EXCERPTS FROM ACTEX STUDY MANUAL FOR SOA EXAM C/CAS EXAM 4   2009

Table of Contents

Introductory Comments

Model Estimation Section 4 - Kernel Smoothing

TABLE OF CONTENTS - VOLUME 1

MODELING

MODEL ESTIMATION

MODEL ESTIMATION

TABLE OF CONTENTS - VOLUME 2

CREDIBILITY

SIMULATION

PRACTICE EXAMS AND SOLUTIONS

INTRODUCTORY COMMENTS

This study guide is designed to help in the preparation for the Society of Actuaries Exam C and Casualty Actuarial Society Exam 4. The exam covers the topics of modeling, model estimation, construction and selection, credibility, simulation and risk measures.

The study manual is divided into two volumes. The first volume consists of a summary of notes, illustrative examples and problem sets with detailed solutions on the modeling and model estimation topics. The second volume consists of notes examples and problem sets on the credibility, simulation and risk measures topics, as well as 14 practice exams.

The practice exams all have 35 questions. The level of difficulty of the practice exams has been designed to be similar to that of the past 4-hour exams. Some of the questions in the problem sets are taken from the relevant topics on SOA/CAS exams that have been released prior to 2009 but the practice exam questions are not from old SOA exams.

I have attempted to be thorough in the coverage of the topics upon which the exam is based. I have been, perhaps, more thorough than necessary on a couple of topics, such as maximum likelihood estimation, Bayesian credibility and applying simulation to hypothesis testing.

Because of the time constraint on the exam, a crucial aspect of exam taking is the ability to work quickly. I believe that working through many problems and examples is a good way to build up the speed at which you work. It can also be worthwhile to work through problems that have been done before, as this helps to reinforce familiarity, understanding and confidence. Working many problems will also help in being able to more quickly identify topic and question types. I have attempted, wherever possible, to emphasize shortcuts and efficient and systematic ways of setting up solutions. There are also occasional comments on interpretation of the language used in some exam questions. While the focus of the study guide is on exam preparation, from time to time there will be comments on underlying theory in places that I feel those comments may provide useful insight into a topic.

The notes and examples are divided into sections anywhere from 4 to 14 pages, with suggested time frames for covering the material. There are over 330 examples in the notes and over 800 exercises in the problem sets, all with detailed solutions. The 14 practice exams have 35 questions each, also with detailed solutions. Some of the examples and exercises are taken from previous SOA/CAS exams. Questions in the problem sets that have come from previous SOA/CAS exams are identified as such. Some of the problem set exercises are more in depth than actual exam questions, but the practice exam questions have been created in an attempt to replicate the level of depth and difficulty of actual exam questions. In total there are aver 1600 examples/problems/sample exam questions with detailed solutions. ACTEX gratefully acknowledges the SOA and CAS for allowing the use of their exam problems in this study guide.

I suggest that you work through the study guide by studying a section of notes and then attempting the exercises in the problem set that follows that section. My suggested order for covering topics is
(1) modeling (includes risk measures), (2) model estimation , (Volume 1) ,
(3) credibility theory , and (4) simulation , (Volume 2).

It has been my intention to make this study guide self-contained and comprehensive for all Exam C topics, but there are occasional references to the Loss Models reference book (3rd edition) listed in the SOA/CAS catalog. While the ability to derive formulas used on the exam is usually not the focus of an exam question, it is useful in enhancing the understanding of the material and may be helpful in memorizing formulas. There may be an occasional reference in the review notes to a derivation, but you are encouraged to review the official reference material for more detail on formula derivations. In order for the review notes in this study guide to be most effective, you should have some background at the junior or senior college level in probability and statistics. It will be assumed that you are reasonably familiar with differential and integral calculus. The prerequisite concepts to modeling and model estimation are reviewed in this study guide. The study guide begins with a detailed review of probability distribution concepts such as distribution function, hazard rate, expectation and variance.

Of the various calculators that are allowed for use on the exam, I am most familiar with the BA II PLUS. It has several easily accessible memories. The TI-30X IIS has the advantage of a multi-line display. Both have the functionality needed for the exam.

