Institute
and Faculty
of Actuaries

Subject CT3
Probability and Mathematical Statistics
Core Technical

Syllabus

for the 2018 exams

1 June 2017

Aim

The aim of the Probability and Mathematical Statistics subject is to provide a grounding in the aspects of statistics and in particular statistical modelling that are of relevance to actuarial work.

Links to other subjects

Subjects CT4 – Models and CT6 – Statistical Methods: use the statistical concepts and models covered in this subject. These are then developed further in other subjects in particular Subject ST1 – Health and Care Specialist Technical, Subject ST7 – General Insurance – Reserving and Capital Modelling Specialist Technical and Subject ST8 – General Insurance – Pricing Specialist Technical.

Objectives

On completion of the subject the trainee actuary will be able to:

(i) Summarise the main features of a data set (exploratory data analysis).

1. Summarise a set of data using a table or frequency distribution, and display it graphically using a line plot, a box plot, a bar chart, histogram, stem and leaf plot, or other appropriate elementary device.

2. Describe the level/location of a set of data using the mean, median, mode, as appropriate.

3. Describe the spread/variability of a set of data using the standard deviation, range, interquartile range, as appropriate.

4. Explain what is meant by symmetry and skewness for the distribution of a set of data.

(ii) Explain the concepts of probability.

1. Explain what is meant by a set function, a sample space for an experiment, and an event.

2. Define probability as a set function on a collection of events, stating basic axioms.

3. Derive basic properties satisfied by the probability of occurrence of an event, and calculate probabilities of events in simple situations.

4. Derive the addition rule for the probability of the union of two events, and use the rule to calculate probabilities.

5. Define the conditional probability of one event given the occurrence of another event, and calculate such probabilities.

6. Derive Bayes' Theorem for events, and use the result to calculate probabilities.

7. Define independence for two events, and calculate probabilities in situations involving independence.

(iii) Explain the concepts of random variable, probability distribution, distribution function, expected value, variance and higher moments, and calculate expected values and probabilities associated with the distributions of random variables.

1. Explain what is meant by a discrete random variable, define the distribution function and the probability function of such a variable, and use these functions to calculate probabilities.

2. Explain what is meant by a continuous random variable, define the distribution function and the probability density function of such a variable, and use these functions to calculate probabilities.

3. Define the expected value of a function of a random variable, the mean, the variance, the standard deviation, the coefficient of skewness and the moments of a random variable, and calculate such quantities.

4. Evaluate probabilities (by calculation or by referring to tables as appropriate) associated with distributions.

5. Derive the distribution of a function of a random variable from the distribution of the random variable.

(iv) Define a probability generating function, a moment generating function, a cumulant generating function and cumulants, derive them in simple cases, and use them to evaluate moments.

1. Define and determine the probability generating function of discrete, integer-valued random variables.

2. Define and determine the moment generating function of random variables.

3. Define the cumulant generating function and the cumulants, and determine them for random variables.

4. Use generating functions to determine the moments and cumulants of random variables, by expansion as a series or by differentiation, as appropriate.

5. Identify the applications for which a probability generating function, a moment generating function, a cumulant generating function and cumulants are used, and the reasons why they are used.

(v) Define basic discrete and continuous distributions, be able to apply them and simulate them in simple cases.

1. Define and be familiar with the discrete distributions: geometric, binomial, negative binomial, hypergeometric, Poisson and uniform on a finite set.

2. Define and be familiar with the continuous distributions: normal, lognormal, exponential, gamma, chi-square, t, F, beta and uniform on an interval.

3. Define a Poisson process and note the connection between Poisson processes and the Poisson distribution, and that a Poisson process may be equivalently characterised as: (1) the distribution of waiting times between events, (2) the distribution of process increments and (3) the behaviour of the process over an infinitesimal time interval.

4. Generate basic discrete and continuous random variables using simulation methods.

(vi) Explain the concepts of independence, jointly distributed random variables and conditional distributions, and use generating functions to establish the distribution of linear combinations of independent random variables.

1. Explain what is meant by jointly distributed random variables, marginal distributions and conditional distributions.

2. Define the probability function/density function of a marginal distribution and of a conditional distribution.

3. Specify the conditions under which random variables are independent.

4. Define the expected value of a function of two jointly distributed random variables, the covariance and correlation coefficient between two variables, and calculate such quantities.

