



Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

J.A. Smith

Statistical modeling of daily rainfall occurrences

Department of Mathematical Statistics

Report MS-R8605

August



Bibliotheek
Centrum voor Wiskunde en Informatica
Amsterdam

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Statistical Modeling of Daily Rainfall Occurrences

James A. Smith

*Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands*

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.

1980 Mathematics Subject Classification: 60G55, 62M99, 62F12.

Key Words & Phrases: discrete point processes, Markov Bernoulli models, maximum likelihood estimators, likelihood ratio tests, daily rainfall data.

1. INTRODUCTION

Markov chain models have played a major role in modeling wet-dry 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 [84]). In this paper we propose a new family of models for wet dry 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 [85] in developing "discrete point process" models of rainfall occurrences (see also FOUFOULA-GEORGIOU and LETTENMAIER [86]).

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 [81]) 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 [83]) 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-dry sequences (see WAYMIRE and GUPTA [81] 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 ARMA model. Our approach to model

Report MS-R8605

Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

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-dry 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 [85]). R. SMITH [84] 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 4.2 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 χ^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

$$Y_t^i(x) = 1(Z_t^i > x), \quad t = 1, \dots, T; \quad i = 1, 2, \dots \quad (2.1)$$

that is, $Y_t^i(x)$ equals 1 if Z_t^i is greater than the threshold x , otherwise $Y_t^i(x)$ equals 0. The threshold x is generally determined by the nature of the application and/or sampling thresholds of the station being used. The counting process

$$N_t^i(x) = \sum_{s=1}^t Y_s^i(x), \quad t = 1, \dots, T; \quad i = 1, 2, \dots \quad (2.2)$$

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

$$T_1^i(x) = \inf \left\{ t : Y_t^i(x) = 1 \right\}, \quad N_T^i(x) \geq 1 \quad (2.3)$$

$$T_k^i(x) = \inf \left\{ t > T_{k-1}^i(x) : Y_t^i(x) = 1 \right\}, \quad k \leq N_T^i(x).$$

The data set of wet day occurrences is denoted by the σ -algebra

$$\mathfrak{G}_t^i(x) = \sigma \left\{ Y_s^i(x) ; s \leq t \right\} \cup \sigma \left\{ Y_s^j(x) ; s = 1, \dots, T, j < i \right\}, \quad t = 1, \dots, T ; i = 1, 2, \dots \quad (2.4)$$

The data set $\mathfrak{G}_t^i(x)$ 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 stochastic intensity $\{\lambda_t^i\}$ of a point process $\{Y_t^i\}$ is defined by

$$\lambda_t^i = P \left\{ Y_t^i = 1 \mid \mathfrak{G}_{t-1}^i \right\}, \quad t = 1, \dots, T ; i = 1, 2, \dots \quad (2.5)$$

The value of λ_t^i is the conditional probability that day t of year i is a wet day given the history of preceding wet days.

In Section 4 statistical inference procedures based on the stochastic intensity (2.5) are developed for discrete point processes. SMITH and KARR [85] 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 $y^i = (y_1^i, \dots, y_T^i)$, $i = 1, \dots, n$,

$$\begin{aligned} P \left\{ Y_T^n = y_T^n, \dots, Y_1^n = y_1^n, \dots, Y_T^1 = y_T^1, \dots, Y_1^1 = y_1^1 \right\} \\ = \prod_{i=1}^n \prod_{s=1}^T P \left\{ Y_s^i = y_s^i \mid Y_{s-1}^i = y_{s-1}^i, \dots, Y_1^i = y_1^i \right\} \\ = \prod_{i=1}^n \prod_{s=1}^T y_s^i \lambda_s^i + (1 - y_s^i)(1 - \lambda_s^i). \end{aligned} \quad (2.6)$$

EXAMPLE 2.2 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

$$\lambda_t^i = p(t), \quad t = 1, \dots, T ; i = 1, 2, \dots \quad (2.7)$$

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 $p(t)$.

