

# Statistical Modeling of Daily Rainfall Occurrences

JAMES A. SMITH

*Interstate Commission on the Potomac River Basin, Rockville, Maryland*

In this paper, likelihood-based inference procedures for discrete point process models are developed, and a new family of discrete point process models for daily rainfall occurrences is proposed. The model, which is termed a Markov Bernoulli process, can be viewed as a sequence of Bernoulli trials with randomized success probabilities. Contained within the family of Markov Bernoulli models are Markov chain and Bernoulli trial models. Asymptotic properties of maximum likelihood estimators of Markov Bernoulli model parameters are derived. These results provide the basis for assessing standard errors and correlation of parameter estimators and for developing likelihood ratio tests to choose among Markov Bernoulli, Markov chain, and Bernoulli trial models. Inference procedures are applied to a data set from Washington, D. C.

## 1. INTRODUCTION

Markov chain models have played a major role in modeling wet-day sequences. Two of the most attractive features of Markov chain models are the ease with which seasonality is accommodated and availability of effective statistical inference procedures for parameter estimation and model selection [see *Stern and Coe*, 1984]. In this paper we propose a new family of models for wet-day sequences, which we term Markov Bernoulli processes. A Markov Bernoulli process can be viewed as a sequence of Bernoulli trials with randomized success probabilities. Although Markov Bernoulli processes contain Markov chain models, we are motivated less by ideas adopted from the Markov chain literature than by ideas from the continuous point process literature. In this respect we follow the approach of *Foufoula-Georgiou* [1985] in developing "discrete point process" models of rainfall occurrences (see also *Foufoula-Georgiou and Lettenmaier* [1986] and *Guttorp* [1986]).

An attractive feature of (some) continuous point process models is that physical significance can be readily attributed to model components. Models that have been introduced as physically based rainfall occurrence models generalize Poisson processes in one of two ways. "Poisson cluster processes" [see *Kavvas and Delleur*, 1981] are constructed by distributing rainfall events about a Poisson process of "cluster centers." *Kavvas and Delleur* have shown that components of a Poisson cluster process have a natural interpretation based on frontogenesis: Cluster centers correspond to fronts, each of which has a random number of "storm" events associated with it. "Cox processes" [see *Smith and Karr*, 1983] can be viewed as Poisson processes with a randomly varying rate of occurrence. For modeling rainfall occurrences the random rate of occurrence has been interpreted as a "stochastic climatological process." In the Cox process model developed by *Smith and Karr* the random rate of occurrence is related to the frequency and duration of anticyclonic events.

The discrete-time analog of Poisson processes is Bernoulli trials. Generalizations of Bernoulli trials, including Markov chain models are typically based on a time series approach

rather than the point process approach. One notes that much of the literature on wet-day sequences (see *Waymire and Gupta* [1981] for a review) is concerned with determining the appropriate order of a Markov chain model, much as the literature on time series modeling of streamflow is concerned with determining the appropriate number of autoregressive terms in an autoregressive moving average model. Our approach to model construction is to generalize Bernoulli trials via the point process approach. A distinguishing feature of the point process approach is reliance on the "stochastic intensity" for description and statistical analysis of models. For a model of daily rainfall occurrences the value of the stochastic intensity on a given day is the conditional probability of a wet day given the preceding history of wet-dry days.

By analogy with Cox processes we generalize Bernoulli trials by randomizing the sequence of success probabilities. For the Markov Bernoulli process the sequence of randomized success probabilities is particularly simple; it is a seasonal Markov chain which, for a given day of year, has two states. More complex models of wet-day sequences can be constructed by selecting more complex sequences of randomized success probabilities. We illustrate in section 3 that a discrete-time analog of Neyman-Scott models can be constructed in this fashion. We also note in section 3 that the "Poisson cluster" construction has no discrete-time analog. An attractive feature of Cox processes is availability of statistical inference procedures [Smith and Karr, 1985]. *Smith* [1984] suggests that Cox processes are amenable to incorporation of seasonality. These features provide ground for hope that Markov Bernoulli processes may possess the attractive features of Markov chain models.

Development of procedures for parameter estimation and model selection is also motivated by the point process approach. The key result is theorem 4.1, in which the log likelihood function of a discrete point process is represented as a simple function of the stochastic intensity of the point process. It follows from theorem 4.1 that likelihood-based inference procedures can be used whenever the stochastic intensity can be computed.

Contents of the sections are as follows. In section 2 we introduce the general framework for a seasonal point process model of daily rainfall occurrences. The stochastic intensity of a point process is introduced and used to define Bernoulli point processes and Markov chain point processes. Markov

Bernoulli point processes are the topic of section 3. The most important distributional result for Markov Bernoulli processes is a recursive formula for computing its stochastic intensity (theorem 3.1). Relationships with other discrete point process models are derived and an important invariance property of Markov Bernoulli processes is presented. In section 4, likelihood-based inference procedures are developed for seasonal point process models. In theorem 4.1 it is shown that maximum likelihood estimators of Markov Bernoulli model parameters are consistent and asymptotically normal. Furthermore, it is shown that log likelihood ratios have a limiting  $\chi^2$  distribution. These results provide the tools for (1) estimating parameters of Markov Bernoulli models, (2) assessing standard errors and correlation of parameter estimators, and (3) carrying out tests to choose among Markov Bernoulli, Markov chain, and Bernoulli trial models. To conclude section 4, inference procedures for Markov Bernoulli models are applied to a data set from Washington, D. C.

2. DEFINITIONS AND NOTATION

Consider a sequence of nonnegative random variables  $Z_1^1, \dots, Z_T^1, Z_1^2, \dots, Z_T^2, \dots$  with the interpretation that  $Z_t^i$  represents total rainfall on day  $t$  of year  $i$ . The total number of days during the year is  $T$  (which we will take to be 366 throughout the paper). The point process of wet-day occurrences is specified by

$$Y_t^i(x) = 1(Z_t^i > x) \quad t = 1, \dots, T \quad i = 1, 2, \dots \quad (1)$$

that is,  $Y_t^i(x)$  equals 1 if  $Z_t^i$  is greater than the threshold  $x$ ; otherwise,  $Y_t^i(x)$  equals zero. The threshold  $x$  is generally determined by the nature of the application and/or sampling thresholds of the station being used. The counting process

$$N_t^i(x) = \sum_{s=1}^t Y_s^i(x) \quad t = 1, \dots, T \quad i = 1, 2, \dots \quad (2)$$

provides cumulative wet days over the course of year  $i$ . The occurrence times of wet days are denoted

$$T_1^i(x) = \min \{t: Y_t^i(x) = 1\} \quad N_T^i(x) \geq 1$$

$$T_k^i(x) = \min \{t > T_{k-1}^i(x): Y_t^i(x) = 1\} \quad k \leq N_T^i(x) \quad (3)$$

The data set of wet day occurrences is denoted by the  $\sigma$  algebra

$$\mathcal{H}_t^i(x) = \sigma\{Y_s^i(x); s \leq t\} \cup \sigma\{Y_s^i(x); s = 1, \dots, T, j < i\}$$

$$t = 1, \dots, T \quad i = 1, 2, \dots \quad (4)$$

The data set  $\mathcal{H}_t^i(x)$  contains data on wet day occurrences up to and including day  $t$  of year  $i$ . The precipitation threshold  $x$  plays only a "supporting" role in model development. To simplify notation, we will suppress dependence on the threshold  $x$  unless explicitly stated otherwise. Thus  $Y_t^i$  will denote the point process of wet days (larger than  $x$ ).

