The Annals of Statistics
1986, Vol. 14, No. 4, 1583–1590

ASYMPTOTIC INFERENCE FOR A CHANGE-POINT POISSON PROCESS

BY V. E. AKMAN[1] AND A. E. RAFTERY

Trinity College, Dublin

Easily implemented asymptotic off-line procedures for the change-point Poisson process with $\lambda(t)$, the intensity at time t, equal to λ_1 if $t \leq \tau$ and to λ_2 if $t > \tau$, are developed. They may also be applied to a problem of estimation of the location of a discontinuity in density discussed by Chernoff and Rubin (1956). A test for change is noted, a test of the hypothesis that $\tau = \tau_0$ is proposed, and point and interval estimates of τ, λ_1, and λ_2 are provided. The small-sample performance of the proposed procedures is studied using simulation, and an example is given.

1. Introduction. Off-line inference about a Poisson process with a change-point is considered. Inference is to be based on an observation period $[0, T]$ during which n events have been observed at times $t_1, \dots, t_n$. The rate of occurrence at time t, denoted by $\lambda(t)$, is equal to λ_1 if $0 \leq t \leq \tau$, and to λ_2 if $\tau < t \leq T$. The change-point t is unknown, as are λ_1 and λ_2.

Our approach here is non-Bayesian, based largely on asymptotic approximations. In order to obtain asymptotic results, it is necessary to embed the problem in a sequence of problems, and there are several ways in which this can be done. Here we assume that $\tau \to \infty$ and $T \to \infty$ in such a way that $\tau/T = \theta$, a constant.

Under this assumption, it was pointed out by Rubin (1961) that the present problem is equivalent to that of inference for a random sample from a density which is equal to β on $[0, \alpha]$, to γ on $[\alpha, 1]$, and to zero elsewhere, where α, β, and γ are unknown. The latter problem was considered by Chernoff and Rubin (1956), who reduced the problem of finding the asymptotic distribution of the maximum likelihood estimators to the distribution of a corresponding function of a stochastic process. According to Rubin (1961), Breakwell and Chernoff, in some unpublished memoranda, obtained the asymptotic distribution of those estimators.

Deshayes (1984) stated a similar result for the change-point Poisson process, but noted that it does not yield operational inference procedures for the change-point. He did obtain inference procedures under the additional assumption that $\lambda_2/\lambda_1 \to 1$, but this approach does not seem designed to provide good approximations if λ_1 and λ_2 are, in fact, widely separated.

Received May 1985; revised January 1986.

[1]Supported by postgraduate research awards from Trinity College, Dublin and the Irish Department of Education.

AMS 1980 *subject classifications.* Primary 62M02, 62M05, 60G55; secondary 62F03, 62F04, 62F05, 62F10, 62F11, 62F12, 62F25.

Key words and phrases. Coal-mining disasters; confidence interval; consistent estimator; discontinuous density; hypothesis test; simulation; small-sample properties.

Our main purpose here is to propose simple, easily implemented, inference procedures for the change-point which do not require that $\lambda_2/\lambda_1 \to 1$; this is done in Section 3. These may also, of course, be applied to the discontinuous density estimation problem of Chernoff and Rubin (1956), described earlier. Before that, in Section 2, we introduce estimators for θ, λ_1, and λ_2 which are related to the inference procedures proposed. We show that they are consistent and that the estimators of λ_1 and λ_2 satisfy a central limit theorem with random norming. In Section 4 we investigate the small-sample performance of the proposed procedures by means of a Monte-Carlo study, and in Section 5 they are applied to a real data set.

Other approaches to the present problem include those of Leonard (1978), who proposed estimating the parameters by minimising the integrated squared difference between the estimated rate and a different, nonparametric, estimate of the rate function; Kendall and Kendall (1980) who discussed a test for change; and Raftery and Akman (1986) who developed a Bayesian approach. Of course, if the change is known to have occurred at an event time, the problem reduces to that of a change-point in a sequence of independent exponential random variables, and methods designed for that case, such as may be derived directly from the results of Hinkley (1970), can be used. This assumption underlies the work of Commenges and Seal (1985).

