	1
12	12	10	11	-1	1
1	9	12	12	0	0
			totals	0	84.5
	r=1-((6*sumd2)/(n(n2-1)))				
		r=	0.704545		
Pairs	Critical values regardless of sign				
	5%	1%			
10	0.648	0.794		0.7>0.59	
12	0.591	0.777		sig 5% chance of error	
14	0.544	0.715			
16	0.506	0.665			

Figure 8 Spearman's rank correlation spreadsheet template.

4 CONCLUSIONS

The AQA AS/A-Level Biology specification [3] prescribes the selection and use of three methods for statistical analysis associated with A-Level content.

The statistical tests introduce learners to both parametric and non-parametric statistics.

AS level modules introduce statistical analysis through a consideration of experimental design, accuracy and precision, and the use of means and standard deviations to identify possible anomalous results.

The specification is less prescriptive regarding activities and required practical work to be used for statistical analysis. Only one exemplar involving χ^2 and mendelian genetics is written into both content statements and suggested skills development activities.

That said, there are ample opportunities for practicing statistical analyses associated with the content described in the AS and A-Level modules, and the associated suggestions for activities for skills development.

The specification gives little guidance on the selection and use of the statistical tests. Neither is guidance offered on reaching valid statistical conclusions by using a null hypothesis (H^0) and the internationally agreed standard of only when rejecting this when there is a 5% or less chance of error.

The assessment emphasis is not on remembering the formulae for statistical tests of significance, nor their calculation from instructions. However, the latter could be an approach used to satisfy other mathematical skills requirements within a statistical analysis examination question.

Activities used to practice the selection and use of statistical tests have been found to work well using teacher constructed spreadsheet templates that deal with time consuming and repetitive calculations for learners. This allows them to concentrate on the interpretation of the results.

An understanding of probability and statistical thinking is fundamental for AS/A-Level Biology learners and the AQA specification [3] could be more explicit in developing this through content, activities for skill development and required practical work.

REFERENCES

[1] Department for Education, *The Importance of Teaching: The Schools White paper 2010*, London: The Stationary Office, 2010.

[2] Parker, R., *Introductory statistics for biology*, Edward Arnold Ltd, 1971.

[3] AQA AS and A-Level Biology 7401, 7402 Exams from: 2016 (AS), 2017 (A-level), Version 1.5 26 November 2021 Available from: https://www.aqa.org.uk/subjects/science/as-and-a-level/biology-7401-7402.

[4] AQA GCSE Biology 8461, Exams from June 2018. Available from: https://www.aqa.org.uk/subjects/science/gcse/biology-8461

[5] Langley, R., Practical Statistics: for non-mathematical people, Lichfield: Pan Piper, 1968.

[6] Ellis, P. D., The essential guide to effect sizes: Statistical Power, Meta-Analysis, and the interpretation of Research Results, Cambridge University Press, 2010.

[7] British Educational Research Association (BERA), *Ethical Guidelines for Educational Research*, BERA, 2018.

[8] Tynan, R., "The importance of ecology fieldwork in English secondary school curricula," *EDULEARN23 Proceedings*, pp. 2789-2797, 2023.

[9] Tynan, R., "Using the school playing field to meet the practical requirements associated with ecology substantive knowledge in current biology examination specifications in England," *INTED2024 Proceedings*, pp, 705-414, 2024.

[10] https://www.field-studies-council.org/

[11] Sutton, D., *Trees of Britain and Europe*, Kingfisher, 1990.