The double indexing by year and day of year is used to emphasize two aspects of rainfall occurrences that are of particular importance in developing statistical inference procedures. From year to year the random vectors of rainfall occurrences (each with  $T$  components) are very nearly independent and identically distributed (IID) copies of each other. Slight year-to-year dependence arises by virtue of the fact that rainfall early in one year will be related to rainfall at the end of the previous year. By contrast, within a year there is not only day-to-day dependence but also pronounced seasonal in-

homogeneity. This contrast plays an important role in determining the form of statistical inference procedures developed in section 4.

The stochastic intensity  $\{\lambda_t^i\}$  of a point process  $\{Y_t^i\}$  is defined by

$$\lambda_t^i = P\{Y_t^i = 1 | \mathcal{H}_{t-1}^i\} \quad t = 1, \dots, T \quad i = 1, 2, \dots \quad (5)$$

The value of  $\lambda_t^i$  is the conditional probability that day  $t$  of year  $i$  is a wet day given the history of preceding wet days. Equation (5) contains the assumption that the probability of rainfall on a given day does not depend on the preceding history of rainfall amounts, if we know the preceding history of wet days.

In section 4, statistical inference procedures based on the stochastic intensity (5) are developed for discrete point processes. *Smith and Karr [1985]* develop intensity-based inference procedures for continuous point process models of rainfall occurrences. It will be seen that motivation and mechanics for using intensity-based inference procedures are much simpler in the discrete case.

Utility of the stochastic intensity for inference problems stems largely from the facts that (1) the stochastic intensity uniquely determines the distribution of a point process and (2) the likelihood function has a simple representation in terms of the stochastic intensity (see theorem 4.1). We have for  $n$  vectors of 0-1 values  $y^i = (y_1^i, \dots, y_T^i) \quad i = 1, \dots, n$

$$P\{Y_T^n = y_T^n, \dots, Y_1^n = y_1^n, \dots, Y_T^1 = y_T^1, \dots, Y_1^1 = y_1^1\}$$

$$= \prod_{i=1}^n \prod_{s=1}^T P\{Y_s^i = y_s^i | Y_{s-1}^i = y_{s-1}^i, \dots, Y_1^i = y_1^i\}$$

$$= \prod_{i=1}^n \prod_{s=1}^T y_s^i \lambda_s^i + (1 - y_s^i)(1 - \lambda_s^i) \quad (6)$$

2.1. Bernoulli Point Process

Let  $p$  map  $E = \{1, \dots, T\}$  into the interval  $[0, 1]$ . The point process  $\{Y_t^i\}$  is termed a Bernoulli point process with success probability  $p$  if

$$\lambda_t^i = p(t) \quad t = 1, \dots, T \quad i = 1, 2, \dots \quad (7)$$

Thus for a Bernoulli point process the value of the stochastic intensity is deterministic, depending only on day of year. On day  $t$  of any year the stochastic intensity equals the success probability  $p(t)$ .

2.2. Markov Chain Point Process

Let  $q_0$  and  $q_1$  map  $E$  into the interval  $(0, 1)$ . The point process  $\{Y_t^i\}$  is termed a Markov chain point process with transition probabilities  $q_0$  and  $q_1$  if

$$\lambda_t^i = q_1(t)Y_{t-1}^i + (1 - q_0(t))(1 - Y_{t-1}^i) \quad (8)$$

$$t = 1, \dots, T \quad i = 1, 2, \dots$$

The value of the stochastic intensity on day  $t$  of year  $i$  is  $q_1(t)$  if the previous day was wet and  $(1 - q_0(t))$  if the previous day was dry. In other words,  $q_1(t)$  is the conditional probability of a wet day given that the preceding day was wet and  $1 - q_0(t)$  is the conditional probability of a wet day given that the preceding day was dry.

Note that (8) is not well defined for  $t = 1$ . Clearly, we want

$$Y_{t-1}^i = Y_T^{i-1} \quad (9)$$

whenever  $t = 1$  (and  $i$  is greater than 1). For  $t = 1$ , (9) states that the value of  $Y$  on the day preceding the first day of year  $i$  is the value of  $Y$  on the last day of year  $i-1$ . To simplify presentation of results, we will not explicitly use notation of the form (9) whenever interpretation is clear from context.

### 3. MARKOV BERNOULLI PROCESSES

In this section we introduce a class of discrete point processes which will be termed Markov Bernoulli processes. A Markov Bernoulli process can be viewed as a Bernoulli point process with randomized success probability. The randomized success probability is of particularly simple form; it is a (seasonal) Markov chain that alternates between two states. The main result of this section is theorem 3.1 in which a recursive formula for computing the stochastic intensity of a Markov Bernoulli process is presented. In subsequent corollaries, relationships between Markov Bernoulli processes and other families of discrete point process models are presented and an important invariance property of Markov Bernoulli processes is derived. At the end of the section we indicate how more complex models can be constructed from randomizing the success probabilities of a Bernoulli point process by a Markov chain. In particular, we construct in this fashion a discrete analog of the Neyman-Scott model. We begin by defining a Markov Bernoulli process.

Let  $\{Y_t^i\}$  be a Markov chain point process with transition probabilities  $q_0$  and  $q_1$  and let  $p$  be a function mapping  $E$  into  $(0, 1]$ . The sequence of randomized success probabilities is given by

$$X_t^i = p(t)Y_t^i \quad t = 1, \dots, T \quad i = 1, 2, \dots \quad (10)$$

Note that on day  $t$  of any year the random success probability can take either the value  $p(t)$  or zero. The point process  $\{Y_t^i\}$  is termed a Markov Bernoulli point process with parameters  $p$ ,  $q_0$ , and  $q_1$  if for all positive integers  $n$  and 0-1 vectors  $y^i = (y_1^i, \dots, y_T^i)$   $i = 1, \dots, n$

$$\begin{aligned} P\{Y^n = y^n, \dots, Y^1 = y^1 | X^i, \quad i = 1, \dots, n\} \\ &= \prod_{i=1}^n \prod_{s=1}^T P\{Y_s^i = y_s^i | X_s^i\} \\ &= \prod_{i=1}^n \prod_{s=1}^T y_s^i X_s^i + (1 - y_s^i)(1 - X_s^i) \end{aligned} \quad (11)$$

It follows from (11) that (1) the sequence of wet-dry days is conditionally independent given the sequence  $\{X_t^i\}$  of success probabilities, (2) the conditional probability of a wet day, given that  $X_t^i$  equals  $p(t)$ , is  $p(t)$ , and (3) the conditional probability of a wet day, given that  $X_t^i$  equals zero, is zero. The main distributional result for Markov Bernoulli processes is the following recursive formula for the stochastic intensity.

