Guide d'étude

Examen MAS-I: Modern Actuarial Statistics I
Casualty Actuarial Society (CAS)

Alec James van Rassel

Table des matières

Préliminaire

Information

Objectives

- Set forth, usually in broad terms, what the candidate should be able to do in actual practice ;
- The objectives include methodologies that may be impossible to perform on an exam that the candidate is expected to be able to explain conceptually ;
- *For example* : The Hat Matrix couldn't be calculated, but conceptual questions about it could be asked ;

Learning outcomes

1. It's important to identify some of the key terms, concepts, and methods associated with each of the learning objectives ;
2. They aren't an exhaustive list of the material being tested, but rather illustrate the scope of each learning objective ;

Information additionnelle

- The learning objectives define the behaviours and the knowledge statements illustrate more fully their intended scope ;
- Learning objectives should not be seen as independent units but as building blocks for our understanding ;
- The ranges are just guidelines ;
- The overall section weights should be seen as having more significance than the individual section weights ;
- Tables include :
  - values for the illustrative life tables ;
  - Standard normal distribution ;
  - Abridged inventories of discrete and continuous probability distributions ;
  - Chi-square distribution ;
  - t-distribution ;

  - F-distribution ;
- **There is a guessing adjustment** ;

Sujets à l'étude

1 Probability models (Stochastic Processes and Survival Models) (20% à 35%)

Information

Description

Notes du descriptif principal :

- Stochastic processes
- Survival models
  - Covered in depth as part of probability modeling in generic terms ;
- Markov Chains
  - Provide the means to model how an entity can move through different states ;
- Simplified version of life contingencies
  - Life contingencies problems can be viewed as discounted cash flow problems which include thee effect of probability of payment ;
  - Covered through a study note which link the generic survival model concepts to a subset of life actuarial concepts ;
  - This study note illustrates how to calculate annuities or single premium insurance amounts ;

Notes de la sous-section :

- Résoudre des problèmes de processus aléatoires ;
- Identifier les probabilités et distributions associées avec ces processus ;
  - Particulièrement, être capable d'utiliser un processus de Poisson dans ces applications ;

- Les modèles de survie sont une rallonge aux modèles de probabilité de processus stochastiques ;
  - En lieu, on estime la vie futur d'une entité avec quelques suppositions sur la distribution de la vraisemblance de survie ;
- Chaines de Markov utiles pour modéliser la mobilité entre états dans un processus et souligner les modèles Bayésien MCMC sous-jacent ;
- La simulation est incluse puisqu'elle peut s'avérer essentielle pour arriver à une solution de problème complexe ;

Learning objectives

1. Understand and apply the properties of Poisson processes ;
   - For increments in the homogeneous case ;
   - For interval times in the homogeneous case ;
   - For increments in the non-homogeneous case ;
   - Resulting from special types of events in the Poisson process ;
   - Resulting from sums of independent Poisson processes ;

   Knowledge Statements

   a. Poisson process ;
   b. Non-homogeneous Poisson process ;
   c. Memoryless property of Exponential and Poisson ;
   d. Relationship between Exponential and Gamma ;
   e. Relationship between Exponential and Poisson ;

2. For any Poisson process and the inter-arrival and waiting distributions associated with the Poisson process, calculate :
   - Expected values ;
   - Variances ;
   - Probabilities ;

Knowledge Statements

a. Probability calculations for Poisson process
b. Conditional distribution of arrival times ;
c. Splitting grouped Poisson rate to subsets of population using probability distribution ;
d. Conditional distribution of events by category within a group within a certain time period ;

3. For a compound Poisson process, calculate moments associated with the value of the process at a given time ;

Knowledge Statements

a. Compound Poisson process mean and variance ;
b. Normal approximation and hypothesis testing ;

4. Apply the Poisson process concepts to calculate the hazard function and related survival model concepts ;
   - Relationship between hazard rate, probability density function and cumulative distribution function ;
   - Effect of memoryless nature of Poisson distribution on survival time estimation ;

Knowledge Statements

a. Failure time random variables ;
b. Cumulative distribution functions ;
c. Survival functions ;
d. Probability density functions ;
e. Hazard functions and relationship to Exponential distribution ;
f. Relationships between failure time random variables in the functions above ;
g. Greedy algorithms ;

---

5. Given the joint distribution of more than one source of failure in a system (or life) and using Poisson Process assumptions :
   - Calculate probabilities and moments associated with functions of these random variables' variances ;
   - Understand differences between a series system (joint life) and parallel system (last survivor) when calculating expected time to failure or probability of failure by a certain time ;
   - Understand the effect of multiple sources of failure (multiple decrement) on expected system time to failure (expected lifetime) ;

Knowledge Statements

a. Joint distribution of failure times ;
b. Probabilities and moments ;
c. Time until failure of the system (life) ;
d. Time until failure of the system (life) from a specific cause ;
e. Time until failure of the system (life) for parallel or series systems with multiple components ;
f. Paths that lead to parallel or series system failures for systems with multiple components ;
g. Relationship between failure time and minimal path and minimal cut sets ;
h. Bridge system and defining path to failure ;
i. Random graphs and defining path to failure ;
j. Effect of multiple sources of failure (multiple decrements) on failure time calculations (competing risk) ;
k. Non-uniform probability of component failure (multiple decrement) ;
l. Method of inclusion and exclusion as applied to failure time estimates ;
m. Expected system lifetime as function of component lifetime and properties of expected lifetime estimates ;
n. Linkage between reliability function for a system and future expected lifetime ;