EXAMPLE 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

$$\lambda_t^i = q_1(t)Y_{t-1}^i + (1 - q_0(t))(1 - Y_{t-1}^i), \quad t = 1, \dots, T ; i = 1, 2, \dots \quad (2.8)$$

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 equation (2.8) is not well defined for $t=1$. Clearly, we want

$$Y_{t-1}^i = Y_T^{i-1} \quad (2.9)$$

whenever $t=1$ (and i is greater than 1). For $t=1$, equation (2.9) 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 (2.9) 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 $\{\tilde{Y}_t^i\}$ 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

$$X_t^i = p(t)\tilde{Y}_t^i, \quad t=1, \dots, T; i=1, 2, \dots \quad (3.1)$$

Note that on day t of any year the random success probability can take either the value $p(t)$ or 0. 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 $y^i = (y_1^i, \dots, y_T^i)$, $i=1, \dots, n$

$$\begin{aligned} P \left\{ Y_T^n = y_T^n, \dots, Y_1^n = y_1^n, \dots, Y_T^1 = y_T^1, \dots, Y_1^1 = y_1^1 \mid X_s^i = y_s^i, s=1, \dots, T, i=1, \dots, n \right\} \\ = \prod_{i=1}^n \prod_{s=1}^T P \left\{ Y_s^i = y_s^i \mid X_s^i = y_s^i \right\} \\ = \prod_{i=1}^n \prod_{s=1}^T y_s^i X_s^i + (1 - y_s^i)(1 - X_s^i) \end{aligned} \quad (3.2)$$

It follows from equation (3.2) that:

- 1) The sequence of wet-dry days are 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 $p(t)$.
- 3) The conditional probability of a wet day, given that X_t^i equals 0, is 0.

The main distributional result for Markov Bernoulli processes is the following recursive formula for the stochastic intensity.

THEOREM 3.1

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

$$\lambda_t^i = p(t)[1 - q_0(t)] + p(t)[q_0(t) + q_1(t) - 1] [Y_{t-1}^i - (1 - Y_{t-1}^i) \frac{1 - p(t-1)}{p(t-1)} \lambda_{t-1}^i] \quad (3.3)$$

PROOF.

$$\begin{aligned}
\lambda_t^i &= E[Y_t^i | \mathfrak{C}_{t-1}^i] \\
&= E[E[Y_t^i | X_t^i] | \mathfrak{C}_{t-1}^i] \\
&= E[X_t^i | \mathfrak{C}_{t-1}^i] \\
&= p(t) P \left\{ X_t^i = p(t) | \mathfrak{C}_{t-1}^i \right\} \\
&= p(t) \left[P \left\{ X_t^i = p(t), X_{t-1}^i = p(t-1) | \mathfrak{C}_{t-1}^i \right\} + P \left\{ X_t^i = p(t), X_{t-1}^i = 0 | \mathfrak{C}_{t-1}^i \right\} \right] \\
&= p(t) \left[P \left\{ X_t^i = p(t) | X_{t-1}^i = p(t-1) \right\} P \left\{ X_{t-1}^i = p(t-1) | \mathfrak{C}_{t-1}^i \right\} \right. \\
&\quad \left. + P \left\{ X_t^i = p(t) | X_{t-1}^i = 0 \right\} P \left\{ X_{t-1}^i = 0 | \mathfrak{C}_{t-1}^i \right\} \right] \\
&= p(t) \left[q_1(t) P \left\{ X_{t-1}^i = p(t-1) | \mathfrak{C}_{t-1}^i \right\} + (1 - q_0(t)) (1 - P \left\{ X_{t-1}^i = p(t-1) | \mathfrak{C}_{t-1}^i \right\}) \right] \\
&= p(t) \left[(1 - q_0(t)) + (q_0(t) + q_1(t) - 1) P \left\{ X_{t-1}^i = p(t-1) | \mathfrak{C}_{t-1}^i \right\} \right] \\
&= p(t) (1 - q_0(t)) \\
&\quad + p(t) (q_0(t) + q_1(t) - 1) [Y_{t-1}^i + (1 - Y_{t-1}^i) (1 - p(t-1)) P \left\{ X_{t-1}^i = p(t-1) | \mathfrak{C}_{t-2}^i \right\}] \\
&= p(t) (1 - q_0(t)) + p(t) (q_0(t) + q_1(t) - 1) [Y_{t-1}^i + (1 - Y_{t-1}^i) \frac{(1 - p(t-1))}{p(t-1)} \lambda_{t-1}^i].
\end{aligned} \tag{3.4}$$