Different, but related, problems have been worked on by Kalbfleisch and Struthers (1982) who considered a Poisson process analogue of intervention analysis, Kutoyants (1984) who studied the statistical analysis of a Poisson process with a periodic discontinuous rate function, Matthews and Farewell (1982) and Nguyen, Rogers, and Walker (1984) who analysed hazard rates with change-points, and Schulze (1984) who gave results for growth curves with change-points.

2. Estimation. The estimation and inference procedures we consider are all based on the process

$$Y(s; c, d) = \{s'(1 - s')\}^{1/2}\left\{\frac{\Pi(s) - \Pi(c)}{s'} - \frac{\Pi(d) - \Pi(s)}{1 - s'}\right\}, \quad c \le s \le d,$$

where $0 \le c < d \le 1$, $s' = (s - c)/(d - c)$, $\Pi(s) = N(sT)$, and $N(t)$ is the number of events that occurred in the time interval $(0, t]$. $Y(s; c, d)$ is the normalised difference between the mean intensities on $(cT, sT]$ and $(sT, dT]$. Also, if $\{t_1, \dots, t_n\}$ is viewed as the set of order statistics from a random sample, then $Y(s; 0, 1)$ is essentially the process which underlies the goodness-of-fit tests of Anderson and Darling (1952) in the case of testing for uniformity.

We assume that there are known constants a and b such that $0 < a \le \theta \le b < 1$. We suppose that θ, λ_1, and λ_2 are unknown, but that it is known that $\lambda_1 > \lambda_2$. The analysis is similar if it is known that $\lambda_1 < \lambda_2$. We consider the estimators

$$\hat{\theta} = \inf\left\{s\colon Y(s;0,1) = \sup_{a \le u \le b} Y(u;0,1)\right\},$$
$$\hat{\lambda}_1 = \hat{\nu}/\hat{\tau}, \quad \hat{\lambda}_2 = (N(T) - \hat{\nu})/(T - \hat{\tau}),$$

where $\hat{\tau} = \hat{\theta}T$ and $\hat{\nu} = N(\hat{\tau})$.

THEOREM 1. *The following asymptotic statements, which refer to limits taken as $T \to \infty$ and $\tau \to \infty$ in such a way that $\tau/T = \theta$ remains constant, hold*:

(i) *The estimators $\hat{\theta}$, $\hat{\lambda}_1$, and $\hat{\lambda}_2$ are consistent.*

(ii)

$$(\lambda_1\hat{\tau} - \hat{\nu})/\hat{\nu}^{1/2} \to_d N(0,1), \tag{2.1}$$

$$\{\lambda_2(T - \hat{\tau}) - (N(T) - \hat{\nu})\}/(N(T) - \hat{\nu})^{1/2} \to_d N(0,1). \tag{2.2}$$

REMARK. The consistency referred to is weak consistency; clearly strong consistency cannot be meaningfully defined in the present context. Confidence intervals for λ_1 and λ_2 may be obtained from (2.1) and (2.2), while confidence intervals for θ may be obtained by inverting a significance test given in the next section.

PROOF. (i) We first show that $\hat{\theta}$ is consistent. Let $Z(t) = T^{-1/2}Y(t/T;0,1)$ and $g(t) = E[Z(t)]$. Then it is readily shown that $g(t)$ is increasing in $[aT, \tau]$ and decreasing on $[\tau, bT]$, and also that for $\varepsilon > 0$ sufficiently small

$$g(t) - g(\tau \pm \varepsilon T) \ge c_1 T^{1/2},$$

where c_1 is a constant which depends only on ε, a, b, λ_1, and λ_2. Thus

$$P[|\hat{\theta} - \theta| > \varepsilon] \le P[|g(\hat{\tau}) - g(\tau)| > c_1 T^{1/2}]$$
$$\le P\left[\sup_{t\in C}|Z(t) - g(t)| > \tfrac{1}{2}c_1 T^{1/2}\right],$$