There is a set of tables that has been provided with the exam in past sittings. These tables consist of some detailed description of a number of probability distributions along with tables for the standard normal and chi-squared distributions. The tables can be downloaded from the SOA website www.soa.org .

If you have any questions, comments, criticisms or compliments regarding this study guide, please contact the publisher ACTEX, or you may contact me directly at the address below. I apologize in advance for any errors, typographical or otherwise, that you might find, and it would be greatly appreciated if you would bring them to my attention. ACTEX will be maintaining a website for errata that can be accessed from www.actexmadriver.com .

It is my sincere hope that you find this study guide helpful and useful in your preparation for the exam. I wish you the best of luck on the exam.

Samuel A. Broverman April, 2009
Department of Statistics www.sambroverman.com
University of Toronto E-mail: sam@utstat.toronto.edu or 2brove@rogers.com

MODEL ESTIMATION - SECTION 4 - KERNEL SMOOTHING

The material in this section relates to Section 14.3 of "Loss Models". The suggested time frame for this section is 3 hours.

ME-4.1 Definition of Kernel Density Estimator

We continue to assume that data is in the form of complete individual data. This means that we have a random sample of observations (of loss amounts, or of times of death) and we know the value of each observation (and there may be some repeated values) with no censoring or truncation of data.

Our objective with kernel smoothing is to create a density function that will in some way approximate the (discrete) empirical distribution. We are trying to create a continuous random variable (whose density function will be called the kernel smoothed density estimator that we are finding) that is an approximation to the discrete empirical distribution. The method simultaneously constructs an estimate of the density function called the **kernel density estimator of the density function** and an estimate of the distribution function called the **kernel density estimator of the distribution function**.

There are a variety of **kernels** that can be used to construct the estimator. Each kernel results in its own kernel density estimator. The kernel is itself a density function that is used in the smoothing procedure. The "Loss Models" book mentions three possible kernels (uniform, triangle and gamma), but the density function of any random variable can be used as a kernel. Essentially what is being done when kernel smoothing is applied to estimate a density function is that at each point y_i in the empirical distribution, a density function corresponding to that point is created, and this density function is denoted $k_{y_i}(x)$. For each y_i, $k_{y_i}(x)$ is an actual pdf, and satisfies the requirements of a pdf. The kernel smoothed density estimator is then a finite mixture (or weighted average) of these separate density functions. The mixing "weight" applied to $k_{y_i}(x)$ is the empirical probability $p(y_i)$, and the kernel smoothed estimate of the density function is

$$\widehat{f}(x) = \sum_{\text{All } y_j} p(y_j) \cdot k_{y_j}(x) \ . \qquad (4.1)$$

Once we have identified the empirical distribution points (the sample value y_i's) and their empirical probabilities ($p(y_i)$ for each y_i), we choose which kernel pdf $k_{y_i}(x)$ we will use. Each kernel density function $k_{y_i}(x)$ has a corresponding distribution function $K_{y_i}(x)$. The kernel smoothed estimate of the distribution function is $\widehat{F}(x) = \sum_{\text{All } y_j} p(y_j) \cdot K_{y_j}(x)$, (4.2)

the same sort of "weighted average" mixture formulation that we have for the density estimator $\widehat{f}(x)$.

The simple example we will first consider has the following 4-point random sample: $y_1 = 1$, $y_2 = 2$, $y_3 = 4$, $y_4 = 7$. The empirical distribution assigns a probability of .25 to each of these points, so that $p(1) = p(2) = p(4) = p(7) = \frac{1}{4}$. We will apply uniform kernel, triangle kernel and gamma kernel to this data set to show the construction and properties of the kernel density and distribution function estimator.