6. For discrete Markov Chains under both homogeneous and non-homogenous states :
   - Definition of a Markov Chain ;
   - Chapman-Kolmogorov Equations for n-step transition calculations ;

- Accessible states ;
- Ergodic Markov Chains and limiting probabilities ;

Knowledge Statements

a. Random Walk ;
b. Classification of states and classes of states (absorbing, accessible, transition, irreducible, and recurrent) ;
c. Transition step probabilities ;
d. Stationary probabilities ;
e. Recurrent vs. transient states ;
f. Gamblers ruin problem ;
g. Branching processes ;
h. Homogeneous transition probabilities ;
i. Memoryless property of Markov Chains ;
j. Limiting probabilities ;

---

7. Solve Life Contingency problems using a life table in a spreadsheet as the combined result of discount, probability of payment and amount of payment vectors. Understand the linkage between the life table and the corresponding probability models ;
   - Calculate annuities for discrete time ;
   - Calculate life insurance single net premiums (or P & C pure premiums) for discrete time ;
   - Solve for net level premiums (**not** including fractional lives) ;

Knowledge Statements

a. Discounted cash flow ;
b. Relationship between annuity values and insurance premiums ;
c. Life table linkage to probability models ;
d. Equivalence property ;

8. The candidate should be familiar with basic computer simulation methods.
   - Understand the basic framework of Monte Carlo Simulation ;
   - Understand the mechanics of generating uniform random numbers ;
   - Generate random numbers from a variety of distributions using the inversion method ;
   - Be able to explain when and how to use the Acceptance-Rejection method ;

Knowledge Statements

a. Random Number Generation ;
b. Uniform Random Numbers ;
c. Inversion Method ;
d. Acceptance-Rejection Method ;

Related lessons ASM

1. Probability Review
2. Parametric Distributions
3. Mixtures

---

4. Markov Chains : Chapman-Kolmogorov Equations
5. Markov Chains : Classification of States
6. Discrete Markov Chains : Long-Run Proportions and Limiting Probabilities
7. Markov Chains : Time in Transient States
8. Markov Chains : Branching Processes
9. Markov Chains : Time Reversible
10. Exponential Distribution

11. The Poisson Process : Probabilities of Events
12. The Poisson Process : Time To Next Event
13. The Poisson Process : Thinning, or Couting Special Types of Events
14. The Poisson Process : Other Characteristics
15. The Poisson Process : Sums and Mixtures
16. Compound Poisson Processes
17. Reliability : Structure Functions
18. Reliability : Probabilities
19. Reliability : Time to Failure

---

20. Survival Models
21. Contingent Payments

---

22. Simulation—Inverse Transformation Method
23. Simulation—Applications
24. Simulation—Rejection Method

Vidéos YouTube

›

Résumés des chapitres

Probability Review

Chapter 1: Probability Review

Introduction to Mathematical Statistics 1 - 3, 5

›

Chapter 2: Parametric Distributions

Introduction to Mathematical Statistics 2.2, 2.7 Nonlife Actuarial Models—Theory Methods and Evaluation 2.2

›

Chapter 3: Mixtures

Introduction to Mathematical Statistics 3.7 Nonlife Actuarial Models—Theory Methods and Evaluation 2.3.2

›

Stochastic Processes

Chapter 4: Markov Chains : Chapman-Kolmogorov Equations

Ross 4.1 - 4.2, 4.5.1 - 4.5.2

›

Chapter 5: Markov Chains : Classification of States

Ross 4.3

›

Chapter 6: Discrete Markov Chains : Long-Run Proportions and Limiting Probabilities

Ross 4.4

›

Chapter 7: Markov Chains : Time in Transient States

Ross 4.6

›

Chapter 8: Markov Chains : Branching Processes

Ross 4.7

›

Chapter 9: Markov Chains : Time Reversible

Ross 4.8

›

Chapter 10: Exponential Distribution

Ross 5.2

›

Chapter 11: The Poisson Process : Probabilities of Events

Ross 5.3.1 - 5.3.2 Daniel Poisson Study Note 1.1, 1.4.1

›

Chapter 12: The Poisson Process : Time To Next Event

Ross 5.2, 5.3.3 Daniel Poisson Study Note 1.1.1

›

Chapter 13: The Poisson Process : Thinning, or Couting Special Types of Events

Ross 5.3.4 Daniel Poisson Study Note 1.3.1, 1.4.3

›

Chapter 14: The Poisson Process : Other Characteristics

Ross 5.2.3, 5.3.4, 5.3.5

›

Chapter 15: The Poisson Process : Sums and Mixtures

Ross 5.4.3 Daniel Poisson Study Note 1.3.2, 1.3.3

›

Chapter 16: Compound Poisson Processes

Ross 5.4.2 Daniel Poisson Study Note 1.2, 1.4.2

›

Chapter 17: Reliability : Structure Functions

Ross 9.1 - 9.2

›

Chapter 18: Reliability : Probabilities

Ross 9.3 - 9.4

›

Chapter 19: Reliability : Time to Failure

Ross 9.5 - 9.6

›

Life Contingencies

Chapter 20: Survival Models

Struppeck 1, 2, 6, 7

›

Chapter 21: Contingent Payments

Struppeck 3, 4, 5, 6

›

Simulation

Chapter 22: Simulation—Inverse Transformation Method

Ross 11.1, 11.2.1

›

Chapter 23: Simulation—Applications

Ross 11.1, 11.2.1

›

Chapter 24: Simulation—Rejection Method

Ross 11.2.2

›

Notes sur les vidéos YouTube

StatQuest: A Gentle Introduction to Machine Learning

›

2 Statistics (15% à 30%)

Information

Description

Notes du descriptif principal :