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

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, $q_0(t) + q_1(t) - 1 = 0$, $t=1, \dots, T$.*

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 (3.3) with (2.7); the second corollary follows by comparing (3.3) with (2.8). 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 [83]), 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 (i.i.d.). 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

$$f(k) = P \left\{ U_j = k \right\}, \quad k = 1, 2, \dots \tag{3.5}$$

or hazard function

$$h(k) = P\left\{U_j = k | U_j > k - 1\right\}, \quad k \geq 2 \quad (3.6)$$

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

The two are related by

$$h(k) = \frac{f(k)}{1 - \sum_{j=1}^{k-1} f(j)}, \quad k = 1, 2, \dots \quad (3.7)$$

Importance of the hazard function is clear from the following lemma (KARR [86]).

LEMMA *The stochastic intensity of a renewal process is given by*

$$\lambda_t^i = h(V_t^i) \quad (3.8)$$

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.

COROLLARY 3. *A Markov Bernoulli process with parameters p , q_0 and q_1 , which do not vary with time of year is a renewal process with hazard function*

$$h(k) = \frac{p(1-q_0)}{p + (1-p)(q_0 + q_1)} + p \left[\frac{1 - (1-p)q_1}{(1-p)[1 - (q_0 + q_1 - 1)(1-p)]} \right] [(1-p)(q_0 + q_1 - 1)]^k \quad (3.9)$$

PROOF. It follows from (3.3) that

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

where h is the solution to the first order difference equation

$$h(k) = p(1-q_0) - (q_0 + q_1 - 1)(1-p)h(k-1) \quad (3.10)$$

with initial condition

$$h(1) = pq_1$$

The solution to (3.10) is given by (3.9) (see, for example, GOLDBERG [58]).

Remarks.

1) If $q_0 + q_1 - 1$ is greater than 0, h is a decreasing function of k . If $q_0 + q_1 - 1$ is less than 0, 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 0, 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 0, 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 (3.3) that

$$\lambda_t^i = h_t(V_t^i) \quad (3.11)$$

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

$$h_t(k) = p(t)(1 - q_0(t)) - (q_0(t) + q_1(t) - 1)(1 - p(t))h_{t-1}(k-1) \quad (3.12)$$

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

Existence of solutions and general expressions for solutions can be obtained from results in GOLDBERG [58]. While (3.11) does not provide a useful computational tool (in the way that (3.9) 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 (3.8)) through the stochastic intensity using the representation (3.11).

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 $\{\tilde{Z}_t^j\}$ be a Bernoulli point process with success probability p, which is independent of the point process $\{Y_t^j\}$. The point process $\{\tilde{Y}_t^j\}$ specified by

$$\tilde{Y}_t^j = \tilde{Z}_t^j Y_t^j, \quad t=1, \dots, T; \quad i=1, 2, \dots \quad (3.13)$$

is termed a p-thinning of $\{Y_t^j\}$. From (3.13) it is seen that the p-thinning $\{\tilde{Y}_t^j\}$ is constructed by randomly deleting events of $\{Y_t^j\}$; an event of $\{Y_t^j\}$ 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.

