PART TWO

Applications

CHAPTER 9

Epileptic seizures

9.1 Introduction

Albert (1991) and Le, Leroux and Puterman (1992) describe the fitting of two-state Poisson–HMMs to series of daily counts of epileptic seizures in one patient. Such models appear to be a promising tool for the analysis of seizure counts, the more so as there are suggestions in the neurology literature that the susceptibility of a patient to seizures may vary in a fashion that can reasonably be represented by a Markov chain; see Hopkins, Davies and Dobson (1985). Another promising approach, not pursued here, is to use an AR(1) analogue based on thinning; see Franke and Seligmann (1993).

9.2 Models fitted

Table 9.1 *Counts of epileptic seizures in one patient on 204 consecutive days (to be read across rows).*

0 3 0 0 0 0 1	1 0 2 1 1 2 0	0 1 2 1 3 1 3
0 4 2 0 1 1 2	1 2 1 1 1 0 1	0 2 2 1 2 1 0
0 0 2 1 2 0 1	0 1 0 1 0 0 0	0 0 0 0 0 1 0
0 0 0 0 1 0 0	0 1 0 0 0 1 0	0 0 1 0 0 1 0
0 2 1 0 1 1 0	0 0 2 2 0 1 1	3 1 1 2 1 0 3
6 1 3 1 2 2 1	0 1 2 1 0 1 2	0 0 2 2 1 0 1
0 0 2 0 1 0 0	0 1 0 0 1 0 0	0 0 0 0 0 1 3
0 0 0 0 0 1 0	1 1 1 0 0 0 0	0 1 0 1 2 1 0
0 0 0 0 0 1 4	0 0 0 0 0 0 0	0 0 0 0 0 0 0
0 0 0 0 0 0 0	0 0 0 0 0 0 0	0

We analyse here a series of counts of myoclonic seizures suffered by one patient on 204 consecutive days*. The observations are given in Table 9.1 and displayed in Figure 9.1.

* The 225-day series published by Le *et al.* (1992) contained a repeat of the observations for a 21-day period; see MacDonald and Zucchini (1997, p. 208). Table 9.1 gives the corrected series.

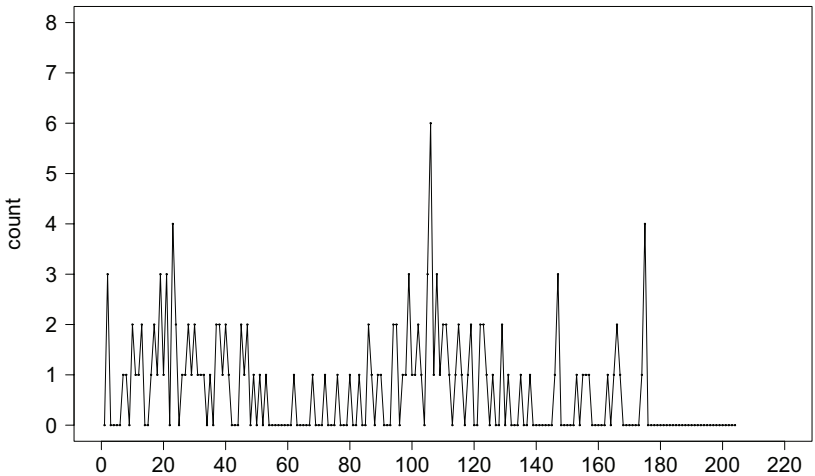

Figure 9.1 *Epileptic seizure counts on 204 days.*

Le *et al.* use an HMM of the type described by Leroux and Puterman (1992). Their model does not assume that the underlying Markov chain is stationary. It is fitted by maximizing the likelihood conditional on the Markov chain starting in a given state with probability one, and then maximizing over the possible initial states.

We consider a similar HMM, but based on a stationary Markov chain and fitted by maximization of the unconditional likelihood of the observations. We investigate models with $m = 1$, 2, 3 and 4 states. (The one-state model is just the model which assumes that the observations are realizations of independent Poisson random variables with a common mean. That mean is the only parameter.)