#### 3.1. Theorem 3.1

The stochastic intensity of a Markov Bernoulli process with parameters  $p$ ,  $q_0$ , and  $q_1$  is given by

$$\begin{aligned} \lambda_t^i &= p(t)[1 - q_0(t)] + p(t)[q_0(t) + q_1(t) - 1] \\ &\quad \cdot \left[ Y_{t-1}^i - (1 - Y_{t-1}^i) \frac{1 - p(t-1)}{p(t-1)} \lambda_{t-1}^i \right] \end{aligned} \quad (12)$$

*Proof*

$$\begin{aligned} \lambda_t^i &= E[Y_t^i | \mathcal{H}_{t-1}^i] \\ &= E\{E[Y_t^i | X_t^i] | \mathcal{H}_{t-1}^i\} \\ &= E[X_t^i | \mathcal{H}_{t-1}^i] \\ &= p(t)P\{X_t^i = p(t) | \mathcal{H}_{t-1}^i\} \\ &= p(t)[P\{X_t^i = p(t), X_{t-1}^i = p(t-1) | \mathcal{H}_{t-1}^i\} \\ &\quad + P\{X_t^i = p(t), X_{t-1}^i = 0 | \mathcal{H}_{t-1}^i\}] \\ &= p(t)[P\{X_t^i = p(t) | X_{t-1}^i = p(t-1)\}P\{X_{t-1}^i = p(t-1) | \mathcal{H}_{t-1}^i\} \\ &\quad + P\{X_t^i = p(t) | X_{t-1}^i = 0\}P\{X_{t-1}^i = 0 | \mathcal{H}_{t-1}^i\}] \\ &= p(t)[q_1(t)P\{X_{t-1}^i = p(t-1) | \mathcal{H}_{t-1}^i\} \\ &\quad + (1 - q_0(t))(1 - P\{X_{t-1}^i = p(t-1) | \mathcal{H}_{t-1}^i\})] \\ &= p(t)[(1 - q_0(t)) + (q_0(t) + q_1(t) - 1)P\{X_{t-1}^i = p(t-1) | \mathcal{H}_{t-1}^i\}] \\ &= p(t)(1 - q_0(t)) \\ &\quad + p(t)(q_0(t) + q_1(t) - 1)[Y_{t-1}^i + (1 - Y_{t-1}^i) \\ &\quad \cdot (1 - p(t-1))P\{X_{t-1}^i = p(t-1) | \mathcal{H}_{t-2}^i\}] \\ &= p(t)(1 - q_0(t)) + p(t)(q_0(t) + q_1(t) - 1) \\ &\quad \cdot \left[ Y_{t-1}^i + (1 - Y_{t-1}^i) \frac{1 - p(t-1)}{p(t-1)} \lambda_{t-1}^i \right] \end{aligned} \quad (13)$$

The following two corollaries characterize relationships with Bernoulli point processes and Markov chain point processes.

#### 3.2. Corollary 1

A Markov Bernoulli process with parameters  $p$ ,  $q_0$ , and  $q_1$  is a Bernoulli point process if and only if  $q_0 + q_1 - 1 = 0$ , that is,  $q_0(t) + q_1(t) - 1 = 0$ ,  $t = 1, \dots, T$ .

#### 3.3. Corollary 2

A Markov Bernoulli process with parameters  $p$ ,  $q_0$ , and  $q_1$  is a Markov chain point process if and only if  $p = 1$ , that is  $p(t) = 1$ ,  $t = 1, \dots, T$ .

Corollary 1 follows by comparing (12) with (7); the second corollary follows by comparing (12) with (8). These results are of particular importance in developing likelihood ratio tests for model selection in section 4.

We now examine relationships between Markov Bernoulli processes and discrete renewal processes. Importance of renewal processes for rainfall modeling stems in large part from data analysis results. For numerous data sets it has been found that correlation of interarrival times for rainfall occurrences (within a season) is very small (see, for example, *Smith and Karr [1983]*), suggesting the plausibility of renewal models.

The point process  $\{Y_t^i\}$  is a renewal process if the interarrival times are independent and identically distributed. For simplicity we will denote the interarrival times of  $\{Y_t^i\}$  by  $U_1, U_2, \dots$ . The distribution of a discrete renewal process can be specified by its probability mass function

$$f(k) = P\{U_j = k\} \quad k = 1, 2, \dots \quad (14)$$

or hazard function

$$h(k) = P\{U_j = k | U_j > k - 1\} \quad k \geq 2 \quad (15)$$

$$h(k) = f(1) \quad k = 1$$

The two are related by

$$h(k) = \frac{f(k)}{1 - \sum_{j=1}^{k-1} f(j)} \quad k = 1, 2, \dots \quad (16)$$

Importance of the hazard function is clear from the following lemma [Karr, 1986].

### 3.4. Lemma

The stochastic intensity of a renewal process is given by

$$\lambda_t^i = h(V_t^i) \quad (17)$$

where  $h$  is the hazard function and  $\{V_t^i\}$  is the backward recurrence time, that is, the time that has elapsed since the most recent event.

### 3.5. Corollary 3

A Markov Bernoulli process with parameters  $p (< 1)$ ,  $q_0$ , and  $q_1$ , which do not vary with time of year is a renewal process with hazard function

$$h(k) = \frac{p(1-q_0)}{p + (1-p)(q_0 + q_1)} + p \left[ \frac{1 - (1-p)q_1}{(1-p)[1 - (q_0 + q_1 - 1)(1-p)]} \right] \cdot [(1-p)(q_0 + q_1 - 1)]^k \quad (18)$$

### 3.6. Proof

It follows from (12) that

$$\lambda_t^i = h(V_t^i)$$

where  $h$  is the solution to the first-order difference equation

$$h(k) = p(1 - q_0) - (q_0 + q_1 - 1)(1 - p)h(k - 1) \quad (19)$$

with initial condition

$$h(1) = pq_1$$

The solution to (19) is given by (18) (see, for example, Goldberg [1958]).

### 3.7. Remarks

1. If  $q_0 + q_1 - 1$  is greater than zero,  $h$  is a decreasing function of  $k$ . If  $q_0 + q_1 - 1$  is less than zero,  $h$  is an oscillating function of  $k$ . In particular,  $h(1)$  is less than  $h(2)$  if  $q_0 + q_1 - 1$  is less than zero, implying that rainfall is more likely two days following a wet day than the day after a wet day. Recall that if  $q_0 + q_1 - 1$  equals zero, the point process is Bernoulli. We can thus view  $q_0 + q_1 - 1$  as a measure of temporal correlation of the point process.

2. In the case that the parameters  $p$ ,  $q_0$ , and  $q_1$  vary with time of year, it follows from (12) that

$$\lambda_t^i = h_t(V_t^i) \quad (20)$$

where the functions  $h_t$ ,  $t = 1, \dots, T$  are solutions to the system of difference equations

$$h_t(k) = p(t)(1 - q_0(t)) - (q_0(t) + q_1(t) - 1)(1 - p(t))h_{t-1}(k - 1) \quad (21)$$

$$h_t(1) = p(t)q_1(t)$$

Existence of solutions and general expressions for solutions can be obtained from results in Goldberg [1958]. While (20) does not provide a useful computational tool (in the way that (18) does), it does indicate an alternative strategy for obtaining "seasonal renewal models." Instead of dividing the year into homogeneous seasons and fitting separate renewal models for each season, one introduces seasonality into a renewal model (specified by its hazard function as in (17)) through the stochastic intensity using the representation (20).

