Statistics for Police Analysts: course outline

The courses

The courses described here comprise a series of modules, which provide a theoretical and practical grounding in these areas. Modules relating to basic statistical theory and communication are strongly recommended for all analysts. It is expected that analysts will pick and choose between the toolkit modules according to their particular interests.

Structure

An integrated modular syllabus for the statistical education of analysts in policing and related fields is set out below. With one exception, the modules are aimed at analysts in areas such as performance, intelligence and survey-based consultation; the exception is Module A1, which is likely to be of interest to police officers and other managers who wish to develop their understanding of the role played by analytical thinking in informed decision making.

The modules are structured in the three blocks, as outlined in the overview.

The ideal length of each module is specified, although some may be ‘crammed’ into a shorter time if necessary. However, when this is done, analysts’ understanding and appreciation of the material will be diminished, unless the analyst has a strong background in mathematics.

Ongoing support and consultation is available for organisations whose analysts embark on any of these courses.

A. Essential statistical theory

Although the subject is often treated merely as a toolkit, there is much more to statistics than that. Statistics allow us to draw conclusions from data, thereby helping us to answer potentially important questions. Statistical theory is essentially a set of rules that governs the conclusions that can be drawn, and - importantly - those that cannot be drawn. The valid use of statistics depends on a confident understanding of these rules.

Module A1: Introduction to analysis and statistics

Broad objectives: To provide an introduction to the use of statistical (and other) evidence in the analysis of police performance, intelligence and consultation. It begins by addressing fundamental problems of interpretation, from which a formal model of analysis - and analytic thinking in general - is developed. The course then moves on to provide an introduction to data and the three main functions of statistical analysis, and provides an overview and basic introduction to some important statistical methods. This will cover:

- basic definitions and distinctions
- problems of interpretation (1) - the problem of randomness
- problems of interpretation (2) - heuristics and bias
- the nature of informed decision making
- a model for analysis
- the nature of data and the functions of statistics
• descriptive statistics (1) - frequencies and averages
• descriptive statistics (2) - the standard deviation
• inferential statistics, including applications of the standard deviation, trends and forecasting.

This course provides the essential introduction to the use of statistics for police analysts. It is also likely to be of interest to police decision-makers who wish to improve their understanding of analysis and their reading of statistical reports.

Length of course: 2 days.

Module A2: Statistical hypothesis testing

Broad objectives: To give analysts a thorough grounding in the theory and practice of hypothesis testing, to build confidence in this area so that analysts are able to make appropriate decisions and interpretations independently. This will cover:

• types of data
• probability distributions, illustrated through the binomial theorem
• normal distribution and using z tables
• essential concepts of hypothesis testing (null and alternative hypotheses, one- and two-tailed tests, Type I and II errors, power efficiency)
• statistical significance testing in principle and practice (illustrated with the chi squared test)

Length of course: 2 days ideally, 1 day as a crammer.

Module A3: Choosing statistical tests

Broad objectives: To provide analysts with a basic toolkit of statistical tests, with guidance on how to choose to apply them to answer different analytic questions. The course uses a series of exercises and real life examples to build analysts’ confidence in using a wider range of statistical techniques. The course on statistical hypothesis testing is a prerequisite for this course. The course does not deal with computation or interpretation of techniques, only with how to make appropriate choices. The course will cover:

• parametric and non-parametric tests
• tests of difference and independence
• choosing the right test for the job
• a basic ‘tool box’ of techniques
• exercises in choosing techniques

Length of course: 1.5 days ideally, 1 day as a crammer. Ongoing online support is available (further exercises and consultation).

B. The toolkit

All statistical tools have their uses, but they also have their limitations, situations in which they cannot be validly used. For example, the mean and standard deviation are very useful in certain circumstances; but in other circumstances they can be misleading, and should not be used. Statistical analysts need to have access to a wider range of techniques than they typically have, and they need to understand how to choose between them according to the sort of data they are analysing, and the questions they need to answer.

The modules in this section examine in detail a range of important and widely used techniques. Unlike the modules in Section A, they are not presented in the order in which they should be taken. Other modules can be developed on request.
Module B1: The standard deviation - potential, limitations and alternatives

Broad objectives: The standard deviation is widely used by police analysts for setting thresholds. Commonly - though mistakenly - called control limits, these thresholds are an informal application of the principles of statistical significance testing. Because the underlying principles are not widely understood, the method is often used inappropriately, and the statistics are often misinterpreted. This course explains these underlying principles, and introduces alternative methods for use with small samples, and low volume figures. This will cover:

• review of the standard deviation
• the limitations of control charts and Statistical Process Control
• alternatives to the standard deviation (1) - Student’s t distribution (small samples)
• alternatives to the standard deviation (2) - Poisson distribution (low volume figures)
• interpreting results against thresholds

Length of course: 1 day as a stand alone. Can be reduced to half a day if done in conjunction with another course, most appropriately statistical hypothesis testing.

Module B2: Confidence intervals and sample sizes

Broad objectives: To give analysts a deeper understanding of sampling, and why the normal distribution is so important, and to explore the potential uses of random sampling that go beyond the familiar survey-based examples. This will cover:

• the central limit theorem
• theoretical foundation of confidence intervals
• determining sample sizes
• finite population correction
• applications to categorical data (including surveys)
• applications to quantitative data
• confidence intervals for differences

Length of course: 2 days ideally, 1 day as a crammer, concentrating on survey data without the applications to quantitative data.

Module B3: Analysing frequency distributions - chi squared and beyond

Broad objectives: To deepen analysts’ understanding of chi squared and its various applications, to examine how it can be extended, what its limitations are, and a range of alternative approaches. This will cover

• the chi squared test in theory and practice
• Yates’s correction for continuity
• Fisher’s Exact probability test
• measures of effect size (phi, Cramer)
• chi squared goodness of fit for simple frequency distributions

Length of course: 1 day.

