Unit title: Statistics for Experimental Design

Unit code: **J/601/0297**

QCF level: **5**

Credit value: **15**

Aim

This unit provides learners with an understanding of the role of statistics in experimental design and hypothesis testing. Learners will be able to use significance testing to make statistical decisions and analyse the relationship between variables.

Unit abstract

The designing of scientific experiments involves multiple stages, often requiring an understanding of statistical analysis.

The unit starts with an overview of experimental design and examines the basic principles of sampling populations, probability and distributions. Learners will examine and address how statistical decisions form an important part of experimental design. Learners will gain an understanding of hypothesis testing, before looking at the differences between parametric and non-parametric models of analysis.

Understanding statistical decisions is extended to cover the role of significance testing, examining one, two and multiple sample tests. The unit concludes with correlation and linear regression; the mathematical processes are covered, as well as the impact of the limitations of correlation analysis on experimental design.

Emphasis throughout the unit is on a practical approach to applications familiar to learners together with an explanation of the theory underpinning the methods used.

Learning outcomes

On successful completion of this unit a learner will:

1 Understand the role of statistics in experimental design

2 Understand how statistical decisions are made using hypothesis testing

3 Be able to make statistical decisions using significance testing

4 Be able to analyse the relationship between variables.

Unit content

1 **Understand the role of statistics in experimental design**

Experimental design: treatments; controls and replicates; factorial designs; specialised designs eg clinical trials, ecological field studies, microbial assays

Population sampling: populations; sampling; sampling error; random sampling; random allocation; use of random number tables

Probability: concept of probability; laws of probability eg addition, multiplication; binomial approximation; Poisson approximation

Probability distributions: normal probability distribution; standard normal variate; central limit theorem; standard error; confidence limits; student's t distributions

2 **Understand how statistical decisions are made using hypothesis testing**

Hypothesis testing: the null hypothesis; significance (alpha) level; type 1 errors; type 2 errors; one tailed tests; two tailed tests; the power of a test; estimation of sample size

Parametric and non-parametric methods: assumptions of parametric analysis; the normal plot; transformation; non-parametric methods; selecting the correct test

3 **Be able to make statistical decisions using significance testing**

One and two sample tests: one sample z test; one sample t test; sign test; two unpaired sample tests; Wilcoxon ranked-sum test; two paired sample tests; Wilcoxon signed-rank test

Multiple sample tests: errors in multiple hypothesis testing; one-way analysis of variance (ANOVA); testing pairs of means; two-way ANOVA; interaction; factorial experimental design; randomised block; latin square; non-parametric; Kruskal-Wallis test

Categorical data: χ^2 goodness of fit test; χ^2 test for association; correction for continuity

4 **Be able to analyse the relationship between variables**

Linear correlation: scatter diagrams; Pearson's correlation coefficient; Spearman rank correlation coefficient; predicting values; assumptions of linear correlation; transformation; testing the significance of coefficients

Linear regression analysis: least squares method; regression equation; assumptions of linear regression; assessing the significance of slope coefficient and intercept; assessing the goodness of fit

Assessment of agreement: reproducibility and repeatability; numerical variables; limitations of correlation analysis; use of paired t tests; limits of agreement; analysis of differences; category variables; Cohen's kappa

Learning outcomes and assessment criteria

Learning outcomes On successful completion of this unit a learner will:		**Assessment criteria for pass** The learner can:	
LO1	Understand the role of statistics in experimental design	1.1	discuss the factors behind experimental design from a statistical view point
		1.2	explain the mechanics of population sampling with regards to controlling error
		1.3	evaluate probabilities using approximation methods
LO2	Understand how statistical decisions are made using hypothesis testing	2.1	assess the use of hypothesis testing in experimental design
		2.2	illustrate the differences between parametric and non-parametric models of analysis
LO3	Be able to make statistical decisions using significance testing	3.1	carry out one and two sample tests on a given population sample
		3.2	use multiple sample tests in experimental design
		3.3	use categorical data to test hypotheses
LO4	Be able to analyse the relationship between variables	4.1	use scatter diagrams to assess linearity with regression lines
		4.2	carry out linear regression analysis with all assumptions clearly indicated
		4.3	discuss how the assessment of agreement is utilised in experimental design

Guidance

Links

This unit has particular links with the following units within this qualification:

- *Unit reference number F/601/0220: Analysis of Scientific Data and Information*

Essential requirements

Delivery

The focus of delivery must be on a practical approach to the application of statistical analysis, although some theoretical background should be included to aid understanding, in particular of probability and parametric methods. There must be a balance between hand calculation (scientific calculator) and software use. Simple examples illustrating the method must be used to build up understanding and confidence before learners use software.

Learners must use the results of analyses as a basis for suggesting improvement in experimental design or extensions of the study in question. They must also be aware of potential errors, for example resulting from data entry, the wrong choice of analysis, misinterpretation and the limitations of the methods used.

Assessment

Learning outcome 1 involves demonstrating an understanding of how statistics are used in designing experiments. Some emphasis must be placed on understanding the role that sampling, probability and probability distributions have in design, rather than on mathematics. Evidence may be drawn from activities that form a normal part of the applied science programme of study. Learning outcome 2 involves understanding the use of hypothesis testing to make statistical decisions in experimental design. Emphasis must be placed on demonstrating understanding of the application of statistics, rather than the fundamental mathematics.

Learning outcome 3 covers the mathematical techniques used to test significance. Emphasis must be on the accurate application of the methods covered. Learning outcome 4 covers the mathematical techniques used to examine correlation between variables. Emphasis must be on the accurate application of the methods covered, with evidence showing how experiments are designed with correlation limitations in mind.

Resources

Learners will need access to IT facilities and appropriate software, to enable them to tackle realistic problems. Many of the operations relevant to applied science programmes can be implemented using a generic spreadsheet package (such as Microsoft Excel). Ideally, this will be supplemented by dedicated mathematical or statistical packages, for example Minitab, SPSS, PASW Statistics or MATLAB.

Employer engagement and vocational contexts

Learners will benefit from visits to industrial and research facilities to observe practical applications of data analysis, or to gain access to learning materials.