5. Define the probability function/density function of the sum of two independent random variables as the convolution of two functions.

6. Derive the mean and variance of linear combinations of random variables.

7. Use generating functions to establish the distribution of linear combinations of independent random variables.

(vii) State the central limit theorem, and apply it.

1. State the central limit theorem for a sequence of independent, identically distributed random variables.

2. Apply the central limit theorem to establish normal approximations to other distributions, and to calculate probabilities.

3. Explain and apply a continuity correction when using a normal approximation to a discrete distribution.

(viii) Explain the concepts of random sampling, statistical inference and sampling distribution, and state and use basic sampling distributions.

1. Explain what is meant by a sample, a population and statistical inference.

2. Define a random sample from a distribution of a random variable.

3. Explain what is meant by a statistic and its sampling distribution.

4. Determine the mean and variance of a sample mean and the mean of a sample variance in terms of the population mean, variance and sample size.

5. State and use the basic sampling distributions for the sample mean and the sample variance for random samples from a normal distribution.

6. State and use the distribution of the t-statistic for random samples from a normal distribution.

7. State and use the F distribution for the ratio of two sample variances from independent samples taken from normal distributions.

(ix) Describe the main methods of estimation and the main properties of estimators, and apply them.

1. Describe the method of moments for constructing estimators of population parameters and apply it.

2. Describe the method of maximum likelihood for constructing estimators of population parameters and apply it.

3. Define the terms: efficiency, bias, consistency and mean squared error.

4. Define the property of unbiasedness of an estimator and use it.

5. Define the mean square error of an estimator, and use it to compare estimators.

6. Describe the asymptotic distribution of maximum likelihood estimators and use it.

(x) Construct confidence intervals for unknown parameters.

1. Define in general terms a confidence interval for an unknown parameter of a distribution based on a random sample.

2. Derive a confidence interval for an unknown parameter using a given sampling distribution.

3. Calculate confidence intervals for the mean and the variance of a normal distribution.

4. Calculate confidence intervals for a binomial probability and a Poisson mean, including the use of the normal approximation in both cases.

5. Calculate confidence intervals for two-sample situations involving the normal distribution, and the binomial and Poisson distributions using the normal approximation.

6. Calculate confidence intervals for a difference between two means from paired data.

(xi) Test hypotheses.

1. Explain what is meant by the terms null and alternative hypotheses, simple and composite hypotheses, type I and type II errors, test statistic, likelihood ratio, critical region, level of significance, probability-value and power of a test.

2. Apply basic tests for the one-sample and two-sample situations involving the normal, binomial and Poisson distributions, and apply basic tests for paired data.

3. Use a χ^2 test to test the hypothesis that a random sample is from a particular distribution, including cases where parameters are unknown.

4. Explain what is meant by a contingency (or two-way) table, and use a χ^2 test to test the independence of two classification criteria.

(xii) Investigate linear relationships between variables using correlation analysis and regression analysis.

1. Draw scatterplots for bivariate data and comment on them.

2. Define and calculate the correlation coefficient for bivariate data, explain its interpretation and perform statistical inference as appropriate.

3. Explain what is meant by response and explanatory variables.

4. State the usual simple regression model (with a single explanatory variable).

5. Derive and calculate the least squares estimates of the slope and intercept parameters in a simple linear regression model.

6. Perform statistical inference on the slope parameter in simple linear regression.

7. Calculate R^2 (coefficient of determination) and describe its use to measure the goodness of fit of a linear regression model.

8. Use a fitted linear relationship to predict a mean response or an individual response with confidence limits.

9. Use residuals to check the suitability and validity of a linear regression model.

10. State the usual multiple linear regression model (with several explanatory variables).

(xiii) Explain the concepts of analysis of variance and use them.

1. Describe the circumstances in which a one-way analysis of variance can be used.

2. State the usual model for a one-way analysis of variance and explain what is meant by the term treatment effects.

3. Perform a simple one-way analysis of variance.

(xiv) Explain the concepts of conditional expectation and compound distribution, and apply them.

1. Define the conditional expectation of one random variable given the value of another random variable, and calculate such a quantity.

2. Show how the mean and variance of a random variable can be obtained from expected values of conditional expected values, and apply this.

3. Derive the moment generating function of the sum of a random number of independent, identically distributed random variables (a compound distribution), and use the result to calculate the mean and variance of such a distribution.

END OF SYLLABUS