COROLLARY 4. *The p-thinning $\{\tilde{Y}_t^j\}$ of a Markov Bernoulli process $\{Y_t^j\}$ with parameters \tilde{p} , q_0 and q_1 is a Markov Bernoulli process with parameters $\tilde{p}p$, q_0 and q_1 .*

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 [84] and WOOLHISER and ROLDAN [82]), that is

$$P\left\{Z_t^i \leq x | Z_t^i > 0, Z_{t-1}^i, \dots, Z_1^i\right\} = P\left\{Z_t^i \leq x | Z_t^i > 0\right\} \\ = H_t(x) \quad (3.14)$$

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

$$p(t) = 1 - H_t(x), \quad t=1, \dots, T \quad (3.15)$$

It follows from the preceding corollary that if the conditional independence property (3.14) 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 (3.14).

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 $\{X_t^i\}$ for a Markov Bernoulli process is a simple Markov chain. By allowing $\{X_t^i\}$ to be more complicated Markov chains 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 Markov chains is that the stochastic intensity of the resulting models can often be computed from "filter equations" analagous to (3.3).

A constructive definition of a discrete Neyman-Scott model would procede along the following lines (for notational simplicity we consider a stationary model on the positive integers). Let $\{T_n\}$ 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 center 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 0 or 1 for all i , we can not avoid the possibility of having more than one event on a single day. This approach clearly will not work.

SMITH and KARR [85] 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.

LEMMA (SMITH and KARR [85]) *Let N be a Neyman-Scott process on $[0, \infty)$ of the following form.*

- 1) *The Poisson process of cluster centers is N ; arrival times are denoted T .*
- 2) *The distribution of cluster sizes is Poisson with parameter a .*
- 3) *The distances from each cluster center to its cluster members are i.i.d exponentially distributed with parameter b .*

Then N is a Cox process directed by the Markov process

$$X(t) = \exp(-bt)X(0) + ab \sum_{i=1}^{\tilde{N}_i} \exp(-b(t - \tilde{T}_i)) \quad (3.16)$$

The Neyman-Scott process N can be interpreted as a Poisson process with randomized rate of occurrence given by the Markov process (3.16).

By analogy with (3.16) we introduce the following definition of a discrete Neyman-Scott model. Let $\{\tilde{Y}_t\}$ be a stationary Bernoulli point process on $Z_+ = \{1, 2, \dots\}$ with success probability p , arrival times $\{\tilde{T}_t\}$ and counting process $\{\tilde{N}_t\}$. Let

$$X_t = b^t X_0 + ab \sum_{i=1}^{\tilde{N}_t} b^{t - \tilde{T}_i} \quad (3.17)$$

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 $\{X_t\}$. A point process $\{Y_t\}$ 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)

$$\begin{aligned} P \left\{ Y_n = y_n, \dots, Y_1 = y_1 | X_t, t \geq 0 \right\} &= \prod_{t=1}^n P \left\{ Y_t = y_t | X_t \right\} \\ &= \prod_{t=1}^n y_t X_t + (1 - y_t)(1 - X_t) \end{aligned} \quad (3.18)$$

4. INTENSITY-BASED STATISTICAL INFERENCE

In this section we develop likelihood-based inference procedures for discrete point processes $\{Y_t^i\}$, 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 χ^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 [84].

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

$$\begin{aligned} p(t) &= f_1(t, \theta) \\ q_0(t) &= f_2(t, \theta), \quad t=1, \dots, T; \theta = (\theta_1, \dots, \theta_d) \\ q_1(t) &= f_3(t, \theta) \end{aligned} \quad (4.1)$$

where $\theta \in \Theta$ and Θ is a compact subset of R^d . Later in this section we will discuss practical aspects of the parameterization (4.1). 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 θ 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, θ .

THEOREM 4.1 *The log-likelihood function of a point process $\{Y_t^i\}$ with stochastic intensity $\{\lambda_t^i\}$ taking values in $(0,1]$, given observations over n years is*

$$L_n(\theta) = \sum_{i=1}^n \sum_{s=1}^T Y_s^i \log \lambda_s^i + (1 - Y_s^i) \log(1 - \lambda_s^i) \quad (4.2)$$

PROOF. The result follows immediately from (2.6).

