Applied Stochastic Modelling

More Difficult Exercises

B.J.T. Morgan

Applied Stochastic Modelling
Exercises 1

1. In relation to question 2.2, consider the use of the "preference.m" prog gram. This is done in Practical 3, but for a different data set. Note that the re-parameterisation for t is not quite right in general. Can you suggest a better one, and perhaps try it out? You may need to experiment with changing the starting value(s). You may like to experiment also with preference.r.

2. In relation to question 2.3, do the graph for a different data set, such as the menarche data set on p 112. Note that the data are all to be found in the data directory. It may be best to write a short program to do this?

3. In relation to question 2.15: instead you could do (ii), ignoring the bit about a fertilization membrane. Note that the explicit solution only exists for experiments that run for a single time.

Applied Stochastic Modelling
Exercises 2

1.   As an alternative to the logit model, you could try the 3-parameter Aranda-Ordaz model, described in p 105 and on p 244. This could be done twice, once using x, and once using $log(x)$. You could do this on the data of Practical 3C, or the time-to-menarche data. The Aranda-Ordaz model has 3 parameters, and a suitable start might be when $\lambda = 1$ and α, β are set to the logit mles. It can be quite hard to get this model to fit/converge, depending on which data set is used.

2.   You might like to try annealing out on a 1D Cauchy example with several optima, and consider how the points visited relate to the positions of the optima. This is challenging, as the annealing code will need editing to cope with a single variable, and the same is true of the Cauchy function. You will also need to consider what to output and when.

3.   This question is from an examination paper in 2002.
Consider the $\Gamma(2, \lambda)$ pdf for survival:

$$f(t) = \lambda^2 e^{-\lambda t} \quad \text{for} \quad t > 0.$$

Obtain the survivor function.
Write down an expression for the log-likelihood when m failure times are known exactly and $(n - m)$ are all right-truncated, with truncation times, $\{t_i, m + 1 \leq i \leq (n - m)\}, \quad \text{for} \quad n > m.$

Applied Stochastic Modelling
Exercises 3

1   Edit the program of Figure 3.7 (on p 56) so that the Newton-Raphson method is replaced by the method-of-scoring. Try out the method-of-scoring and report the results.

2   Provide a profile log-likelihood for the two-parameter Cauchy case. Give the path followed in the parameter space corresponding to the profile (As in Figure 4.7).

3.   As for Exercise 3, but for $\frac{\partial^2 \ell}{\partial \alpha^2}$.

Applied Stochastic Modelling
Exercises 4

1.   Relating to Question 1 of Practical 5, compare the direct estimate of $Var(\hat{\mu})$, resulting from the parameterisation (μ, β) with that obtained from using the (α, β) parameterisation and then using the δ-method.

2.   Try simulating data **r** for the same **n** and **x** as in Question 1 of Practical 5, using "binsim". Repeat Question 2 above. Consider trying this several times and then plotting the resulting estimates of $Var(\hat{\mu})$, exact vs δ-method.

3.   Experiment with program "monte". For example, try the effect of increasing s. Consider whether "fminbnd" might produce a result, applied to this program (which would first of all have to be re-written as a function, delivering a single value). Might simulated annealing produce some kind of optimum in this case?

Applied Stochastic Modelling
Exercises 5

1. Exercise 6.23: suggest a way to simulate data from a beta-binomial distribution.

2. The following data are sampled from an exponential distribution, with all observations greater than 4 censored at 4, and denoted by a *:

2.11	4*	2.45	4*	2.30	2.39	2.25
2.06	4*	0.03	1.49	0.25	3.47	3.25
4*	2.76	2.00	0.99	3.13	3.67	

Write and run a MATLAB program to estimate the exponential distribution parameter by means of SEM (see p 144).

3. Try to edit and run the bootstrap program of Figure 6.12, so that instead of using the bootstrap, the smoothed bootstrap is used. Note (pp 188 and 189) how one might simulate a smoothed bootstrap sample. For $K()$ one might use a $N(0,1)$ distribution. It will be necessary to choose h.

4. Modify and run the program of Figure 7.2, so that it corresponds to the trivariate distribution of Example 7.1.