According to the definition of $\widehat{f}(x)$, the way in which we calculate $\widehat{f}(x)$, is to apply the formula $\widehat{f}(x) = \sum_{\text{All } y_j} p(y_j) \cdot k_{y_j}(x)$. For our example, we have $y_1 = 1$, $y_2 = 2$, $y_3 = 4$, $y_4 = 7$,
and $p(1) = p(2) = p(4) = p(7) = .25$. Then $\widehat{f}(x) = (.25)[k_1(x) + k_2(x) + k_4(x) + k_7(x)]$,
where the $k_y(x)$ functions are the kernel density functions, and
$\widehat{F}(x) = (.25)[K_1(x) + K_2(x) + K_4(x) + K_7(x)]$.
Note that the subscript of k and K is the y_i-value. For instance, $k_4(x)$ is the kernel function associated with the 3rd y-value, $y_3 = 4$ ($k_4(x)$ is not the 4-th kernel pdf).

ME-4.2 Uniform Kernel Estimator

Uniform kernel density estimator $\widehat{f}(x)$ with bandwidth b
One of the kernels introduced in the "Loss Models" book is the **uniform kernel with bandwidth b.** The uniform kernel is based on the continuous uniform distribution. Recall that the uniform distribution on the interval $[c, d]$ has pdf $k(x) = \begin{cases} \frac{1}{d-c} & \text{for } c \le x \le d \\ 0 & \text{otherwise} \end{cases}$. (4.3)

For the uniform kernel with bandwidth b, at each sample point y_i, the kernel density $k_{y_i}(x)$ is the density function for the uniform distribution on the interval $[y_i - b \, , \, y_i + b]$, so that

$$k_{y_i}(x) = \begin{cases} \frac{1}{2b} & \text{for } y_i - b \le x \le y_i + b \\ 0 & \text{otherwise} \end{cases} . \qquad (4.4)$$

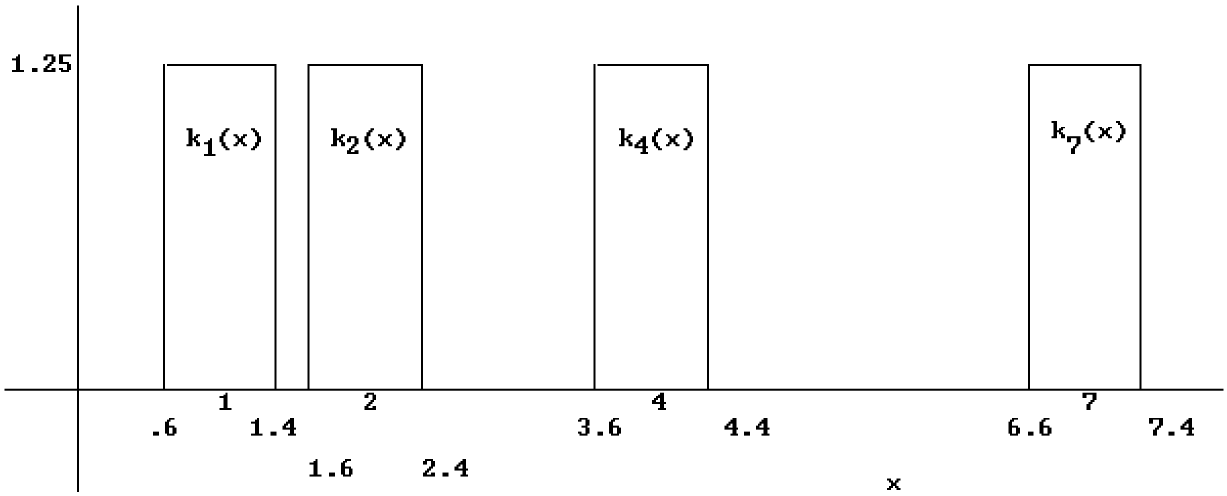
The graph of $k_{y_i}(x)$ is a horizontal line of height $\frac{1}{2b}$ on the interval $[y_i - b \, , \, y_i + b]$, and it is 0 outside that interval; the graph is a rectangle with an area of 1 (since $k_{y_i}(x)$ is a pdf, total area must be 1). Each y_i has its own associated kernel function $k_{y_i}(x)$, which is a uniform density on an interval centered at y_i and the interval length is $2b$.