The score functions are defined by

$$[U_n(\theta)]_j = \frac{\partial L_n(\theta)}{\partial \theta_j}, \quad j=1, \dots, d \quad (4.3)$$

It follows from (4.2) that

$$[U_n(\theta)]_j = \sum_{i=1}^n \sum_{s=1}^T \left(\frac{\partial \lambda_s^i}{\partial \theta_j} \right) [\lambda_s^i (1 - \lambda_s^i)]^{-1} (Y_s^i - \lambda_s^i) \quad (4.4)$$

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

$$U_n(\theta) = 0 \quad (4.5)$$

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

$$[V_n(\theta)]_{j,k} = - \frac{\partial^2 L_n(\theta)}{\partial \theta_j \partial \theta_k} \quad (4.6)$$

The Fisher information matrix $I(\theta)$ is defined by

$$[I(\theta)]_{j,k} = E \left[- \frac{\partial^2 L_1(\theta)}{\partial \theta_j \partial \theta_k} \right]_{j,k} \quad (4.7)$$

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

LEMMA 1 *The observed information matrix is given by*

$$\begin{aligned} [V_n(\theta)]_{j,k} = & - \sum_{i=1}^n \sum_{s=1}^T \left[\frac{\partial^2 \lambda_s^i}{\partial \theta_j \partial \theta_k} \left\{ \lambda_s^i (1 - \lambda_s^i) \right\}^{-1} - \frac{\partial \lambda_s^i}{\partial \theta_j} \frac{\partial \lambda_s^i}{\partial \theta_k} \frac{1 - 2\lambda_s^i}{(\lambda_s^i (1 - \lambda_s^i))^2} \right] (Y_s^i - \lambda_s^i) \\ & + \sum_{i=1}^n \sum_{s=1}^T \frac{\partial \lambda_s^i}{\partial \theta_j} \frac{\partial \lambda_s^i}{\partial \theta_k} [\lambda_s^i (1 - \lambda_s^i)]^{-1} \end{aligned} \quad (4.8)$$

LEMMA 2 *The Fisher information matrix is given by*

$$[I(\theta)]_{j,k} = \sum_{s=1}^T E \left[\frac{\partial \lambda_s^i}{\partial \theta_j} \frac{\partial \lambda_s^i}{\partial \theta_k} (\lambda_s^i (1 - \lambda_s^i))^{-1} \right] \quad (4.9)$$

Lemma 1 follows from direct calculation of (4.6) (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 θ_0 .

THEOREM 4.2 *The following properties hold for maximum likelihood estimators of Markov Bernoulli process parameters.*

$$\hat{\theta} \rightarrow \theta_0 \quad (4.10)$$

$$n^{1/2}(\hat{\theta} - \theta_0) \rightarrow N(0, I(\theta_0)^{-1}) \quad (4.11)$$

$$n^{-1} V_n(\hat{\theta}) \rightarrow I(\theta_0) \quad (4.12)$$

$$2(L_n(\hat{\theta}) - L_n(\theta_0)) \rightarrow \chi_d^2 \quad (4.13)$$

Practical significance of the theorem can be summarized as follows. From (4.10) we are guaranteed that for sufficiently large data sets, maximum likelihood estimators will be close to the true model parameters. Standard errors and correlations of parameter estimates can be assessed using (4.8), (4.11) and (4.12). Finally (4.13) 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 for Markov Bernoulli model parameters does not depend heavily on properties of the Markov Bernoulli model. Precisely the same asymptotic results (4.10 - 4.13) 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 [86], Section 9.2, or OGATA [78]). In effect one needs three types of conditions: a) the stochastic intensity must be a "smooth" function of unknown parameters b) dependence of the stochastic intensity on the past must die out sufficiently quickly, and c) 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: .01", .10", and 1.00".