- Topics which would commonly be covered in a 2-semester Probability & Statistics sequence ;

Learning objectives

1. Perform point estimation of statistical parameters using Maximum likelihood estimation (MLE).
   Apply criteria to estimates such as :
   - Consistency ;
   - Unbiasedness ;
   - Sufficiency ;
   - Efficiency ;
   - Minimum variance ;
   - MSE ;

Knowledge Statements

a. Equations for MLE of mean, variance from a sample ;

b. Estimation of mean and variance based on samples ;

c. General equations for MLE of parameters ;

d. Recognition of consistency property of estimators and alternative measures of consistency ;

e. Application of criteria for measurement when estimating parameters through minimisation of variance, MSE ;

f. Definition of statistical bias and recognition of estimators that are unbiased or biased ;

g. Application of Rao-Cramer Lower Bound and Efficiency ;

h. Relationship between Sufficiency and Minimum Variance ;

i. Develop and estimate a sufficient statistic for a distribution ;

j. Factorization Criterion for sufficiency ;

k. Application of Rao-Cramer Lower Bound and Fisher Information ;

l. Application of MVUE for the exponential class of distributions ;

m. Linkage between Score Function, Fisher Information and maximum likelihood ;

n. Method of Moments ;

o. Percentile Matching ;

p. Kernel Density Estimation ;

q. Maximum Likelihood with Censoring and Truncation ;

Learning objectives

2. Calculate parameter estimates using methods other than maximum likelihood.
3. Test statistical hypotheses including Type I and Type II errors using :
   - Neyman-Pearson theorem ;

     Apply Neyman-Pearson theorem to construct likelihood ratio equation ;
   - Likelihood ratio tests ;
   - First principles ;

     Use critical values from a sampling distribution to test means and variances ;

Knowledge Statements

a. Presentation of fundamental inequalities based on general assumptions and normal assumptions ;

b. Definition of Type I and Type II errors ;

c. Significance levels ;

d. One-sided versus two-sided tests ;

e. Estimation of sample sizes under normality to control for Type I and Type II errors ;

f. Determination of critical regions ;
g. Definition and measurement of likelihood ratio tests ;
h. Determining parameters and testing using tabular values (from a table) ;
i. Recognizing when to apply likelihood ratio tests versus chi-square or other goodness of fit tests ;
j. Apply paired t-test to two samples ;
k. Test for difference in variance under Normal distribution between two samples through the application of F-test ;
l. Test of significance of means from two samples under Normal distribution assumptions in both large and small sample cases ;
m. Test for significance of difference in proportions between two samples under the Binomial distribution assumption in both large and small sample cases ;
n. Application of contingency tables to test independence between effects ;
o. Asymptotic relationship between likelihood ratio tests and the Chi-Square distribution ;
p. Application of Neyman-Pearson theorem to Uniformly Most Powerful hypothesis tests ;
q. Equivalence between critical regions and confidence intervals ;
r. Kolmogorov-Smirnov test ;

Learning objectives

4. For the Exponential, Gamma, Weibull, Pareto, Lognormal, Beta, and mixtures thereof :
   - Identify the applciations to Insurance claim modeling in which each distribution is used and reasons why ;
   - Transformation of distributions ;

Knowledge Statements

a. Frequency, severity and aggregate loss ;
b. Common continuous distributions for modeling claim severity ;
c. Mixing distributions ;
d. Tail properties of claim severity ;
e. Effects of coverage modifications including, for example : limits, deductibles, loss elimination ratios and effects of inflation ;

Learning objectives

5. Calculate Order Statistics of a sample for a given distribution.

Knowledge Statements

a. General form for distribution of n^{th} largest element of a set ;
b. Application to a given distributional form ;

Related lessons ASM

25. Estimator Quality
26. Kernel Density Estimation
27. Method of Moments
28. Percentile Matching
29. Maximum Likelihood Estimators
30. Maximum Likelihood Estimators—Special Techniques
31. Variance of Maximum Likelihood Estimator
32. Sufficient Statistics
33. Hypothesis Testing
34. Confidence Intervals and Sample Size
35. Confidence Intervals for Means
36. Kolmogorov-Smirnov Tests
37. Chi Square Tests
38. Confidence Intervals for Variances

39. Uniformly Most Powerful Critical Regions
40. Likelihood Ratio Tests
41. Q.-Q. Plots

Vidéos YouTube

-

Likely Questions

- Question where we calculate the sample variance with the STAT function of the calculator ;
  – MAS-I F19, # 15 ;

Résumés des chapitres

Chapter 25: Estimator Quality

Introduction to Mathematical Statistics 4.1.3, 5.1, 7.1

Overview of various basic functions to evaluate the quality of an estimator including :

1. The Bias of an estimator (incl. asymptotically unbiased).
   - Shows the sample mean $\bar{x}$ is an unbiased estimator of the true mean μ.
   - Shows the sample variance $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$ is an unbiased estimator of the true variance σ^2.
   - Shows the empirical variance $\hat{\sigma}^2 = \frac{\sum(x_i - \bar{x})^2}{n}$ is an biased estimator of the true variance σ^2.
     However, with a bias of $\frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$, it is *asymptotically* unbiased.
   - Gives the boiling water feet-freezer head analogy for why the bias is not sufficient to assess the quality of an estimator.
2. The consistency of an estimator.