where $C = [aT, bT]$, using the triangle inequality,

$$\le P\left[\sup_{t\in C}|X_t| + b|X_T| > c_2 T\right],$$

where $X_t = N(t) - E[N(t)]$ and $c_2 = (ab)^{1/2}c_1$, again using the triangle inequality,

$$\le c_3 E\left[\sup_{t\in C}|X_t|\right]^2 T^{-2} + c_4 E\left[X_T^2\right]T^{-2}, \tag{2.3}$$

where $c_3 = 2c_2^{-2}$ and $c_4 = b^2 c_3$, by Chebyshev's inequality and the fact that $(x + y)^2 \le 2(x^2 + y^2)$. $\{X_t\}$ is a square integrable martingale and so $\{|X_t|\}$ is a square integrable submartingale by Jensen's inequality. Thus by Kolmogorov's inequality for submartingales (Liptser and Shiryayev (1977), Theorem 3.2) the first term on the right of (2.3) is $O(T^{-1})$, as is clearly true for the second term. Thus $\hat{\theta}$ is consistent.

We now show the consistency of $\hat{\lambda}_1$. For $\varepsilon > 0$ and $0 < \delta < b - a$, we have

$$P[|\hat{\lambda}_1 - \lambda_1| > \varepsilon] \le P[|\hat{\lambda}_1 - \lambda_1| > \varepsilon, |\hat{\theta} - \theta| < \delta] + P[|\hat{\theta} - \theta| > \delta]$$
$$\le P\Big[\sup_{t \in F} |N(t)/t - \lambda_1| > \varepsilon\Big] + P[|\hat{\theta} - \theta| > \delta],$$

where $F = \{t\colon |t - \tau| < \delta T\}$,

$$\le P\Big[\sup_{t \in F} |X(t)/t| > c_5\Big] + P[|\hat{\theta} - \theta| > \delta],$$

where $c_5 = \varepsilon - \delta(\lambda_1 - \lambda_2)/(\theta - a)$,

$$\le E\big[X(\tau - \delta T)^2\big]/c_5^2(\tau + \delta T)^2 + P[|\hat{\theta} - \theta| > \delta],$$

by Kolmogorov's inequality for submartingales. The first term is clearly $O(T^{-1})$, and we have shown the second term to be $o(1)$, so that $\hat{\lambda}_1$ is consistent. The proof of the consistency of $\hat{\lambda}_2$ is similar.

(ii) Let $\eta_j = t_j - t_{j-1}$, where $t_0 = 0$ and $t_{n+1} = T$. Let $S_m = \Sigma_{j+1}^m(\eta_j - E(\eta_j))$. Clearly, for $j \le \nu$, $\eta_j \sim_{\text{iid}} e(\lambda_1)$, where $\nu = N(\tau)$. We have

$$\frac{\lambda_1 S_{\hat{\nu}}}{\hat{\nu}^{1/2}} = \frac{\lambda_1 S_\nu}{\nu^{1/2}}\left(\frac{\nu}{\hat{\nu}}\right)^{1/2} + \frac{\lambda_1(S_{\hat{\nu}} - S_\nu)}{\hat{\nu}^{1/2}}. \tag{2.4}$$

It is shown in Akman (1985, Chapter 5) that $(S_{\hat{\nu}} - S_\nu)/\hat{\nu}^{1/2} \to_P 0$. Also, $(\nu/\hat{\nu}) \to_P 1$ by the consistency of $\hat{\lambda}_1$ and $\hat{\theta}$, and the renewal theorem. Also, by the renewal theorem, $\nu/T \to \lambda_1\theta$, so that $\lambda_1 S_\nu/\nu^{1/2} \to_d N(0,1)$, by Theorem 17.1 of Billingsley (1968). Thus, by (2.4), $\lambda_1 S_{\hat{\nu}}/\hat{\nu}^{1/2} \to_d N(0,1)$, and (2.1) follows. The proof of (2.2) is similar. This completes the proof of Theorem 1. □