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

$$f(t, \theta_1, \theta_2, \theta_3) = \frac{\exp\left\{\theta_1 + \theta_2 \sin\left(\frac{2\pi t}{366}\right) + \theta_3 \cos\left(\frac{2\pi t}{366}\right)\right\}}{1 + \exp\left\{\theta_1 + \theta_2 \sin\left(\frac{2\pi t}{366}\right) + \theta_3 \cos\left(\frac{2\pi t}{366}\right)\right\}} \quad (4.14)$$

where θ_1, θ_2 , and $\theta_3 \in \mathbb{R}$.

The following properties of the parameterization (4.14) are of primary importance.

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

$$f(t, \theta_1, \theta_2, \theta_3) \in (0, 1) \quad , \quad t = 1, \dots, T \quad (4.15)$$

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

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

$$f(t, \theta_1, \theta_2, \theta_3) = \frac{\exp\left\{\theta_1\right\}}{1 + \exp\left\{\theta_1\right\}} \quad (4.16)$$

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

3) For large positive values of θ_1 the parameter function is close to 1; for large negative values of θ_1 the parameter function is close to 0. More precisely,

$$\lim_{\theta_1 \rightarrow \infty} f(t, \theta_1, \theta_2, \theta_3) = 1 \quad , \quad t = 1, \dots, T \quad (4.17)$$

$$\lim_{\theta_1 \rightarrow -\infty} f(t, \theta_1, \theta_2, \theta_3) = 0 \quad , \quad t = 1, \dots, T \quad (4.18)$$

Using (4.14) for the Markov Bernoulli model we obtain 9 parameters $\theta = (\theta_1, \dots, \theta_9)$ specified by

$$p(t) = f(t, \theta_1, \theta_2, \theta_3) \quad (4.19)$$

$$q_0(t) = f(t, \theta_4, \theta_5, \theta_6) \quad (4.20)$$

$$q_1(t) = f(t, \theta_7, \theta_8, \theta_9) \quad (4.21)$$

To estimate the nine parameters of the Markov Bernoulli model the likelihood function is maximized numerically using the representation (4.2) for the likelihood function (in terms of the stochastic intensity) and the representation (3.3) for the stochastic intensity of a Markov Bernoulli model. Initial parameter estimates are: $\theta_1 =$ a "large" positive value (see (4.17)), $\theta_2 = \theta_3 = 0$ (see 4.16), and $\theta_4 - \theta_9$ equal to maximum likelihood estimates of Markov chain parameters obtained using the procedures described by STERN and COE [84]. Parameter estimates for threshold values of ".01", ".10", and "1.00" are given in Table 1.

Threshold	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$\hat{\theta}_7$	$\hat{\theta}_8$	$\hat{\theta}_9$
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

TABLE 1. Markov Bernoulli model parameter estimates for wet day thresholds of 0.01, 0.10 and 1.00 inches.

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 for the smallest threshold value. For this threshold the thinning function varies seasonally about a value of .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 it is seen that temporal correlation decreases with increasing threshold. For the .01 inch threshold, $q_0 + q_1 - 1$ varies seasonally about a value of .3; for the 1.00 inch threshold, $q_0 + q_1 - 1$ varies about a value of .05.

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 (4.13), 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" θ_0 in (4.13) with maximum likelihood estimators of a Markov chain model with parameters q_0 and q_1 specified by (4.20) and (4.21). 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 χ^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.

x	Markov Bernoulli	Markov Bernoulli
	vs	vs
	Markov Chain	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 (>0.99)

TABLE 2. Likelihood ratio test values with significance levels in parenthesis.

Table 2 contains log-likelihood ratios (with approximate significance levels in parenthesis) for tests of Markov Bernoulli 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 clearly preferable to the Markov chain model for the .10 inch threshold. For the .01 inch threshold, superiority of the Markov Bernoulli model is marginal.