- As the number of observations n on which an estimator $\hat{\theta}_n$ is based increases, if both the bias and variance of $\hat{\theta}_n$ go to 0, we can say it is a consistent estimator.
  Although this condition is sufficient, it is not necessary.
- For example, the sample mean $\bar{x}$ is a consistent estimator of the true mean for both the Gamma and Normal distributions.
  However, a Pareto distribution with $\alpha \leq 2$ has an infinite variance ; therefore, the variance of the estimator is infinite and $\bar{x}$ is not a consistent estimator.

3. The efficiency of an estimator.
   - Shows the relative efficiency of one estimator to another.
4. The Mean Square Error (MSE) of an estimator.
   - Shows both how it's defined as the variance of the estimator's predictions of the parameter and the relation with bias and variance.

We then combine these concepts to define the **uniformly minimum variance unbiased estimator (UMVUE)** which has a smaller variance, for any true value θ, than any other unbiased estimator.

Chapter 26: Kernel Density Estimation

Nonlife Actuarial Models : Theory Methods and Evaluation 11.1

-

Chapter 27: Method of Moments

Nonlife Actuarial Models : Theory Methods and Evaluation 12.1.1

When we estimate the parameters for the underlying distribution, the question itself does not impact our estimates.

- For example, if we have data from a coverage with a policy limit of 100$ and want to find the average payment per loss with a deductible of 100$, then we estimate the parameters considering the policy limit and not the deductible.

- This sounds obvious, but it is quite easy to make the mistake.

Complete data Given the exact value of each observation.

Types of incomplete data :

Grouped data Given a set of intervals and told how many observations are in each.

- For example :

Range	Number of observations in range
0-999	85
1000-1999	64
2000 and over	72

Censored data Given that the value of an observation is in a range, but not given the exact value.

- For example, a policy with a limit of 10 000\$ ($Y = \min(X, 10\,000)$).

Truncated data Given the value of an observation only when it is in a certain range ; typically, only above or below a certain number.

- For example, a policy with a deductible of 100\$ has no recorded losses of 100\$ or less ($Y = \{X|X > 100\}$).

To fit a k parameter distribution, we set equal the k first sample moments $\hat{\mu}_k$ to the k first raw moments μ'_k of the distribution.
We may match the variances instead of the second moments but we match the biased empirical variance $\hat{\sigma}^2$ by default and not s^2.

$$\hat{\mu}_j = \mu'_j, \; j = 1, \ldots, k$$

If the data is incomplete, we match the moments of the corresponding distribution.

- If data is censored at u, $\hat{\mu}_k = \mathrm{E}[\min(X; u)^k]$.
- If data is truncated at d, $\hat{\mu}_k = \mathrm{E}[X^k|X > d]$.

Chapter 28: Percentile Matching

Nonlife Actuarial Models : Theory Methods and Evaluation 11.1.1, 12.1.2

Smoothed empirical estimate

We use the sample's order statistics $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ for the following interpolation :

$$\hat{\pi}_g = (1-h)x_{(j)} + hx_{(j+1)}, \quad \text{where}$$
$$j = \lfloor (n+1)g \rfloor \quad \text{and} \quad h = (n+1)g - j$$

- Exams should specify which percentiles to use.
- Percentile matching itself matches $F(\hat{\pi}_g) = g$.
- Exams don't typically ask a lot of percentile matching questions, thus it's not really worth memorizing each distribution's formulas as it can easily be about a random distribution.

- With censored data, the only additional consideration is that the percentile must be within the range of the uncensored portion of the data.
- With truncated data, we match the percentile of the conditional distribution $(X|X > d)$.

Chapter 29: Maximum Likelihood Estimators

Introduction to Mathematical Statistics 4.1, 6.1
Nonlife Actuarial Models : Theory Methods and Evaluation 10.2, 12.3

Whatever the type of data, maximise the function representing the probability of the observation occuring $\boxed{\mathcal{L}(\theta) = \prod g(x_i, \theta)}$.

Therefore $g(x_i, \theta) =$
$f(x_i)$ for complete individual observations.

$F(x_i) - F(x_{i-1})$ for grouped observations.

$S(u)$ for observations censored at u (for those below u, $g(x_i, \theta) = f(x_i)$).

$\frac{f(x_i)}{S(d)}$ for observations truncated at d.

Chapter 30: Maximum Likelihood Estimators—Special Techniques

Introduction to Mathematical Statistics 4.1, 6.1
Nonlife Actuarial Models : Theory Methods and Evaluation 10.2, 12.3

Few shortcuts :
If the likelihood function is of the form $\boxed{\mathcal{L}(\gamma) = \gamma^{-a} \mathrm{e}^{-b/\gamma}}$ then $\boxed{\hat{\gamma}^{\text{MLE}} = \frac{b}{a}}$.

If the likelihood function is of the form $\boxed{\mathcal{L}(\lambda) = \lambda^{a} \mathrm{e}^{-\lambda b}}$ then $\boxed{\hat{\lambda}^{\text{MLE}} = \frac{a}{b}}$.

If the likelihood function is of the form $\boxed{\mathcal{L}(\theta) = \theta^{a} (1-\theta)^{b}}$ then $\boxed{\hat{\theta}^{\text{MLE}} = \frac{a}{a+b}}$.

- The types of data are explained in 10.2 of Tse.
- Applying MLE techniques to incomplete data is covered in 12.3 of Tse.