We will illustrate this method by applying the uniform kernel with a bandwidth of .4 (a somewhat arbitrary choice). At each of the original data points we create a rectangle, with the sample data point value at the center of the base of the rectangle, and with the area of the rectangle being 1. For sample data point y_i in the original random sample, we create a rectangle whose base is from $y_i - b$ to $y_i + b$, and whose height is chosen so that the area of the rectangle is 1. Since the base is $2b$, the height must be $\frac{1}{2b}$. With our chosen value of $b = .4$, the rectangles will all have height $\frac{1}{2(.4)} = 1.25$, and there will be four rectangles with the following bases,

$$[.6 \, , \, 1.4] \; , \; [1.6 \, , \, 2.4] \; , \; [3.6 \, , \, 4.4] \; , \; [6.6 \, , \, 7.4] \, .$$

The notation $k_1(x)$, $k_2(x)$, $k_4(x)$ and $k_7(x)$ describes the four "horizontal-line" functions represented in the graph below (note that the subscript of k is the y-value for the k-th interval). For instance, $k_1(x) = 1.25$ for $.6 \le x \le 1.4$, and $k_1(x) = 0$ for any x outside the interval $[.6 \, , \, 1.4]$. Similar conditions apply to the other three rectangles. The subscript to k is the value of the data point from the original random sample. This identifies which rectangle is being considered. Note that for each sample data point y, $k_y(x)$ is the pdf of the uniform distribution on the interval from $y - b$ to $y + b$.

What we have created is four separate uniform distributions, one for each interval. The graph of these four rectangles is as follows.



The way in which we apply the formula $\widehat{f}(x) = \sum_{\text{All } y_j} p(y_j) \cdot k_{y_j}(x)$ is as follows.

Given a value of x, in order to find $\widehat{f}(x)$, we first determine which rectangle bases contain x. We only need to know the rectangles for which x is in the base because $k_y(x) = 0$ for values of x that are outside of the base rectangle around y. We then find the $k(x)$ value for each rectangle and multiply by the empirical probability for that rectangle's base center point.

For instance, suppose we wish to find the kernel density estimator at $x = 1.1$, i.e., we wish to find $\widehat{f}(1.1)$. $\widehat{f}(1.1)$ is found by first identifying which rectangle bases contain the value 1.1. We see that 1.1 is in the interval $[.6\,,\,1.4]$, so only the kernel function $k_1(x)$ will be non-zero in calculating $\widehat{f}(1.1)$ ($k_2(1.1) = 0$ since 1.1 is not in the interval $[1.6\,,\,2.4]$ centered at $y_2 = 2$, and the same applies to $y_3 = 4$ and $y_4 = 7$).

We find $\widehat{f}(1.1)$ from $\widehat{f}(1.1) = \sum_{\text{All } y_j} p(y_j) \cdot k_{y_j}(1.1)$, which is

$p(1) \cdot k_1(1.1) + p(2) \cdot k_2(1.1) + p(4) \cdot k_4(1.1) + p(7) \cdot k_7(1.1)$
$= (.25)(1.25) + (.25)(0) + (.25)(0) + (.25)(0) = .3125$.

Again, it is important to note that we only used $k_1(1.1)$ since the value $x = 1.1$ was only in the first of the four intervals for the rectangle bases, so $k_1(1.1) = 1.25$, but $k_2(1.1) = 0$ and $k_4(1.1) = 0$ and $k_7(1.1) = 0$.

Suppose we now consider the x-value $x = 3.5$ and we wish to find $\widehat{f}(3.5)$, the kernel density estimator at $x = 3.5$. We see that $x = 3.5$ is not in any of the four intervals formed by the bases of the four rectangles. Therefore, $\widehat{f}(3.5) = 0$, since $k_{y_i}(3.5) = 0$ for each sample data point y_i.

Note that in the simple example we are now considering, since any x is in either one rectangle base or none, $\widehat{f}(x)$ will be .3125 if x is in one of the four rectangle bases, and $\widehat{f}(x) = 0$ if x is not in any of the four rectangle bases. If we were to draw the graph of this $\widehat{f}(x)$ it would look the same as the four rectangles in the graph above, but the heights would be .3125 instead of 1 for each rectangle.