An attractive feature of Markov Bernoulli processes is invariance under random deletion (or  $p$ -thinning) of events. Corollary 4 below states that if we randomly delete wet days from a Markov Bernoulli model, we still have a Markov Bernoulli model. As discussed below, this property insures that model selection is not too strongly tied to the particular precipitation threshold that is chosen. Before presenting corollaries 4 and 5 we define the " $p$ -thinning" of a point process. Let  $\{Z_t^i\}$  be a Bernoulli point process with success probability  $p$ , which is independent of the point process  $\{Y_t^i\}$ . The point process  $\{Y_t^i\}$  specified by

$$Y_t^i = Z_t^i Y_t^i \quad t = 1, \dots, T \quad i = 1, 2, \dots \quad (22)$$

is termed a  $p$ -thinning of  $\{Y_t^i\}$ . From (22) it is seen that the  $p$ -thinning  $\{Y_t^i\}$  is constructed by randomly deleting events of  $\{Y_t^i\}$ ; an event of  $\{Y_t^i\}$  that occurs on day  $s$  of any year is retained with probability  $p(s)$  and deleted with probability  $(1 - p(s))$ . Corollary 5 below is virtually a paraphrase of the definition of a Markov Bernoulli process. Corollary 4 follows from the commutative property of  $p$ -thinning.

### 3.8. Corollary 4

The  $p$ -thinning  $\{Y_t^i\}$  of a Markov Bernoulli process  $\{Y_t^i\}$  with parameters  $p$ ,  $q_0$ , and  $q_1$  is a Markov Bernoulli process with parameters  $\bar{p}$ ,  $\bar{q}_0$ , and  $\bar{q}_1$ .

### 3.9. Corollary 5

The  $p$ -thinning of a Markov chain point process with parameters  $q_0$  and  $q_1$  is a Markov Bernoulli process with parameters  $p$ ,  $q_0$ , and  $q_1$ .

A common assumption in modeling daily precipitation is that precipitation amount is conditionally independent of previous precipitation values given that positive precipitation occurs (see, for example, Stern and Coe [1984] and Woolhiser and Roldan [1982]), that is

$$P\{Z_t^i \leq x | Z_t^i > 0, Z_{t-1}^i, \dots, Z_1^i\} \\ = P\{Z_t^i \leq x | Z_t^i > 0\} = H_t(x) \quad (23)$$

where, for each  $t$ ,  $H_t(x)$  is a distribution function on  $(0, \infty)$ . In this case, the point process of days with total precipitation greater than  $x$ ,  $\{Y_t^i(x)\}$  is a  $p$ -thinning of the point process of days with positive rainfall  $\{Y_t^i(0)\}$ , with

$$p(t) = 1 - H_t(x) \quad t = 1, \dots, T \quad (24)$$

It follows from the preceding corollary that if the conditional independence property (23) holds and if  $\{Y_t^i(x)\}$  is a Markov chain for some  $x \geq 0$ , then for any  $u$  not equal to  $x$ ,  $Y_t^i(u)$  is not a Markov chain. On the other hand, if  $\{Y_t^i(x)\}$  is a Markov Bernoulli process for some  $x \geq 0$ , then  $\{Y_t^i(x + u)\}$  is a Markov Bernoulli process for all  $u \geq 0$ . This invariance property of Markov Bernoulli processes under  $p$ -thinning is very attractive in light of (23).

We conclude this section with a discussion of extensions to the Markov Bernoulli model. At the beginning of the section we noted that the Markov Bernoulli model can be viewed as a Bernoulli point process with randomized success probabilities. The random sequence of success probabilities  $\{X_i\}$  for a Markov Bernoulli process is a simple Markov chain. By allowing  $\{X_i\}$  to be a more complicated Markov chain we obtain a broad family of discrete point process models, including discrete analogs to Neyman-Scott models. The reason for restricting the sequence of randomized success probabilities to be a Markov chain is that the stochastic intensity of the resulting models can often be computed from "filter equations" analogous to (12).

A constructive definition of a discrete Neyman-Scott model would proceed along the following lines (for notational simplicity we consider a stationary model on the positive integers). Let  $\{\tilde{T}_n\}$  be arrival times of a Bernoulli point process on the nonnegative integers, representing "cluster centers." Let  $\{\tilde{N}_i\}$  be a sequence of nonnegative integer-valued random variables with  $\tilde{N}_i$  representing the number of cluster members associated with the cluster located at time  $\tilde{T}_i$ . We now attempt to distribute cluster members about cluster centers. The problem we encounter is that unless  $\tilde{N}_i$  equals zero or 1 for all  $i$ , we cannot avoid the possibility of having more than one event on a single day. One possibility for sidestepping this problem is to count multiple events as one. Instead, we pursue a totally different approach.

Smith and Karr [1985] show that a large class of Neyman-Scott processes can be represented as Cox processes (that is, Poisson processes for which the rate of occurrence is a random process). The following result suggests a different approach for constructing a discrete Neyman-Scott model.

### 3.10. Lemma [Smith and Karr, 1985]

Let  $N$  be a Neyman-Scott process on  $[0, \infty)$  of the following form: (1) the Poisson process of cluster centers is  $\tilde{N}$ ; arrival times are denoted  $\tilde{T}_i$ , (2) the distribution of cluster sizes is Poisson with parameter  $a$ , and (3) the distances from each cluster center to its cluster members are IID exponentially distributed with parameter  $b$ . Then  $N$  is a Cox process directed by the Markov process

$$X(t) = \exp(-bt)X(0) + ab \sum_{i=1}^{\tilde{N}_i} \exp(-b(t - \tilde{T}_i)) \quad (25)$$

The Neyman-Scott process  $N$  can be interpreted as a Poisson process with randomized rate of occurrence given by the Markov process (25).

By analogy with (25) we introduce the following definition of a discrete Neyman-Scott model. Let  $\{\tilde{Y}_i\}$  be a stationary Bernoulli point process on  $Z_+ = \{1, 2, \dots\}$  with success probability  $p$ , arrival times  $\{\tilde{T}_i\}$  and counting process  $\{\tilde{N}_i\}$ . Let

$$X_i = b^i X_0 + ab \sum_{i=1}^{\tilde{N}_i} b^{i - \tilde{T}_i} \quad (26)$$

where  $a, b > 0$ ,

$$a \sum_{k=1}^{\infty} b^k \leq 1$$

and  $X_0$  is chosen to have the stationary distribution of the Markov chain  $\{X_i\}$ . A point process  $\{Y_i\}$  on  $Z_+$  is a discrete Neyman-Scott process with parameters  $a, b$ , and  $p$  if for all

positive integers  $n$  and 0-1 vectors  $(y_1, \dots, y_n)$

$$P\{Y_n = y_n, \dots, Y_1 = y_1 | X_r, r \geq 0\} \\ = \prod_{i=1}^n P\{Y_i = y_i | X_i\} = \prod_{i=1}^n y_i X_i + (1 - y_i)(1 - X_i) \quad (27)$$

Guttorp [1986] notes that the "number of discrete time models of rainfall is rather limited." He points in particular to the need to "develop discrete time cluster processes." In this section we have illustrated a rather flexible approach for accomplishing this task. In the next section we turn our attention to the important problems of parameter estimation and model selection with the hope that the attractive statistical inference procedures available for Markov chain models can be generalized for Markov Bernoulli models.