Chapter 31: Variance of Maximum Likelihood Estimator

Introduction to Mathematical Statistics 6.2, 6.5

- The Rao-Cramér Lower Bound and the Fisher information matrix are the main subjects.
- The book is very mathematical for this section and it’s difficult to understand.

Chapter 32: Sufficient Statistics

Introduction to Mathematical Statistics 7

Chapter 33: Hypothesis Testing

Introduction to Mathematical Statistics 4.5, 4.6

- Description of hypothesis tests (null vs alternative) and the types (bilateral vs unilateral and simple vs composite).
- Description of the critical region.
- Description of the types of errors.
- Description of the power of a test.

Chapter 34: Confidence Intervals and Sample Size

Introduction to Mathematical Statistics 4.5, 4.6

- Covered the basics of sample size, the confidence interval, the confidence coefficient.
- Distinction between point estimates vs interval estimates.

Chapter 35: Confidence Intervals for Means

Introduction to Mathematical Statistics 4.2

- Reminder of the Chi-Square distribution.
- Reminder of the T distribution.
- Difference of mean confidence intervals.

Chapter 36: Kolmogorov-Smirnov Tests

Nonlife Actuarial Models : Theory Methods and Evaluation 13.2.1

Chapter 37: Chi Square Tests

Introduction to Mathematical 4.7
Nonlife Actuarial Models : Theory Methods and Evaluation 13.2.3

Chapter 38: Confidence Intervals for Variances

Introduction to Mathematical Statistics 8.3

Chapter 39: Uniformly Most Powerful Critical Regions

Introduction to Mathematical Statistics 8.1 - 8.2

Chapter 40: Likelihood Ratio Tests

Introduction to Mathematical Statistics 8.3

Chapter 41: Q.-Q. Plots

Introduction to Mathematical Statistics 4.4
Larsen Study Note

Notes sur les vidéos YouTube

StatQuest: A Gentle Introduction to Machine Learning

>

3 Extended Linear Models (30% à 50%)

Information

Description

Notes du descriptif principal :

- Include GLMs which are commonly used to construct classification plans ;
- OLS model is covered as one member of the exponential family ;
- `R` is useful to better visualise and conceptualise the material ;

Notes de la sous-section :

- OLS treated as *one* type of model that may be used when the dependant variable follows the Normal distribution and the observations are (iid) with a constant variance ;
- All models assume data is (iid) from the exponentional family ;
- Assume linear relationship between dependant and independant variables ;
- Assume variance is a function of the mean ;
- **VIF** formula found on p. 102 of James et al. and p. 101 of Dobson is used and not hte one on p. 101 of James and et al. ;

$$\text{VIF}(b_j) = \frac{1}{1 - R^2_{(j)}}$$

- Questions may contain parameter tables and plots (of the type shown in texts) with which we should familiarise ourselves ;

Learning objectives

1. Understand the assumptions behind different forms of the Extended Linear Model and be able to select the appropriate model from list

below :

- OLS ;
- GLM ;
- ANOVA ;
- GAM ;
- Local Regression ;
- Lasso ;
- Ridge Regression ;
- Partial Least Squares ;
- PCA regression ;

Knowledge Statements

a. Understand the relationship between mean and variance by model family member for the exponential distribution ;

b. Understand how to select the appropriate distribution function for the dependent variable and the implication for the appropriate model form ;

c. Link functions (Identity, Log, Logit, Power, Inverse) ;

d. Characteristics of Exponential Family (Binomial, Normal, Exponential, Gamma, Poisson, Inverse Gaussian, Negative Binomial, and Tweedie) ;

e. Canonical Forms of link function and effect of non-canonical link function on bias ;

f. Penalized Regression as implemented using the Lasso or Ridge Regression ;

g. Understand concept of models within models for Generalized Additive Models (GAM) ;

h. Understand dimension reduction using Partial Least Squares (PLS) or Principal Components Regression (PCR) ;

2. Evaluate models developed using Extended Linear Model approach ;

Knowledge Statements

a. Raw or studentized Residuals ;
b. R^2 ;
c. Cook's Distance and outliers ;
d. Influential points ;
e. Leverage ;
f. AIC and BIC penalized log likelihood measures ;
g. Standardized / Studentized Residuals ;
h. Deviance, Deviance Residuals and relationship to likelihood ;
i. Pearson Residuals vs. Deviance residuals ;
j. Scatter, QQ and Box Plots ;
k. Type III Sequential Chi-Square test ;
l. T-test and Wald test for significance of regression coefficients ;
m. Prediction intervals for response variable ;
n. MSE and standard error ;
o. Calculation and validity of F test to compare 2 models (under OLS) ;
p. Cross-Validation ;
q. Test vs. Train Error ;
r. Bootstrapping to test model validity ;
s. Prediction vs. Forecast Error ;
t. Overfitting ;
u. Bias-Variance Tradeoff ;
v. Evaluate collinearity using VIF ;
w. Evaluate appropriateness of underlying assumptions including homoscedasticity and autocorrelation of residuals ;

3. Understand the algorithms behind the numerical solutions for the different forms of the Extended Linear Model family to enable interpretation of output from the statistical software employed in modeling and to make appropriate modeling choices when selecting modeling options ;

