Institute of Actuaries of India

Subject
CT6 – Statistical Methods

For 2017
Examinations

Aim

The aim of the Statistical Methods subject is to provide a further grounding in mathematical and statistical techniques of particular relevance to financial work.

Links to other subjects

Subject CT3 – Probability and Mathematical Statistics: provides a grounding in probability and statistics.

Subject CA1 – Actuarial Risk Management – develops some of the concepts introduced in this subject.

Subject ST1 – Health and Care Specialist Technical , Subject ST7 – General Insurance – Reserving and Capital Modelling Specialist Technical, Subject ST8 – General Insurance – Pricing Specialist Technical and Subject ST9 – Enterprise Risk Management use the mathematics developed in this subject.

Objectives

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

(i)   Explain the concepts of decision theory and apply them.

     1.   Determine optimum strategies under the theory of games.

     2.   Explain what is meant by a decision function and a risk function.

     3.   Apply decision criteria to determine which decision functions are best with respect to a specified criterion. In particular consider the minimax criterion and the Bayes criterion.

(ii)   Calculate probabilities and moments of loss distributions both with and without limits and risk-sharing arrangements.

     1.   Describe the properties of the statistical distributions which are suitable for modelling individual and aggregate losses.

     2.   Derive moments and moment generating functions (where defined) of loss distributions including the gamma, exponential, Pareto, generalised Pareto, normal, lognormal, Weibull and Burr distributions.

     3.   Apply the principles of statistical inference to select suitable loss distributions for sets of claims.

     4.   Explain the concepts of excesses (deductibles), and retention limits.

     5.   Describe the operation of simple forms of proportional and excess of loss reinsurance.

6. Derive the distribution and corresponding moments of the claim amounts paid by the insurer and the reinsurer in the presence of excesses (deductibles) and reinsurance.

7. Estimate the parameters of a failure time or loss distribution when the data is complete, or when it is incomplete, using maximum likelihood and the method of moments.

(iii) Construct risk models involving frequency and severity distributions and calculate the moment generating function and the moments for the risk models both with and without simple reinsurance arrangements.

1. Construct models appropriate for short term insurance contracts in terms of the numbers of claims and the amounts of individual claims.

2. Describe the major simplifying assumptions underlying the models in 1.

3. Derive the moment generating function of the sum of N independent random variables; in particular when N has a binomial, Poisson, geometric or negative binomial distribution.

4. Define a compound Poisson distribution and show that the sum of independent random variables each having a compound Poisson distribution also has a compound Poisson distribution.

5. Derive the mean, variance and coefficient of skewness for compound binomial, compound Poisson and compound negative binomial random variables.

6. Derive formulae for the moment generating functions and moments of aggregate claims over a given time period for the models in 1. In terms of the corresponding functions for the distributions of claim numbers and claim amounts, stating the mathematical assumptions underlying these formulae.

7. Repeat 5. for both the insurer and the reinsurer after the operation of simple forms of proportional and excess of loss reinsurance.

(iv) Explain the concept of ruin for a risk model. Calculate the adjustment coefficient and state Lundberg's inequality. Describe the effect on the probability of ruin of changing parameter values and of simple reinsurance arrangements.

1. Explain what is meant by the aggregate claim process and the cash-flow process for a risk.

2. Use the Poisson process and the distribution of inter-event times to calculate probabilities of the number of events in a given time interval and waiting times.

3. Define a compound Poisson process and derive the moments and moment generating function for such a process.

4.   Define the adjustment coefficient for a compound Poisson process and for discrete time processes which are not compound Poisson, calculate it in simple cases and derive an approximation.

5.   Define the probability of ruin in infinite/finite and continuous/discrete time and state and explain relationships between the different probabilities of ruin.

6.   State Lundberg's inequality and explain the significance of the adjustment coefficient.

7.   Describe the effect on the probability of ruin, in both finite and infinite time, of changing parameter values.

8.   Analyse the effect on the adjustment coefficient and hence on the probability of ruin of simple reinsurance arrangements.

(v)   Explain the fundamental concepts of Bayesian statistics and use these concepts to calculate Bayesian estimators.