Table 9.2 gives the AIC and BIC values for the models. From the table we see that, of the four models considered, the three-state model is chosen by AIC, but the two-state model is chosen, by a large margin, by BIC. We concentrate on the two-state model, the details of which are as follows. The Markov chain has transition probability matrix

$$\begin{pmatrix} 0.965 & 0.035 \\ 0.027 & 0.973 \end{pmatrix},$$

Table 9.2 *Epileptic seizure counts: comparison of several stationary Poisson–HMMs by means of AIC and BIC.*

no. of states	k	$-l$	AIC	BIC
1	1	232.15	466.31	469.63
2	4	211.68	431.36	**444.64**
3	9	205.55	**429.10**	458.97
4	16	201.68	435.36	488.45

Table 9.3 *Sample ACF for the epileptic seizure counts.*

k	1	2	3	4	5	6	7	8
$\hat{\rho}(k)$	0.236	0.201	0.199	0.250	0.157	0.181	0.230	0.242

and starts from the stationary distribution (0.433, 0.567). The seizure rates in states 1 and 2 are 1.167 and 0.262 respectively.

The ACF of the model can be computed by the results of Exercise 4 of Chapter 2. It is given, for all positive integers k, by

$$\begin{aligned} \rho(k) &= \left(1 + \frac{\boldsymbol{\delta}\boldsymbol{\lambda}'}{(\lambda_2 - \lambda_1)^2 \delta_1 \delta_2}\right)^{-1} (1 - \gamma_{12} - \gamma_{21})^k \\ &= 0.235 \times 0.939^k. \end{aligned}$$

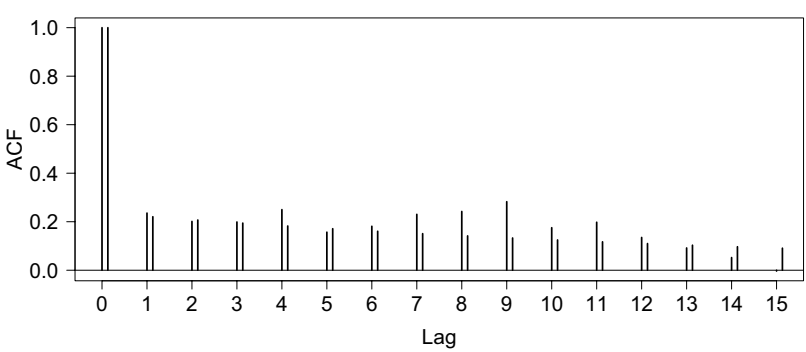

Figure 9.2 *Sample and theoretical ACF for the epileptic seizures data. At each lag the left bar represents the sample ACF, and the right bar the ACF of a stationary two-state Poisson–HMM.*

Table 9.4 *Observed and expected numbers of days with* $r = 0, 1, 2, \ldots$ *epileptic seizures.*

r	observed no.	expected no.
0	117	116.5
1	54	55.4
2	23	21.8
3	7	7.5
4	2	2.1
5	0	0.5
≥ 6	1	0.1
	204	203.9

Table 9.3 gives the corresponding sample ACF. Figure 9.2, which displays both, shows that the agreement between sample and theoretical ACF is reasonably close.

The marginal properties of the model can be assessed from Table 9.4, which gives the observed and expected numbers of days on which there were 0, 1, 2, ..., 6 or more seizures. Agreement is excellent.

9.3 Model checking by pseudo-residuals

We now use the techniques of Section 6.2.2 to check for outliers under the two-state model we have chosen. Figure 9.3 is a plot of ordinary normal pseudo-residual segments. From it we see that three observations of the 204 stand out as extreme, namely those for days 106, 147 and 175. They all yield pseudo-residual segments lying entirely within the top $\frac{1}{2}\%$ of their respective distributions.

It is interesting to note that observations 23 and 175, both of which represent four seizures in one day, yield rather different pseudo-residual segments. The reason for this is clear when one notes that most of the near neighbours (in time) of observation 175 are zero, which is not true of observation 23 and its neighbours. Observation 23 is much less extreme relative to its neighbours than is 175, and this is reflected in the pseudo-residual. Similarly, observation 106 (six seizures in a day) is less extreme relative to its neighbours than is observation 175 (four seizures).