Knowledge Statements

a. Maximum Likelihood and OLS ;
b. Fisher Scoring (iterative weighted least squares as implemented using the Information and Score functions from section B. 1) ;
c. Quasi-Likelihood and relationship to maximum likelihood ;
d. Collinearity (Aliasing) and model stability ;
e. Hat matrix $\boldsymbol{H}$;
f. Design matrix $\boldsymbol{X}$;
g. Fitting adjoining, overlapping observations in groups for Local Regression ;
h. Supervised vs. Unsupervised learning methods ;
i. Modeling functions within functions for GAMs ;
j. Penalty function in Penalized regression models (Lasso and Ridge Regression) ;
k. PLS supervised learning vs. PCR unsupervised learning ;

4. Understand and be able to select the appropriate model structure for an Extended Linear Model given the behavior of the data set to be modeled ;

Knowledge Statements

a. Predictor variables ;
b. Response variables ;
c. Regression through the origin ;
d. Transformation of variables ;
e. Categorical vs. continuous explanatory variables ;
f. Interaction terms ;
g. Significance and model comparison statistics ;
h. Residuals and model parameter selection ;
i. Piecewise linear and Smoothing Splines ;
j. Smoothing parameter for splines ;
k. Basis Functions ;
l. Knot Selection for Splines ;
m. Weighting function for local regression ;
n. Selection of functions within functions for GAMs ;
o. Selection of appropriate tuning factor for Lasso or Ridge Regression ;
p. Select either Lasso or Ridge Regression depending on desired effect from penalized regression ;
q. Curse of High Dimensionality ;
r. Forward or backward or best subset selection ;

Related lessons ASM

42. Introduction to Extended Linear Models
43. How a Generalized Linear Model Works
44. How a Generalized Linear Model Works : Categorical Response
45. Generalized Linear Model : Estimating Parameters
46. Generalized Linear Model : Measures of Fit
47. Normal Linear Model : Standard Error, R-squared, and t-statistic

48. Normal Linear Model : F and VIF
49. Normal Linear Model : Validation
50. Normal Linear Model : Predictions
51. ANOVA
52. Generalized Linear Model : Measures of Fit II
53. Resampling Methods
54. Normal Linear Model : Subset Selection
55. Normal Linear Model : Shrinkage and Dimension Reduction
56. Extensions to the Linear Model

Vidéos YouTube

›

Résumés des chapitres

Chapter 42: Introduction to Extended Linear Models

Introduction to Mathematical Statistics 4.4 Introduction to Generalized Linear Models 2 Introduction to Statistical Learning with R 2 (except 2.2.3)

›

Chapter 43: How a Generalized Linear Model Works

Introduction to Generalized Linear Models 3.9 Introduction to Statistical Learning with R 3.3.1 - 3.3.2 Larsen Study Note

›

Chapter 44: How a Generalized Linear Model Works : Categorical Response

Introduction to Generalized Linear Models 7, 8 Introduction to Statistical Learning with R 4

›

Chapter 45: Generalized Linear Model : Estimating Parameters

Introduction to Generalized Linear Models 4 Introduction to Statistical Learning with R 3.1.1, 3.2.1 Larsen Study Note

›

Chapter 46: Generalized Linear Model : Measures of Fit

Introduction to Generalized Linear Models 5

›

Chapter 47: Normal Linear Model : Standard Error, R-squared, and t-statistic

Introduction to Generalized Linear Models 6.1 - 6.3 Introduction to Statistical Learning with R 3.1.2 - 3.1.3, 3.2.2 - 3.2.3

›

Chapter 48: Normal Linear Model : F and VIF

Introduction to Generalized Linear Models 6.2 Introduction to Statistical Learning with R 3

›

Chapter 49: Normal Linear Model : Validation

Introduction to Generalized Linear Models 6.2 Introduction to Statistical Learning with R 3.3.3

›

Chapter 50: Normal Linear Model : Predictions

›

Chapter 51: ANOVA

Introduction to Generalized Linear Models 6.4 - 6.5, 9.3 - 9.7 Introduction to Mathematical Statistics 9.1 - 9.5

›

Chapter 52: Generalized Linear Model : Measures of Fit II

Introduction to Generalized Linear Models 7, 8, 9

›

Chapter 53: Resampling Methods

Introduction to Statistical Learning with R 5

›

Chapter 54: Normal Linear Model : Subset Selection

Introduction to Statistical Learning with R 6.1

›

Chapter 55: Normal Linear Model : Shrinkage and Dimension Reduction

Introduction to Statistical Learning with R 6.2 - 6.4

›

Chapter 56: Extensions to the Linear Model

Introduction to Statistical Learning with R 7

›

Notes sur les vidéos YouTube

StatQuest: A Gentle Introduction to Machine Learning

>

4 Time Series with Constant Variance (10% à 20%)

Information

Description

Notes du descriptif principal :

- Covers an introduction to modeling activity, such as financial results or stock prices, over time ;
- The model used is the Auto Regressive Integrated Moving Average (ARIMA) where activity in a given period may be linked to activity in subsequent time periods ;
- The connection between adjacent time periods violates one of the assumptions behind the Extended Linear Model techniques ;
- The ARIMA appproach incorporates that linkage as an aid for it's predictions ;
- Also covers the application of regression models to time series analysis ;

Notes de la sous-section :

- Section covers basic applications of the ARIMA time series model ;

Learning objectives

1. Use time series to model trends ;
   - Estimation, data analysis, and forecasting ;
   - Forecast errors and confidence intervals ;

Knowledge Statements

(a) Mean-reverting time series ;

(b) Elimination of trends using differencing ;

(c) Relationship between seasonality and autocorrelation ;