1.   Use Bayes' Theorem to calculate simple conditional probabilities.

2.   Explain what is meant by a prior distribution, a posterior distribution and a conjugate prior distribution.

3.   Derive the posterior distribution for a parameter in simple cases.

4.   Explain what is meant by a loss function.

5.   Use simple loss functions to derive Bayesian estimates of parameters.

6.   Explain what is meant by the credibility premium formula and describe the role played by the credibility factor.

7.   Explain the Bayesian approach to credibility theory and use it to derive credibility premiums in simple cases.

8.   Explain the empirical Bayes approach to credibility theory, in particular its similarities with and its differences from the Bayesian approach.

9.   State the assumptions underlying the two models in 8.

10.   Calculate credibility premiums for the two models in 8.

(vi)   Describe and apply techniques for analysing a delay (or run-off) triangle and projecting the ultimate position.

1.   Define a development factor and show how a set of assumed development factors can be used to project the future development of a delay triangle.

2.   Describe and apply the basic chain ladder method for completing the delay triangle.

3.   Show how the basic chain ladder method can be adjusted to make explicit allowance for inflation.

4.   Discuss alternative ways for deriving development factors which may be appropriate for completing the delay triangle.

5.   Describe and apply the average cost per claim method for estimating outstanding claim amounts.

6.   Describe and apply the Bornhuetter-Ferguson method for estimating outstanding claim amounts.

7.   Describe how a statistical model can be used to underpin a run-off triangles approach.

8.   Discuss the assumptions underlying the application of the methods in 1. to 7. above.

(vii)   Explain the fundamental concepts of a generalised linear model (GLM), and describe how a GLM may apply.

1.   Be familiar with the principles of Multiple Linear Regression and the Normal Linear Model

2.   Define an exponential family of distributions. Show that the following distributions may be written in this form: binomial, Poisson, exponential, gamma, normal.

3.   State the mean and variance for an exponential family, and define the variance function and the scale parameter. Derive these quantities for the distributions in 2.

4.   Explain what is meant by the link function and the canonical link function, referring to the distributions in 2.

5.   Explain what is meant by a variable, a factor taking categorical values and an interaction term. Define the linear predictor, illustrating its form for simple models, including polynomial models and models involving factors.

6.   Define the deviance and scaled deviance and state how the parameters of a GLM may be estimated. Describe how a suitable model may be chosen by using an analysis of deviance and by examining the significance of the parameters.

7.   Define the Pearson and deviance residuals and describe how they may be used.

8.   Apply statistical tests to determine the acceptability of a fitted model: Pearson's Chi-square test and the Likelihood ratio test

(viii)   Define and apply the main concepts underlying the analysis of time series models.

   1.   Explain the concept and general properties of stationary, $I(0)$, and integrated, $I(1)$, univariate time series.

   2.   Explain the concept of a stationary random series.

   3.   Explain the concept of a filter applied to a stationary random series.

   4.   Know the notation for backwards shift operator, backwards difference operator, and the concept of roots of the characteristic equation of time series.

   5.   Explain the concepts and basic properties of autoregressive (AR), moving average (MA), autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) time series.

   6.   Explain the concept and properties of discrete random walks and random walks with normally distributed increments, both with and without drift.

   7.   Explain the basic concept of a multivariate autoregressive model.

   8.   Explain the concept of cointegrated time series.

   9.   Show that certain univariate time series models have the Markov property and describe how to rearrange a univariate time series model as a multivariate Markov model.

   10. Outline the processes of identification, estimation and diagnosis of a time series, the criteria for choosing between models and the diagnostic tests that might be applied to the residuals of a time series after estimation.

   11. Describe briefly other non-stationary, non-linear time series models.

   12. Describe simple applications of a time series model, including random walk, autoregressive and cointegrated models as applied to investment variables.

   13. Develop deterministic forecasts from time series data, using simple extrapolation and moving average models, applying smoothing techniques and seasonal adjustment when appropriate.

(ix)   Explain the concepts of "Monte Carlo" simulation using a series of pseudo-random numbers.

   1.   Explain the disadvantages of using truly random, as opposed to pseudo-random, numbers.

   2.   Describe how pseudo-random drawings from specified distributions can be generated.

   3.   Explain the circumstances in which the same set of random numbers would be used for two sets of simulations and the circumstances in which different sets would be used.

4.   Discuss how to decide how many simulations to carry out in order to estimate a quantity of interest.

END OF SYLLABUS