However, a more interesting exercise is to see whether, if a model had been fitted to (say) the first 100 observations only, day-by-day monitoring thereafter by means of forecast pseudo-residuals would have identified any outliers. The two-state model fitted from the first 100 observa-

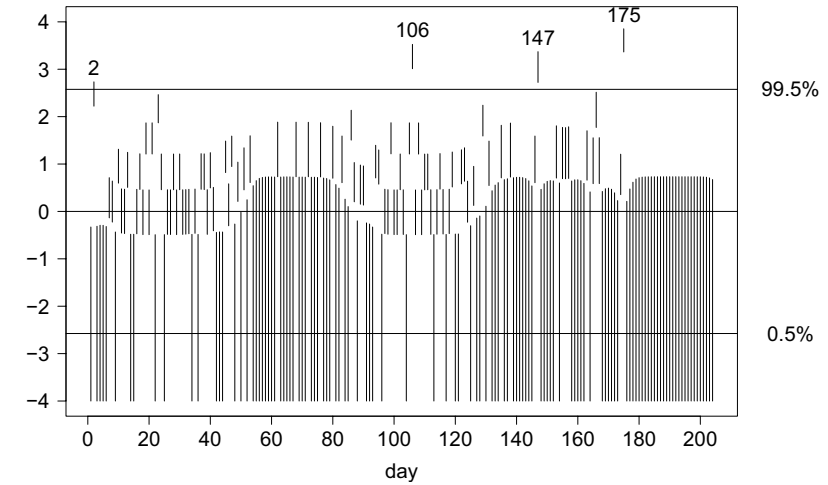

Figure 9.3 *Epileptic seizures data: ordinary (normal) pseudo-residual segments, relative to stationary two-state Poisson–HMM fitted to all 204 observations.*

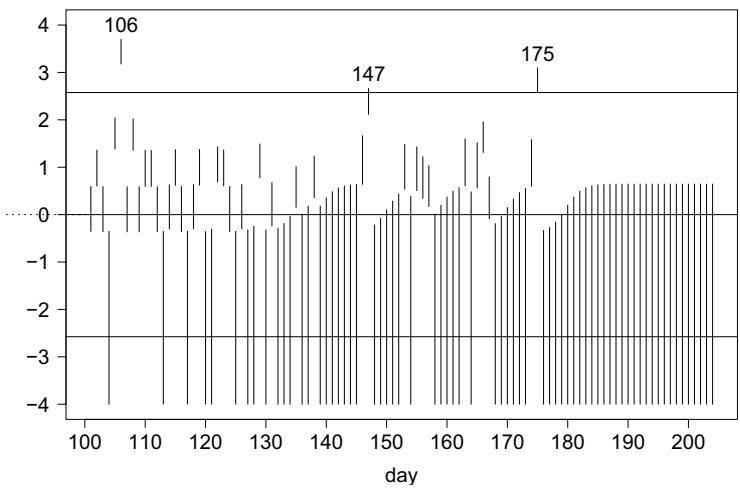

Figure 9.4 *Epileptic seizures data: forecast pseudo-residual segments, relative to stationary two-state Poisson–HMM fitted to data for days 1–100 only.*

tions has transition probability matrix

$$\begin{pmatrix} 0.983 & 0.017 \\ 0.042 & 0.958 \end{pmatrix},$$

and seizure rates 1.049 and 0.258.

From a plot of forecast pseudo-residuals, Figure 9.4, we see that the same three observations stand out: days 106, 147 and 175. Observation 106 emerges from such a monitoring procedure as the clearest outlier relative to its predecessors, then 175, then 147.

Exercises

1. Consider the two-state model for the epileptic seizures.
   (a) Compute the probabilities $\Pr(C_t = i \mid \mathbf{X}^{(T)})$ for this model, for $i = 1, 2$ and all t.
   (b) Perform both local and global decoding to estimate the most likely states. Do the results differ?
   (c) Perform state prediction for the next three time points; i.e. find the probabilities $\Pr(C_{T+h} = i \mid \mathbf{X}^{(T)})$ for $h = 1, 2, 3$.
   (d) Compute the forecast distribution $\Pr(X_{T+h} = x \mid \mathbf{X}^{(T)})$ for $h = 1, \ldots, 10$.

2.(a) Fit a stationary three-state Poisson–HMM to the epileptic seizures.
   (b) Find the general expression for the ACF of this model.
   (c) For lags 1 to 8, compare this model ACF with the sample ACF given in Table 9.3.