3. Inference about the change-point.

Inference about the change-point will be based on the quantity

$$D(c, d; a, b) = \sup_{s \in C'} |Y(s; c, d)/(\Pi(d) - \Pi(c))^{1/2}|,$$

where $C' = [c + a(d - c), c + b(d - c)]$. To test the null hypothesis that no change occurred we use the test statistic $\Delta = D(0, 1; a, b)$. It follows from the results of Anderson and Darling (1952) that, under the null hypothesis, $P[\Delta > c | \Pi(1) = R]$ converges to $P[\sup_{0 \le t \le \alpha} |U(t)| > c]$ as $R \to \infty$, where $\{U(t)\}$ is the standard Ornstein–Uhlenbeck process and $\alpha = \frac{1}{2}\log\{b(1 - a)/a(1 - b)\}$. Kendall and Kendall (1980) also used this test statistic in the present context, but they found critical values of the test by simulation, also noting an approximation due to Mandl. Here we use the following, simpler approximation due to Dirkse (1975),

$$P[\Delta > c] \sim (2/\pi)^{1/2}\exp(-c^2/2)(\alpha c - \alpha c^{-1} + c^{-1}),$$

when c is large.

We now derive a test of H_0: $\theta = \theta_0$ against H_1: $\theta \neq \theta_0$. Now, $\Delta_1(\theta_0) = D(0, \theta_0; a_1, b_1)$ is the statistic used for testing for a change-point in the interval $[0, \theta_0 T]$, while $\Delta_2(\theta_0) = D(\theta_0, 1; a_2, b_2)$ is the corresponding statistic for the interval $[\theta_0 T, T]$. Clearly, therefore, under H_0, $\Delta_1(\theta_0)$ and $\Delta_2(\theta_0)$ are independent random variables with the same asymptotic distribution as that of Δ under the no-change hypothesis. Our test statistic is $G(\theta_0) = \max\{\Delta_1(\theta_0), \Delta_2(\theta_0)\}$, for which approximate critical values may be found from the relation

$$P[G(\theta_0) > c] \sim 1 - \prod_{i=1}^{2} \left\{1 - (2/\pi)^{1/2} \exp(-c^2/2)\left(\alpha_i c - \alpha_i c^{-1} + c^{-1}\right)\right\},$$

where

$$\alpha_i = \tfrac{1}{2}\log\{b_i(1 - a_i)/a_i(1 - b_i)\}, \qquad i = 1, 2.$$

Clearly, one-tailed tests result from taking as test statistic $\Delta_1(\theta_0)$ when H_1: $\theta < \theta_0$, and $\Delta_2(\theta_0)$ when H_1: $\theta > \theta_0$.

The test based on $G(\theta_0)$ may be inverted, as discussed, for example, by Bickel and Doksum ((1977), Section 5.3), to yield a confidence set for θ. Although this need not be an interval, it often is (see, e.g., Section 5). If it is not, the frequent occurrence of spikes in $G(\theta)$ as a function of θ suggests adopting the slightly conservative solution of using as confidence interval the smallest interval which contains the confidence set.

4. Small-sample results. In order to evaluate the small-sample performance of the procedures proposed in Sections 2 and 3, a Monte-Carlo experiment was performed. To check the size of the test for change based on Δ, 1,000 realisations of a homogeneous Poisson process with $\lambda(t) = 1$ were generated for each of $T = 50, 100, 200$. The proportions of rejections at the 95% level with $a = 0.01$ and $b = 0.99$ were 0.044, 0.052, and 0.035, respectively, which constitutes fairly good agreement with the theoretical size.

In order to evaluate the power of the test for change, and the point and interval estimators of τ, 1,000 realisations of each of four different change-point Poisson processes were generated, with the results shown in Table 1. The variance and bias of $\hat{\tau}$ decreased rapidly with the expected number of events.

TABLE 1

Results from 1,000 *realisations of four different change-point Poisson processes with* $\lambda_1/\lambda_2 = 3$. *All results are at the* 95% *level with* $a = 0.01$ *and* $b = 0.99$.