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 dry 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 χ^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 .01, .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.

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

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 $\{Y_t^i - \lambda_t^i\}$ is a martingale difference series with respect to the data set $\{\mathfrak{C}_t^i\}$, that is,

$$E[(Y_t^i - \lambda_t^i) | \mathfrak{C}_{t-1}^i] = 0, \quad t=1, \dots, T; i=1, 2, \dots \quad (\text{A.1})$$

The representation for the Fisher information matrix presented in Lemma 2 follows from (A.1) as follows. We have, from (4.2) and (4.3) that

$$[I(\theta_0)]_{j,k} = - \sum_{s=1}^T E[\tilde{Z}_s^1 (Y_s^1 - \lambda_s^1)] + \sum_{s=1}^T E\left[\frac{\partial \lambda_s^1}{\partial \theta_j} \frac{\partial \lambda_s^1}{\partial \theta_k} (\lambda_s^1 (1 - \lambda_s^1))^{-1}\right], \quad (\text{A.2})$$

where

$$\tilde{Z}_s^1 = \frac{\partial^2 \lambda_s^1}{\partial \theta_j \partial \theta_k} (\lambda_s^1 (1 - \lambda_s^1))^{-1} - \frac{\partial \lambda_s^1}{\partial \theta_j} \frac{\partial \lambda_s^1}{\partial \theta_k} \frac{1 - 2\lambda_s^1}{(\lambda_s^1 (1 - \lambda_s^1))^2} \quad (\text{A.3})$$

depends only on observations in \mathfrak{C}_{s-1}^1 . It follows that

$$[I(\theta)]_{j,k} = - \sum_{s=1}^T E[E[\tilde{Z}_s^1 (Y_s^1 - \lambda_s^1) | \mathfrak{C}_{s-1}^1]] + \sum_{s=1}^T E\left[\frac{\partial \lambda_s^1}{\partial \theta_j} \frac{\partial \lambda_s^1}{\partial \theta_k} (\lambda_s^1 (1 - \lambda_s^1))^{-1}\right] \quad (\text{A.4})$$

The lemma follows by noting that

$$\begin{aligned} E[E[\tilde{Z}_s^1 (Y_s^1 - \lambda_s^1) | \mathfrak{C}_{s-1}^1]] &= E[\tilde{Z}_s^1 E[(Y_s^1 - \lambda_s^1) | \mathfrak{C}_{s-1}^1]] \\ &= 0. \end{aligned} \quad (\text{A.5})$$

An important feature of Markov Bernoulli models is that $\{\lambda_i^1\}$ has the same distribution as $\{\lambda_1^i\}$ for all i . 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 θ_0 . We obtain,

$$U_n(\theta) - U_n(\theta_0) = -V_n(\theta^*)(\theta - \theta_0) \quad (\text{A.6})$$

where θ^* is on the line segment between θ and θ_0 . Substituting the maximum likelihood estimator $\hat{\theta}$ in (A.6) we obtain

$$-U_n(\theta_0) = -V_n(\theta^*)(\hat{\theta} - \theta_0) \quad (\text{A.7})$$

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

$$n^{-1/2}U_n(\theta_0) = [n^{-1}V_n(\theta^*)]n^{1/2}(\hat{\theta} - \theta_0) \quad (\text{A.8})$$

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

$$n^{-1}V_n(\theta^*) \rightarrow I(\theta_0) \quad (\text{A.9})$$

and

$$n^{-1/2}U_n(\theta_0) \rightarrow N(0, I(\theta_0)) \quad (\text{A.10})$$

We begin with (A.10). Setting

$$Z_i = \sum_{s=1}^T \left(\frac{\partial \lambda_s^i}{\partial \theta} \right) [\lambda_s^i (1 - \lambda_s^i)]^{-1} (Y_s^i - \lambda_s^i) \quad (\text{A.11})$$

we have from (4.4) that

$$U_n(\theta_0) = \sum_{i=1}^n Z_i \quad (\text{A.12})$$

Note that Z_i is a random vector of dimension d . Using computations analogous to (A.5) it is straightforward to show that

$$E[Z_i] = 0 \quad (\text{A.13})$$

$$E[Z_i^2]_{j,k} = [I(\theta_0)]_{j,k} \quad (\text{A.14})$$

$$E[Z_i Z_l]_{j,k} = [0]_{j,k} \quad (\text{A.15})$$

for i different from l .