2. Model relationships of current and past values of a statistic / metric ;
   - Estimation, data analysis, and forecasting ;
   - Forecast errors and confidence intervals ;

Knowledge Statements

(a) Calculation and use of lag k autocorrelation statistic and cross correlation statistics in determining model structure ;

(b) Stationary series ;

(c) AR(1) models ;

(d) ARIMA models—AR(p), MA(q), ARMA(p, q), ARIMA(p, d, q), ARMA vs ARIMA ;

(e) Invertible time series and relationship between AR and MA models ;

(f) Converting between AR and MA models ;

(g) Interpretation of auto-correlation function as aid to model selection (AR vs. MA and number of lags to include in model) ;

(h) Relationship between time series input and item modeled for AR vs. MA ;

3. Understand forecasts produced by ARIMA ;

Knowledge Statements

(a) Forecast using ARIMA models ;
(b) **One** step ahead *prediction* vs **many** step ahead *projection* ;
(c) Change in variance in prediction by AR vs MA model ;

---

4. Time Series with Regression ;

Knowledge Statements

(a) Deterministic vs. Stochastic Trend ;
(b) Serial correlation in regression error results ;
(c) Correction in regression via Generalized Least Squares ;
(d) Transformation of data using natural logarithms for regression modelling ;
(e) Forecast error correction under natural logarithms transformation ;

Related lessons ASM

57. Time Series : Trend and Seasonality
58. Time Series : Correlation
59. Time Series : White Noise and Random Walks
60. Time Series : Autoregressive Models
61. Time Series : Regression
62. Time Series : Moving Average Models
63. Time Series : ARMA Models
64. Time Series : ARIMA and SARIMA Models

Vidéos YouTube

›

Résumés des chapitres

57. Time Series : Trend and Seasonality

Introduction to Time Series with `R` 1

1.1 : Time Series Data—Purpose ;

- Understand the past and predict the future ;
- Usage by Singapore Airlines to increase their fleet based on future goals (expansion) and past trends (growing popularity) ;
- Usage by gas suppliers in the UK who must purchase the supply the day before it's demanded from offshore fields ;

  The price thus varies according to the temperature, the time of year, etc. and even the wind temperature.

1.2 : Time Series Data—Time Series ;

- Idea that variables are measured sequentially in time everywhere ;

  Banks record interest and exchange rates on a daily basis ;

  Governments measure GDP on a yearly basis ;
- Time series model, a.k.a. **discrete-time stochastic process**, is a **sequence** of *random variables* defined at **fixed sampling intervals** ;
- Main **features** of time series are **trends** and **seasonal variations** ;
- Another important feature is that most observations which are close together in time tend to be correlated *(serially dependent)* ;
- This can be applied to making predictions and test whether, for example, fluctuations in sales are due to an underlying change or just regular variation ;
- Sampling intervals can differ in their relation to data.
  - The data could have been aggregated over the period (for example, the number of foreign tourists arriving per day) ;
  - The data could have been sampled over the period (for example, the daily close price of Apple stock).

1.3 : Time Series Data—`R language` ;

1.4 : Time Series Data—Plots, trends, and seasonal variation ;

- **Trend** : systemic change in a time series which doesn't appear to be periodic ;

  The simplest model for trends is a linear (de)increase and is often an adequate approximation ;

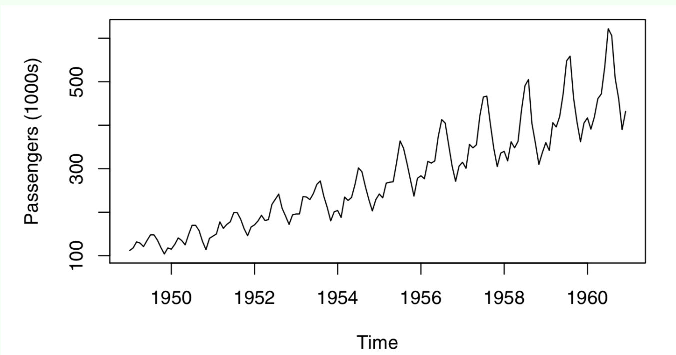
- **Seasonal variation** : Repeating pattern within each year ;

  For example, restaurant bookings varying according to the week-day ;

- **Cycles** : do not correspond to some **fixed** natural period

  For example, the economic cycle ;

  No cycles are apparent in the image below :

- The trend has many possible explanations (post-WWII prosperity, cheaper flights, etc.) but if there are non then it would be a **stochastic trend** ;

  In this case, regression's not appropriate ;

- Forecasting uses extrapolation supposing trends continue at a slow pace ;

- Therefore, linear extrapolation is a reasonable approximation for a **few steps** ahead ;

  There's no empirical way to verify this, so we must explain the trend to justify extrapolation ;

- Forecasts beyond a year are thus better described as ***scenarios*** ;

- **Seasonal effects** can be *removed* by *aggregating to an annual basis* and evaluating the resulting time series ;

- Example of Chocolate/Beer/Electricity in Australia for *Multiple Time Series* ;
- Because autocorrelated series will often have the same trends and seasonal variations, they're often removed before comparison ;
- These trends will often be *deterministic*

  Australian population increase leads to an increase in electricity usage ;
- In contrast, finance will often have *stochastic* trends ;
- Sometimes these can be modeled with a **random walk** ;
- Idea of there being two distinct trends