θ	**0.4375**	**0.25**	**0.5**	**0.5**
$E[N(T)]$	50	50	100	200
Empirical $E(\hat{\theta})$	0.390	0.237	0.471	0.486
Empirical s.d. $(\hat{\theta})$	0.122	0.117	0.069	0.037
Proportion of times θ was in the confidence set	0.929	0.934	0.940	0.944
Power of the test for change	0.743	0.817	0.973	1.000

The test for change was powerful against the alternatives considered, even when the expected number of events was quite small.

5. An illustrative example. We now apply the procedures proposed in Sections 2 and 3 to the data set consisting of intervals between coal-mining disasters during the period 1851–1962 given by Jarrett (1979), who corrected and extended the data set given by Maguire, Pearson, and Wynn (1952). Previous authors, including Barnard (1953), Cox and Lewis (1966), Jarrett (1979), and Berman (1981), have fitted smoothly decreasing log-linear rates of occurrence to the data. However, both the plot of the cumulative number of disasters against time in Jarrett (1979, Figure 1), and the histogram with the smallest risk function estimate given by Rudemo (1982, Figure 12A) suggest that a change-point model may also be appropriate.

The estimators of Section 2 give $\hat{\tau}$ = March 10, 1890, $\hat{\lambda}_1 = 3.21$, and $\hat{\lambda}_2 = 0.92$ disasters per year, with 95% confidence intervals [2.64, 3.77] and [0.70, 1.14] for λ_1 and λ_2, respectively. The test for change yields $\Delta = 8.78$ with $a = 0.01$, $b = 0.99$, so that the null hypothesis of no change is strongly rejected. Figure 1 shows the graph of $G(\theta)$; the resulting 95% confidence set for τ, which is, in fact, an interval, is [October 6, 1886, December 17, 1898].

It may be of interest to compare these results with those of the Bayesian analysis developed by Raftery and Akman (1986). This yielded very high

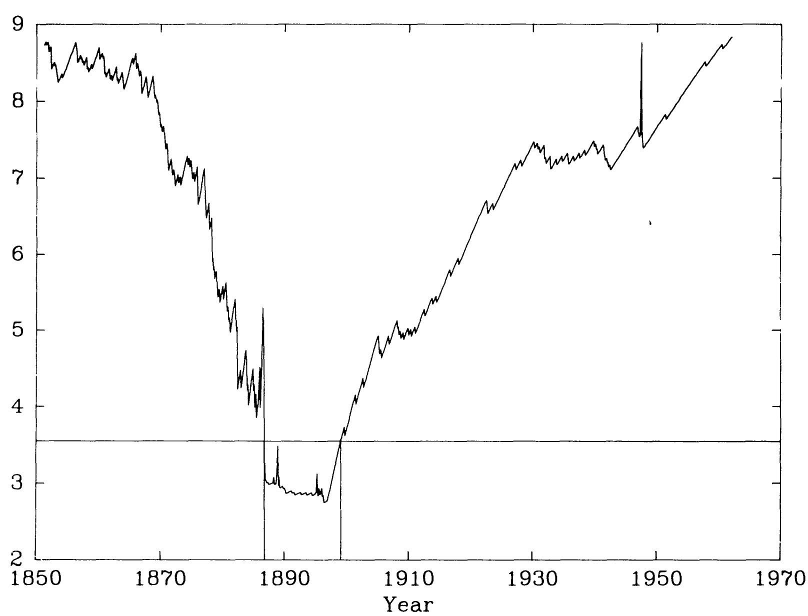

FIG. 1. *Plot of $G(\theta_0)$. The horizontal line indicates the critical value for the test of $\theta = \theta_0$ at the 95% level; the vertical lines delimit the 95% confidence interval for τ.*

posterior odds for a change, with the posterior mode for τ at March 10, 1890, and the 95% Bayesian estimation interval for τ [May 15, 1887, August 3, 1895]. The two analyses give similar results, although the Bayesian estimation interval for τ is somewhat shorter than the confidence interval derived here.