Condition (A.10) follows now from a standard central limit theorem for stationary ergodic sequences (see, for example, Karr [86]). Similarly, consistency of the maximum likelihood estimator follows from (A.12), (A.13), and the strong law of large numbers for stationary ergodic sequences. By (A.9) 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

$$E[n^{-1}V_n(\theta_0)] = [I(\theta_0)]_{j,k} \quad (\text{A.16})$$

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 $\{Y_i^i - \lambda_i^i\}$.

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

$$L_n(\theta) = L_n(\hat{\theta}) + (\theta - \hat{\theta})U_n(\hat{\theta}) - \frac{1}{2}(\theta - \hat{\theta})V_n(\theta^*)(\theta - \hat{\theta})^T \quad (\text{A.17})$$

where θ^* is on the line segment between θ and $\hat{\theta}$. Evaluating (A.17) at the true parameter θ_0 , and using the fact the $U_n(\hat{\theta})$ is by definition equal to 0, we obtain

$$L_n(\theta_0) - L_n(\hat{\theta}) = -\frac{1}{2}(\theta_0 - \hat{\theta})V_n(\theta^*)(\theta_0 - \hat{\theta})^T \quad (\text{A.18})$$

It follows that

$$-2(L_n(\theta_0) - L_n(\hat{\theta})) = n^{1/2}(\theta_0 - \hat{\theta})[n^{-1}V_n(\theta^*)]n^{1/2}(\theta_0 - \hat{\theta})^T. \quad (\text{A.19})$$

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

REFERENCES

- 1 FOUFOULA-GEORGIU E. *Discrete-time point process models for daily rainfall*, Ph.D Dissertation, University of Florida, 1985.
- 2 FOUFOULA-GEORGIU, E. AND D. LETTENMAIER *Continuous-time versus discrete time point process models for daily rainfall*, Water Resources Research, to appear.
- 3 GOLDBERG, S. *Introduction to Difference Equations*, John Wiley and Sons, New York, 1958.
- 4 KARR, A. *Point Processes and their Statistical Inference*, Marcel Dekker, New York, 1986.
- 5 KAVVAS, M. L. AND J. DELLEUR *A stochastic cluster model for daily rainfall sequences*, Water Resources Research, 17(4), 1151-1160, 1981.
- 6 OGATA Y. The asymptotic behavior of maximum likelihood estimators for stationary point processes, *Ann. Inst. Statist. Math.*, 30, 243-261, 1978.
- 7 SMITH, J. AND A. KARR *A point process model of summer season rainfall occurrences*, Water Resources Research, 19(1), 95-103, 1983.
- 8 SMITH, J. AND A. KARR *Statistical inference for point process models of rainfall*, Water Resources Research, 21(1), 73-79, 1985.
- 9 SMITH, R. Discussion of; A model fitting analysis of daily rainfall data, *J. R. Statist. Soc. A*, 147(1), 24-25, 1984.
- 10 STERN, R. AND R. COE A model fitting analysis of daily rainfall data, *J. R. Statist. Soc. A*, 147(1), 1-34, 1984.
- 11 WAYMIRE, E. AND V. GUPTA The mathematical structure of rainfall representations. 1. *A review of the stochastic rainfall models*, Water Resources Research, 17, 1261-1272, 1981.
- 12 WOOLHISER, D. AND J. ROLDAN Stochastic daily precipitation models, 2. *A comparison of distribution of amounts*, Water Resources Research, 18(5), 1461-1468, 1982.