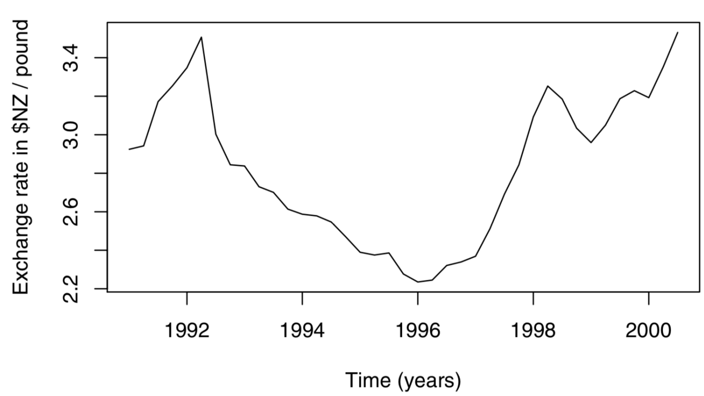
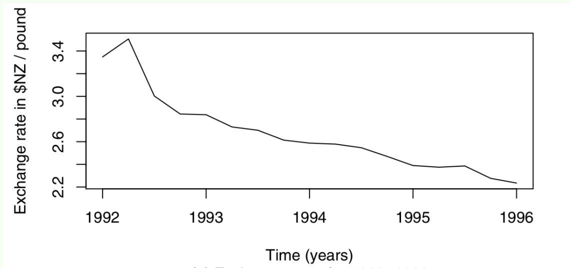


Fig. 1.9. Quarterly exchange rates for the period 1991–2000.

- Thereby, split the series in 2 (`window` in `R`) ;



(a) Exchange rates for 1992–1996

(b) Exchange rates for 1996–1998

- This also highlights the importance of not extrapolating—without additional information we don't know whether the trend will continue ;
- Idea that 2 unrelated time series will be correlated if they both contain a trend thus we can't attribute global warming to fossil fuel increase without a physical explanation ;
- As per scientists, we judge appropriate to attribute a **causal relationship** and to expect mean global temperature to continue rising if greenhouse gas emissions aren't reduced ;

1.5 : Time Series Data—Decomposition of series ;

- Time series of length n : $\{x_t : t = 1, \ldots, n\} = \{x_1, x_2, \ldots, x_n\}$;

  It is a sequence of random variables x_t sampled at n discrete times $1, 2, \ldots, n$;
- Forecast at time t for a future value $t + k$: $\hat{x}_{t+k|t}$;

  k is the **lead time** ;
- If the seasonal effect tends to increase as the trend increases, a multiplicative model may be more appropriate ;

- If there's a multiplicative factor modelled with the random variable (for example we have really large numbers) then an additive decomposition model for $\log(x_t)$ may be more appropriate ;
- Moving average centered around x_t is one of the simplest ways to estimate a trend ;

  Because the average time is $t = 6.5$ and we have integer values, we have one half of both x_{t-6} and x_{t+6} ;
- We can estimate the seasonal effect of each month $\bar{s}_t$ by averaging $\hat{s}_t$ per month ;
- We can take the mean of the averages and substract it from the time series and obtain the **seasonally adjusted series** ;
- **loess** : locally weighted regression technique ;

  « *local* » as it uses a « small » number of points « around » it ;

  Thereby, this reduces the impact of outliers ;

58. Time Series : Correlation

Introduction to Time Series with R 2 - 3.1

2 : Correlation

3.1 : Forecasting Strategies—Purpose

1. Second order properties of a time series ;
2. Relationships of different time series ;

59. Time Series : White Noise and Random Walks

Introduction to Time Series with R 4.1 - 4.4

4.1 : Basic Stochastic Models—Purpose

4.2 : Basic Stochastic Models—White Noise

4.3 : Basic Stochastic Models—Random Walks

4.4 : Basic Stochastic Models—Fitted models and diagnostic plots

1. White noise ;
2. Random walks ;

60. Time Series : Autoregressive Models

Introduction to Time Series with `R` 4.5 - 4.8

4.5 : Basic Stochastic Models—AR models

4.6 : Basic Stochastic Models—Fitted models

4.7 : Basic Stochastic Models—Summary of `R` commands

4.8 : Basic Stochastic Models—Exercices

Correlograms and partial correlograms ;

Stationnarity ;

Forecasting with $AP(p)$ series ;

61. Time Series : Regression

Introduction to Time Series with `R` 5

5 : Regression

1. Correcting for autocorrelation ;
2. Seasonality ;
3. Logarithmic transformations ;
4. Error correction factors ;

62. Time Series : Moving Average Models

Introduction to Time Series with `R` 6.1 - 6.4

6.1 : Stationary Models—Purpose

6.2 : Stationary Models—Strictly stationary series

6.3 : Stationary Models—MA models

6.4 : Stationary Models—Fitted MA models

>

63. Time Series : ARMA Models

Introduction to Time Series with R 6.5 - 6.8

6.5 : Stationary Models—Mixed models : The ARMA process

6.6 : Stationary Models—ARMA models : Empirical Analysis

6.7 : Stationary Models—Summary of R commands

6.8 : Stationary Models—Exercices

>

64. Time Series : ARIMA and SARIMA Models

Introduction to Time Series with R 7.1 - 7.3

7.1 : Non-stationary Models—Purpose

7.2 : Non-stationary Models—Non-Seasonal ARIMA models

7.3 : Non-stationary Models—Seasonal ARIMA models

>

Notes sur les vidéos YouTube

StatQuest: A Gentle Introduction to Machine Learning

>