Module B4: Correlation

Broad objectives: To give analysts an understanding of the concept of correlation, the various ways it can be measured statistically, and how it should be interpreted. This will cover
• the concept of correlation and the correlation coefficient
• Pearson’s correlation for quantitative data
• coefficient of determination (r-squared)
• testing the significance of the correlation coefficient
• Spearman’s rank correlation
• Kendall’s correlation for weak ordinal data

Length of course: 1 day, although the material can potentially be trimmed and integrated in a two-day course on correlation and regression (see Module B5).

Module B5: Regression analysis: trends, forecasting and modelling

Broad objectives: To give analysts a confident grounding in the theory and practice of line fitting, and the derivation of other models. Analysts will learn how to fit, evaluate and interpret a regression line, and to take account of other factors, such as seasonal variations. This will cover

• the line of best fit and how it is defined mathematically
• fitting a trend line and finding the value of r-squared
• forecasts and other predictions
• the confidence interval for an estimate
• taking account of seasonal variation
• basic introduction to multiple regression

Length of course: 2 days ideally, 1 day as a crammer. Can be integrated with Module B4 as 2 day course on correlation and regression.

Module B6: Analysis of variance

Broad objectives: To introduce analysts to analysis of variance (ANOVA), which is an important technique in its own right, and has important applications in other areas of statistics, most notable in regression analysis and modelling. This will cover:

• principles of analysis of variance
• variance ratio (F) test
• simple (one way) analysis of variance
• higher order (two and three way) analysis of variance
• analysis of variance in modelling
• deriving r-squared from the ANOVA table

Length of course: 1 day.

Other bespoke toolkit modules

Modules on a range of other statistical topics can be developed and provided on request. Modules can also be provided on material relating to questionnaire surveys, including questionnaire design and sampling.

C. Communication

The most valid and insightful statistical analysis will not result in informed decisions if the conclusions are not clearly and accurately communicated to those who make the decisions. This is more difficult than it might first appear, and depends not only on the disciplined and appropriate use of language on the part of the analyst, but also on the analyst having a clear understanding of what can and cannot be said on the basis of the analysis.
Module C1: Presenting findings

Broad objective: To help analysts develop their skills and methods for presenting statistical (and other analytic) findings in performance, intelligence and other reports. This will cover

- report structure
- demarcation of results and interpretation
- writing style for findings and interpretation
- graphical presentation
- adapting to the audience

Length of course: 1 day, although a condensed version of the material can be combined with Module C2.

Module C2: Report writing workshop

Broad objective: To help analysts further develop their report writing and presentation skills through critical feedback on their own work, submitted in advance. Not for the faint-hearted!

Length of course: 1 day, although a condensed version of the material can be combined with Module C1.

Statistics for police analysts - what course participants have said

These courses have been provided to analysts in the British Transport Police and Lancashire Police. Here is what some of them have said:

My department includes crime and performance analysts and I was struck by the structured development of crime analysts and the limited development for performance teams. Based on this I worked closely with Malcolm Hibberd to develop a bespoke course to BTP performance teams that was over the course of a year. The team were able to learn new skills and also change their approach to performance management as the now understand some of the weaknesses to presenting and analysing performance data in certain ways. Involving Malcolm in the process revolutionised the way that BTP manage performance using more sophisticated methods and professionalising the role of a performance analyst. Whether you are a crime analyst or a performance one, I would highly recommend any course delivered by Malcolm.

Vanita Patel, Head of Analysis & Performance, British Transport Police

I unreservedly recommend all analysts to attend Malcolm’s statistical training. The aim of analysis should always be to inform the decision making process. Malcolm’s training, which is both interesting and career developing, will give you a whole new insight into concepts and techniques whereby analysis can be made more scientific, rational and objective. I learnt to recognise the weaknesses of some of the traditional ways in which analysts present and interpret data. I am also more informed about the policing debate surrounding performance and process control charts.

Paul Watson, Performance Analyst, British Transport Police

Malcolm’s statistics courses were much better than any of the statistics teaching I received at school or even university. It was easy to see how the various techniques could be used in my day to day work, thanks to plenty of engaging, real-world examples. Malcolm has great knowledge of how the use of statistics and targets in policing has changed over time. I left feeling more enthused about my work. I am now better at conducting scientifically rigorous analysis and feel more confident in explaining and defending my findings to senior police staff and officers.

Nick Mouat, Performance Analyst, British Transport Police
Having been on previous courses run by Malcolm I had no hesitation in contacting him to provide statistical analysis training to me and my team. Malcolm’s teaching style and detailed course structure ensure that knowledge and understanding is gained throughout the course through group discussions, presentation and practical exercises. Malcolm’s personality ensures that people are engaged in the training and open to new ways of thinking about analysis. We are now applying the knowledge and principles gained from the course to our future analytical work. A thoroughly recommended course which will be beneficial to individuals, groups and organisations.

Ryan Bretherton, Satisfaction & Analysis Manager, Lancashire Police

Malcolm delivers statistical training in an easy-listening and informed manner which is interspersed with practical exercises. This model of teaching provided a break from listening to give you time to use the information so that you could prove your understanding during the course. He was happy to return to any areas that needed affirmation and clarity. He added his personality and enthusiasm for the subject to his teaching and provided examples appropriate to our line of work. In contrast, the course is an intense one containing a lot of information so the additional reading / homework is recommended. In a word the course was “excellent”. Thank you.

Rebecca Eckersley, Intelligence Analyst Partnerships, Lancashire Police