Acknowledgments. We are most grateful to P. Deheuvels, A. F. M. Smith, R. L. Smith, M. Stuart, the Editor (M. Perlman), an Associate Editor, and an anonymous referee for very helpful comments. Section 2 was based on Chapter 5 of V. E. Akman's Ph.D. thesis written at Trinity College, Dublin under the supervision of A. E. Raftery.

REFERENCES

AKMAN, V. E. (1985). Some aspects of statistical inference for stochastic point processes. Ph.D. thesis, Trinity College, Dublin.

ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–212.

BARNARD, G. A. (1953). Time intervals between accidents—a note on Maguire, Pearson and Wynn's paper. *Biometrika* **40** 212–213.

BERMAN, M. (1981). Inhomogeneous and modulated gamma processes. *Biometrika* **68** 143–152.

BICKEL, P. J. and DOKSUM, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics.* Holden-Day, San Francisco.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

CHERNOFF, H. and RUBIN, H. (1956). The estimation of the location of a discontinuity in density. *Proc. 3rd Berkeley Symp. Math. Statist. Probab.* **1** 19–37. Univ. California Press.

COMMENGES, D. and SEAL, J. (1985). The analysis of neuronal discharge sequences: Change-point estimation and comparison of variances. *Statist. Medicine* **4** 91–104.

COX, D. R. and LEWIS, P. A. W. (1966). *The Statistical Analysis of Series of Events.* Methuen, London.

DESHAYES, J. (1984). Ruptures de modèles pour des processus de Poisson. *Ann. Sci. Univ. Clermont-Ferrand II* **78** 1–7.

DIRKSE, J. P. (1975). An absorption probability for the Ornstein–Uhlenbeck process. *J. Appl. Probab.* **12** 595–599.

HINKLEY, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika* **57** 1–17.

JARRETT, R. G. (1979). A note on the intervals between coal-mining disasters. *Biometrika* **66** 191–193.

KALBFLEISCH, J. D. and STRUTHERS, C. A. (1982). An analysis of the Reynolds Metal Company data. *Canad. J. Statist.* **10** 240–247.

KENDALL, D. G. and KENDALL, W. S. (1980). Alignments in two-dimensional random sets of points. *Adv. Appl. Probab.* **12** 380–424.

KUTOYANTS, Y. A. (1984). *Parameter Estimation for Stochastic Processes.* Heldermann, Berlin.

LEONARD, T. (1978). Density estimation, stochastic processes and prior information (with discussion). *J. Roy. Statist. Soc. Ser. B* **40** 113–146.

LIPTSER, R. S. and SHIRYAYEV, A. N. (1977). *Statistics of Random Process* **1**. Springer, New York.

MAGUIRE, B. A., PEARSON, E. S. and WYNN, A. H. A. (1952). The time intervals between industrial accidents. *Biometrika* **39** 168–180.

MANDL, P. (1962). Rozložení doby prvého překročení jedné hranice procesem uhlenbecka. *Aplikace Matematiky* **7** 141–148.

MATTHEWS, D. E. and FAREWELL, V. T. (1982). On testing for a constant hazard against a change-point alternative. *Biometrics* **38** 463–468.

NGUYEN, H. T., ROGERS, G. S. and WALKER, E. A. (1984). Estimation in change-point hazard rate models. *Biometrika* **71** 299–304.

RAFTERY, A. E. and AKMAN, V. E. (1986). Bayesian analysis of a Poisson process with a change-point. *Biometrika* **73** 85–89.

RUBIN, H. (1961). The estimation of discontinuities in multivariate densities, and related problems in stochastic process. *Proc. 4th Berkeley Symp. Math. Statist. Probab.* **1** 563–574. Univ. California Press.

RUDEMO, M. (1982). Empirical choice of histograms and kernel density estimators. *Scand. J. Statist.* **9** 65–78.

SCHULZE, U. (1984). A method of estimation of change-points in multiphasic growth models. *Biometrical J.* **5** 495–504.

DEPARTMENT OF STRATEGIC PLANNING
LISA
ZINCIRLIKUYU
ISTANBUL, TURKEY

DEPARTMENT OF STATISTICS GN-22
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195