## 4. INTENSITY-BASED STATISTICAL INFERENCE

In this section we develop likelihood-based inference procedures for discrete point processes  $\{Y_i\}$ , establish asymptotic properties for maximum likelihood estimators of Markov Bernoulli model parameters, and apply inference procedures to a data set from Washington, D. C. The principal results are theorem 4.1, which exhibits the log likelihood function of a discrete point process as a simple function of the stochastic intensity and theorem 4.2, which asserts that maximum likelihood estimators of Markov Bernoulli model parameters are consistent and asymptotically normal and that log likelihood ratios have a limiting  $\chi^2$  distribution. Theorem 4.1 tells us how to estimate parameters; theorem 4.2 tells us how to assess standard errors and correlation of parameter estimates and how to select among competing classes of models. Theorem 4.2 guarantees that generalization of the Markov chain model to the Markov Bernoulli model is not done at the cost of losing the attractive statistical inference procedures described by Stern and Coe [1984].

For the Markov Bernoulli model presented in section 3 it is impractical to estimate all  $3T$  parameters of the parameter functions  $p, q_0$ , and  $q_1$ . For implementation of the model it is necessary to parameterize  $p, q_0$ , and  $q_1$  in the form

$$p(t) = f_1(t, \theta) \\ q_0(t) = f_2(t, \theta) \quad t = 1, \dots, T \quad \theta = (\theta_1, \dots, \theta_d) \quad (28) \\ q_1(t) = f_3(t, \theta)$$

where  $\theta \in \Theta$  and  $\Theta$  is a compact subset of  $R^d$ . Later in this section we will discuss practical aspects of the parameterization (28). Now we are concerned with establishing conditions under which asymptotic properties of maximum likelihood estimators of Markov Bernoulli model parameters hold. For this purpose we need to assume that the functions  $f_1, f_2$ , and  $f_3$  are continuous in  $\theta$  and possess continuous derivatives up to order 3. We now present a representation for the likelihood function of a discrete point process depending on  $d$  real-valued parameters,  $\theta$ .

### 4.1. Theorem 4.1

The log likelihood function of a point process  $\{Y_i\}$  with stochastic intensity  $\{\lambda_i\}$  taking values in  $(0, 1]$ , given observations over  $n$  years is

$$L_n(\theta) = \sum_{i=1}^n \sum_{t=1}^T Y_t^i \log \lambda_t^i + (1 - Y_t^i) \log (1 - \lambda_t^i) \quad (29)$$

4.2. Proof

The result follows immediately from (6). Model parameters "appear" on the right-hand side of (29) through the stochastic intensity.

The score functions are defined by

$$[U_n(\theta)]_j = \frac{\partial L_n(\theta)}{\partial \theta_j} \quad j = 1, \dots, d \quad (30)$$

It follows from (29) that

$$[U_n(\theta)]_j = \sum_{i=1}^n \sum_{s=1}^T \left( \frac{\partial \lambda_s^i}{\partial \theta_j} \right) [\lambda_s^i (1 - \lambda_s^i)]^{-1} (Y_s^i - \lambda_s^i) \quad (31)$$

The maximum likelihood estimator  $\hat{\theta}$  is the solution to the system of equations

$$U_n(\hat{\theta}) = 0 \quad (32)$$

The observed information matrix  $V_n(\theta)$  is defined by

$$[V_n(\theta)]_{j,k} = - \frac{\partial^2 L_n(\theta)}{\partial \theta_j \partial \theta_k} \quad (33)$$

The Fisher information matrix  $I(\theta)$  is reasonably defined for our model (see (38) and (39) below) by

$$[I(\theta)]_{j,k} = E \left[ - \frac{\partial^2 L_1(\theta)}{\partial \theta_j \partial \theta_k} \right] \quad (34)$$

Before presenting the main theorem we give results for the observed information matrix and Fisher information matrix.

4.3. Lemma 1

The observed information matrix is given by

$$[V_n(\theta)]_{j,k} = - \sum_{i=1}^n \sum_{s=1}^T \left[ \frac{\partial^2 \lambda_s^i}{\partial \theta_j \partial \theta_k} (\lambda_s^i (1 - \lambda_s^i))^{-1} - \frac{\partial \lambda_s^i}{\partial \theta_j} \frac{\partial \lambda_s^i}{\partial \theta_k} \frac{1 - 2\lambda_s^i}{(\lambda_s^i (1 - \lambda_s^i))^2} (Y_s^i - \lambda_s^i) + \sum_{i=1}^n \sum_{s=1}^T \frac{\partial \lambda_s^i}{\partial \theta_j} \frac{\partial \lambda_s^i}{\partial \theta_k} [\lambda_s^i (1 - \lambda_s^i)]^{-1} \right] \quad (35)$$

4.4. Lemma 2

The Fisher information matrix is given by

$$[I(\theta)]_{j,k} = \sum_{i=1}^T E \left[ \frac{\partial \lambda_s^i}{\partial \theta_j} \frac{\partial \lambda_s^i}{\partial \theta_k} (\lambda_s^i (1 - \lambda_s^i))^{-1} \right] \quad (36)$$

Lemma 1 follows from direct calculation of (33) (using the differentiability assumptions on  $f_1, f_2$ , and  $f_3$ ). Proofs of lemma 2 and theorem 4.2 are sketched in the Appendix. We denote below the true model parameters by  $\theta_0$ . As in theorem 4.1, the subscript  $n$  is for years.

4.5. Theorem 4.2

The following properties hold for maximum likelihood estimators of Markov Bernoulli process parameters.

$$\hat{\theta} \xrightarrow{P} \theta_0 \quad (37)$$

$$n^{1/2}(\hat{\theta} - \theta_0) \xrightarrow{D} N(0, I(\theta_0)^{-1}) \quad (38)$$

$$n^{-1} V_n(\hat{\theta}) \xrightarrow{P} I(\theta_0) \quad (39)$$

$$2(L_n(\hat{\theta}) - L_n(\theta_0)) \xrightarrow{D} \chi_d^2 \quad (40)$$

Practical significance of the theorem can be summarized as follows. From (37) we are guaranteed that for sufficiently large data sets, maximum likelihood estimators will, with high probability, be close to the true model parameters. Standard errors and correlations of parameter estimates can be assessed using (35), (38), and (39). Finally, (40) provides a tool for model selection that is particularly useful when competing models are embedded in one another (as, for example, Markov chain models are embedded in Markov Bernoulli models).

