# 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 ;
  - Abridiged inventories of discrete and continuous probability distributions ;
  - Chi-square distribution ;
  - $t$-distribution ;

- $F$-distribution ;

› **There is a guessing adjustement** ;

# 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-homogenous case ;
   - Resulting from special types of events in the Poisson process ;
   - Resulting from sums of independant Poisson processes ;

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

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

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;

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);

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;

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);

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 ;

**Related lessons ASM**

1. Basics of Statistical Learning

**Vidéos YouTube**

-

## Résumés des chapitres

### 1. Basics of Statistical Learning

-

## 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 ;
   - Efficiency ;
   - Unbiasedness ;
   - Minimum variance ;
   - Sufficiency ;
   - MSE ;

   Calculate parameter estimates using methods other than maximum likelihood ;
2. 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 ;

3. 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 ;
4. Calculate Order Statistics of a sample for a given distribution ;

**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 ;

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;

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 ;

---

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

**Related lessons ASM**

1. Basics of Statistical Learning

**Vidéos YouTube**

›

## Résumés des chapitres

**1. Basics of Statistical Learning**

›

## 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 ;

2. Evaluate models developed using Extended Linear Model approach ;

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 ;

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 ;

5.

**Related lessons ASM**

1. Basics of Statistical Learning

**Vidéos YouTube**

-

## Résumés des chapitres

### 1. Basics of Statistical Learning

›

## 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 ;
2. Model relationships of current and past values of a statistic / metric ;

- Estimation, data analysis, and forecasting ;
- Forecast errors and confidence intervals ;

3. Understand forecasts produced by ARIMA ;

4. Time Series with Regression ;

**Related lessons ASM**

1. Basics of Statistical Learning

**Vidéos YouTube**

-

## Résumés des chapitres

### 1. Basics of Statistical Learning

-

## Notes sur les vidéos YouTube

### StatQuest: A Gentle Introduction to Machine Learning

-