Note that the form of asymptotic properties of maximum likelihood estimators does not depend heavily on properties of the Markov Bernoulli model. Precisely the same asymptotic results ((37)-(40)) will hold in a much broader setting. Conditions under which asymptotic properties will hold can be obtained by mimicking conditions under which asymptotic properties of maximum likelihood estimators for continuous point processes hold (see Karr [1986], Section 9.2, or Ogata [1978]). In effect, one needs three types of conditions: (1) the stochastic intensity must be a "smooth" function of unknown parameters, (2) dependence of the stochastic intensity on the past must die out sufficiently quickly, and (3) the Fisher information matrix must be invertible. Using conditions of this type, one can derive asymptotic properties of maximum likelihood estimators for a broad class of discrete point process models, including the discrete Neyman-Scott model presented in section 3.

We conclude this section by applying inference procedures developed in this section to a 10-year record of daily precipitation data (1971-1980) from Washington, D. C. Our primary tasks are (1) to estimate parameters of a Markov Bernoulli model and (2) decide whether a Markov Bernoulli model is better than a Markov chain or Bernoulli trials model. Furthermore, we want to carry out these tasks for a range of precipitation thresholds. Below we present results for three "wet day" thresholds: 0.01, 0.10, and 1.00 inches. Figure 1 shows the unconditional wet day success probability by season for each of the wet day thresholds. Seasonality is clearly an important feature of the data.

Before implementing parameter estimation and model selection procedures we must parameterize the seasonal parameter functions  $p(t)$ ,  $q_0(t)$ , and  $q_1(t)$ . We will take our seasonal parameter functions to be of the form

$$f(t, \theta_1, \theta_2, \theta_3) = \frac{\exp \left\{ \theta_1 + \theta_2 \sin \left( \frac{2\pi t}{366} \right) + \theta_3 \cos \left( \frac{2\pi t}{366} \right) \right\}}{1 + \exp \left\{ \theta_1 + \theta_2 \sin \left( \frac{2\pi t}{366} \right) + \theta_3 \cos \left( \frac{2\pi t}{366} \right) \right\}} \quad (41)$$

where  $\theta_1, \theta_2$ , and  $\theta_3 \in R$ .

The following properties of the parameterization (41) are of primary importance.

1. The parameter functions are "probabilities," that is,

$$f(t, \theta_1, \theta_2, \theta_3) \in (0, 1) \quad t = 1, \dots, T \quad (42)$$

This condition is necessary because all of our parameter functions are either "transition probabilities" or "success probabilities."

2. For  $\theta_2 = \theta_3 = 0$ ,

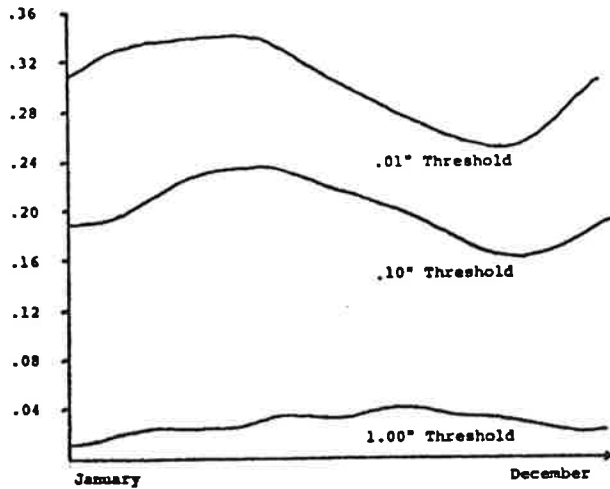


Fig. 1. Unconditional wet day success probability by season for Washington, D. C. (for thresholds of 0.01, 0.10, and 1.00 inches).

$$f(t, \theta_1, \theta_2, \theta_3) = \frac{\exp\{\theta_1\}}{1 + \exp\{\theta_1\}} \quad (43)$$

that is, the parameter functions do not vary with time of year.

3. For large positive values of  $\theta_1$  the parameter function is close to 1; for large negative values of  $\theta_1$  the parameter function is close to zero. In other words,

$$\lim_{\theta_1 \rightarrow \infty} f(t, \theta_1, \theta_2, \theta_3) = 1 \quad t = 1, \dots, T \quad (44)$$

$$\lim_{\theta_1 \rightarrow -\infty} f(t, \theta_1, \theta_2, \theta_3) = 0 \quad t = 1, \dots, T \quad (45)$$

Using (41) for the Markov Bernoulli model we obtain nine parameters,  $\theta = (\theta_1, \dots, \theta_9)$ , specified by

$$p(t) = f(t, \theta_1, \theta_2, \theta_3) \quad (46)$$

$$q_0(t) = f(t, \theta_4, \theta_5, \theta_6) \quad (47)$$

$$q_1(t) = f(t, \theta_7, \theta_8, \theta_9) \quad (48)$$

To estimate the nine parameters of the Markov Bernoulli model, the likelihood function is maximized numerically using the representation (29) for it (in terms of the stochastic intensity) and the representation (12) for the stochastic intensity of a Markov Bernoulli model. Initial parameter estimates are:  $\theta_1$  = a "large" positive value (see (44)),  $\theta_2 = \theta_3 = 0$  (see (43)) and  $\theta_4 - \theta_9$  equal to maximum likelihood estimates of Markov chain parameters obtained using the procedures described by Stern and Coe [1984]. Parameter estimates for threshold values of 0.01, 0.10, and 1.00 inches are given in Table 1.

From Table 1 it is seen that the estimated thinning function  $p(t)$  is a decreasing function of precipitation threshold. The Markov Bernoulli model is most similar to the Markov chain

TABLE 1. Markov Bernoulli Model Parameter Estimates for Wet Day Thresholds of 0.01, 0.10, and 1.00 Inches

Threshold	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$	$\theta_6$	$\theta_7$	$\theta_8$	$\theta_9$
0.01	2.1	-0.4	-0.2	1.0	-0.4	0.0	0.0	0.2	0.0
0.10	1.6	-0.2	-0.2	1.4	-0.3	0.0	-0.5	0.4	0.0
1.00	-0.5	-0.3	-0.2	3.4	-0.3	0.1	-2.4	0.2	0.1

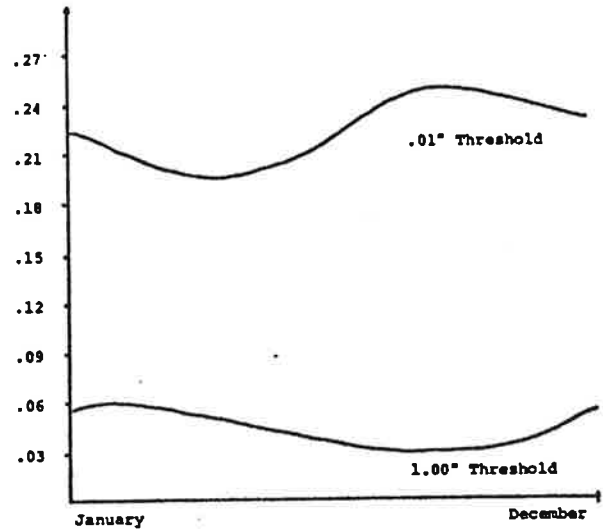


Fig. 2. Estimated seasonally varying values of the "temporal correlation"  $q_0 + q_1 - 1$  for thresholds of 0.01 and 1.00 inches.

for the smallest threshold value. For this threshold the thinning function varies seasonally about a value of 0.88.

Recall from section 3 that a measure of temporal correlation for the Markov Bernoulli model is  $q_0 + q_1 - 1$ . From Table 1 is seen that temporal correlation decreases with increasing threshold. For the 0.01-inch threshold,  $q_0 + q_1 - 1$  varies seasonally about a value of 0.22; for the 1.00-inch threshold,  $q_0 + q_1 - 1$  varies about a value of 0.05 (see Figure 2).

Qualitative features of parameter estimates suggest the following dependence of model selection on precipitation threshold. For small precipitation thresholds the Markov Bernoulli model is similar to the Markov chain model. For large thresholds the Markov Bernoulli model is indistinguishable from a Bernoulli trials model. For "moderate" thresholds, Markov Bernoulli models are different (and superior) to both Markov chain and Bernoulli trial models.

Likelihood ratio tests, based on (40) can be used to more formally assess questions of model suitability. A likelihood ratio test for model selection between Markov Bernoulli and Markov chain models is obtained by replacing the "true parameter"  $\theta_0$  in (40) with maximum likelihood estimators of a Markov chain model with parameters  $q_0$  and  $q_1$  specified by (47) and (48). Recall from section 3 that a Markov Bernoulli model with thinning function  $p$  identically equal to 1 is a Markov chain. The distribution of the test statistic is approximately  $\chi^2$  with 3 degrees of freedom. In similar fashion, a likelihood ratio test is constructed for model selection between Markov Bernoulli and Bernoulli trial models.

Table 2 contains log likelihood ratios (with approximate significance levels in parentheses) for tests of Markov Ber-

TABLE 2. Likelihood Ratio Test Values With Significance Levels in Parentheses

	Markov Bernoulli Versus Markov Chain	Markov Bernoulli Versus Bernoulli
0.01	2.5 (0.6)	>20 (>0.99)
0.10	3.8 (0.7)	>20 (>0.99)
1.00	<1.0 (<0.1)	<1 (<0.1)



noulli versus Markov chains and Markov Bernoulli versus Bernoulli trials. For a threshold of one inch, selection of a model more complex than a seasonal Bernoulli trials model is clearly unsupported. The Markov Bernoulli model is preferable to the Markov chain model for both the 0.01- and 0.10-inch thresholds. Preference for the Markov Bernoulli model is strongest at the 0.10-inch threshold.

5. SUMMARY AND CONCLUSIONS

The main points of the paper are summarized below.

1. A new family of discrete point process models of rainfall occurrences, termed Markov Bernoulli processes, is proposed. A Markov Bernoulli process can be viewed as a sequence of Bernoulli trials with randomized success probabilities. The family of Markov Bernoulli processes not only contains Markov chain and Bernoulli trial models, but also, both families can be represented as Markov Bernoulli models via simple parameterizations. Seasonality is easily accommodated in Markov Bernoulli models, as is the case with Markov chain models of wet-dry sequences. Unlike Markov chain models, Markov Bernoulli processes are invariant under random deletion of wet days. Random deletion arises in modeling wet-dry sequences when more than one precipitation threshold is used to define a wet day. Markov Bernoulli models can be easily generalized to more complex models of wet-day sequences by appropriate choice of the sequence of "randomized success probabilities." As an example, we present in section 3 a discrete analog to the Neyman-Scott model.

2. Likelihood-based inference procedures are developed for discrete point process models of wet-dry sequences. We obtain asymptotic properties for maximum likelihood estimators of Markov Bernoulli model parameters. In particular, we show in theorem 4.2 that maximum likelihood estimators are consistent and asymptotically normal and that log likelihood ratios have a limiting  $\chi^2$  distribution. These results provide the necessary tools for assessing standard errors and correlation of Markov Bernoulli model parameter estimators and for developing likelihood ratio tests for deciding whether Markov Bernoulli models are better than Markov chain and Bernoulli trial models.

3. Inference procedures developed for Markov Bernoulli models are applied to a data set from Washington, D. C. We present results for wet-dry sequences with precipitation thresholds of 0.01, 0.10, and 1.00 inches. Results illustrate dependence of model selection on precipitation threshold. For large precipitation thresholds there is little justification for resorting to more complex models than Bernoulli trial models. For the smallest threshold the Markov Bernoulli model is very nearly a Markov chain. With moderate thinning, the Markov Bernoulli model diverges from the Markov chain model.

The inference procedures we present provide not only quantitative tests for model selection but also tools that can be used to qualitatively evaluate significance of parameter estimates. These tools are of particular value in situations where physical interpretation is attributed to parameter estimates. Precipitation modeling has increasingly moved in the direction of physically based models. As increasingly sophisticated models of rainfall are developed, it is important that development of statistical tools keep pace.

APPENDIX

In this Appendix we sketch the proofs of lemma 2 and theorem 4.2. The key to obtaining asymptotic results for

intensity-based inference procedures is that the sequence  $\{Y_t' - \lambda_t'\}$  is a martingale difference series with respect to the data set  $\{\mathcal{H}_t'\}$ , that is,

$$E[(Y_t' - \lambda_t') | \mathcal{H}_{t-1}'] = 0 \quad t = 1, \dots, T \quad (A1)$$

$i = 1, 2, \dots$

The representation for the Fisher information matrix presented in lemma 2 follows from (A1) as follows. We have, from (29) and (30), that

$$I(\theta_0)_{j,k} = - \sum_{s=1}^T E[\mathcal{Z}_s'(Y_s^1 - \lambda_s^1)] + \sum_{s=1}^T E\left[\frac{\partial \lambda_s^1}{\partial \theta_j} \frac{\partial \lambda_s^1}{\partial \theta_k} (\lambda_s^1(1 - \lambda_s^1))^{-1}\right] \quad (A2)$$

where

$$\mathcal{Z}_s' = \frac{\partial^2 \lambda_s^1}{\partial \theta_j \partial \theta_k} (\lambda_s^1(1 - \lambda_s^1))^{-1} - \frac{\partial \lambda_s^1}{\partial \theta_j} \frac{\partial \lambda_s^1}{\partial \theta_k} \frac{1 - 2\lambda_s^1}{(\lambda_s^1(1 - \lambda_s^1))^2} \quad (A3)$$

depends only on observations in  $\mathcal{H}_{s-1}'$ . It follows that

$$I(\theta)_{j,k} = - \sum_{s=1}^T E\{E[\mathcal{Z}_s'(Y_s^1 - \lambda_s^1) | \mathcal{H}_{s-1}']\} + \sum_{s=1}^T E\left\{\frac{\partial \lambda_s^1}{\partial \theta_j} \frac{\partial \lambda_s^1}{\partial \theta_k} [\lambda_s^1(1 - \lambda_s^1)]^{-1}\right\} \quad (A4)$$

The lemma follows by noting that

$$E\{E[\mathcal{Z}_s'(Y_s^1 - \lambda_s^1) | \mathcal{H}_{s-1}']\} = E\{\mathcal{Z}_s' E[(Y_s^1 - \lambda_s^1) | \mathcal{H}_{s-1}']\} = 0 \quad (A5)$$

An important feature of Markov Bernoulli models is that  $\{\lambda_1^1\}$  has the same distribution as  $\{\lambda_t^1\}$  for all  $t$ . This property is used in defining the Fisher information and in the proof of theorem 4.2, which we sketch below.

To prove asymptotic normality, we first take a Taylor series expansion of the score function about  $\theta_0$ . We obtain

$$U_n(\theta) - U_n(\theta_0) = -V_n(\theta^*) (\theta - \theta_0) \quad (A6)$$

where  $\theta^*$  is on the line segment between  $\theta$  and  $\theta_0$ . Substituting the maximum likelihood estimator  $\hat{\theta}$  in (A6), we obtain

$$-U_n(\hat{\theta}_0) = -V_n(\hat{\theta}^*) (\hat{\theta} - \theta_0) \quad (A7)$$

Multiplying both sides by  $n^{-1/2}$  yields

$$n^{-1/2} U_n(\hat{\theta}_0) = [n^{-1} V_n(\hat{\theta}^*)] n^{1/2} (\hat{\theta} - \theta_0) \quad (A8)$$

To complete the proof of asymptotic normality, we need to show that

$$n^{-1} V_n(\hat{\theta}^*) \rightarrow I(\theta_0) \quad (A9)$$

and

$$n^{-1/2} U_n(\hat{\theta}_0) \rightarrow N[0, I(\theta_0)] \quad (A10)$$

We begin with (A10). Setting

$$Z_i = \sum_{s=1}^T \left(\frac{\partial \lambda_s^1}{\partial \theta}\right) [\lambda_s^1(1 - \lambda_s^1)]^{-1} (Y_s^1 - \lambda_s^1) \quad (A11)$$

we have, from (31), that

$$U_n(\hat{\theta}_0) = \sum_{i=1}^n Z_i \quad (A12)$$



Note that  $Z_i$  is a random vector of dimension  $d$ . Using computations analogous to (A5), it is straightforward to show that

$$E[Z_i] = 0 \quad (\text{A13})$$

$$E[Z_i^2]_{j,k} = [I(\theta_0)]_{j,k} \quad (\text{A14})$$

$$E[Z_i Z_j]_{j,k} = [0]_{j,k} \quad (\text{A15})$$

for  $i$  different from  $j$ .

Condition (A10) follows now from a standard central limit theorem for stationary ergodic sequences (see, for example, Karr [1986]). Similarly, consistency of the maximum likelihood estimator follows from (A12), (A13), and the strong law of large numbers for stationary ergodic sequences. By (A9) the proof of asymptotic normality is complete if we prove consistency of the observed information matrix as an estimator of the Fisher information matrix. In proving lemma 2 we carried out computations which show that

$$E[n^{-1}V_n(\theta_0)] = [I(\theta_0)]_{j,k} \quad (\text{A16})$$

The consistency result follows from lemma 1 of section 4 and the strong law of large numbers for martingale difference series using the martingale difference property of  $\{Y_t^i - \lambda_t^i\}$

To show that log likelihood ratios have a limiting  $\chi^2$  distribution, we take a Taylor series expansion of  $L_n(\theta)$  about the maximum likelihood estimator  $\hat{\theta}$ , obtaining,

$$L_n(\theta) = L_n(\hat{\theta}) + (\theta - \hat{\theta})U_n(\hat{\theta}) - \frac{1}{2}(\theta - \hat{\theta})V_n(\theta^*)\chi(\theta - \hat{\theta})^T \quad (\text{A17})$$

where  $\theta^*$  is on the line segment between  $\theta$  and  $\hat{\theta}$ . Evaluating (A17) at the true parameter  $\theta_0$  and using the fact the  $U_n(\hat{\theta})$  is by definition equal to zero, we obtain

$$L_n(\theta_0) - L_n(\hat{\theta}) = -\frac{1}{2}(\theta_0 - \hat{\theta})V_n(\theta^*)\chi(\theta_0 - \hat{\theta})^T \quad (\text{A18})$$

It follows that

$$-2[L_n(\theta_0) - L_n(\hat{\theta})] = n^{1/2}(\theta_0 - \hat{\theta})[n^{-1}V_n(\theta^*)]n^{1/2}(\theta_0 - \hat{\theta})^T \quad (\text{A19})$$

The result follows from asymptotic normality and consistency of the estimators  $\hat{\theta}$ .

*Acknowledgments.* This research was carried out at the Centre for Mathematics and Computer Science in Amsterdam with the support of a Fulbright postdoctoral research grant.

#### REFERENCES

- Foufoula-Georgiou, E., Discrete-time point process models for daily rainfall, Ph.D. dissertation, Univ. of Fla., Gainesville, 1985.
- Foufoula-Georgiou, E., and D. Lettenmaier, Continuous-time versus discrete-time point process models for daily rainfall, *Water Resour. Res.*, 22(4), 531-542, 1986.
- Goldberg, S., *Introduction to Difference Equations*, John Wiley, New York, 1958.
- Guttorp, P., On binary time series obtained from continuous point processes describing rainfall, *Water Resour. Res.*, 22(6), 897-904, 1986.
- Karr, A., *Point Processes and Their Statistical Inference*, Marcel Dekker, New York, 1986.
- Kavvas, M. L., and J. Delleur, A stochastic cluster model for daily rainfall sequences, *Water Resour. Res.*, 17(4), 1151-1160, 1981.
- Ogata, Y., The asymptotic behavior of maximum likelihood estimators for stationary point processes, *Ann. Inst. Stat. Math.*, 30, 243-261, 1978.
- Smith, J., and A. Karr, A point process model of summer season rainfall occurrences, *Water Resour. Res.*, 19(1), 95-103, 1983.
- Smith, J., and A. Karr, Statistical inference for point process models of rainfall, *Water Resour. Res.*, 21(1), 73-79, 1985.
- Smith, R., Discussion of: A model fitting analysis of daily rainfall data, *J. R. Stat. Soc. A*, 147(1), 44-45, 1984.
- Stern, R., and R. Coe, A model fitting analysis of daily rainfall data, *J. R. Stat. Soc. A*, 147(1), 1-34, 1984.
- Waymire, E., and V. Gupta, The mathematical structure of rainfall representations, 1, A review of the stochastic rainfall models, *Water Resour. Res.*, 17(5), 1261-1272, 1981.
- Woolhiser, D., and J. Roldan, Stochastic daily precipitation models, 2, A comparison of distribution of amounts, *Water Resour. Res.*, 18(5), 1461-1468, 1982.

J. A. Smith, Interstate Commission on the Potomac River Basin, 6110 Executive Boulevard, Suite 300, Rockville, MD 20852.

(Received July 17, 1986;  
revised January 20, 1987;  
accepted January 21, 1987.)