Tests for an Increasing Trend in the Intensity of a Poisson Process: A Power Study

LEE J. BAIN, MAX ENGELHARDT, and F. T. WRIGHT*

This article concerns a comparison of several tests for testing the hypothesis of a constant intensity against the alternative of an increasing intensity function in a nonhomogeneous Poisson process (NHPP). The study includes the well-known Laplace test statistic, the most powerful test for the shape parameter in a Poisson process with Weibull intensity, the likelihood ratio test against arbitrary NHPP alternatives, two nonparametric tests for trends based on Kendall's tau and Spearman's rho, and a test based on an F statistic. The powers of the tests are determined by Monte Carlo simulation against alternatives that are increasing at an exponential rate, a power rate (Weibull intensity), and a logarithmic rate. Alternatives that are step functions with one jump are also considered. In a few cases, the exact powers are also obtained analytically.

KEY WORDS: Poisson processes; Increasing intensities; Weibull Poisson processes; Laplace test; Likelihood ratio tests.

1. INTRODUCTION

Nonhomogeneous Poisson processes (NHPP's) provide models for a variety of physical phenomena. For instance, if at each failure a system is repaired to its condition at the time of failure and placed in service again, then the failures are often modeled by a NHPP, provided the repair times can be neglected. In some of these situations, it may be reasonable to assume that the intensity, $\lambda(\cdot)$, is nondecreasing, so tests of H_0: $\lambda(\cdot)$ is constant versus H_1: $\lambda(\cdot)$ is increasing are of interest. The results of such tests could indicate whether the simple homogeneous Poisson process (HPP) may be adequate or whether a more general NHPP model is required.

Suppose that a NHPP is observed for T^* units of time and that the number of failures, which is a random variable, is denoted by N. The tests considered are conditional on $N > 0$, and for an observed number of failures, $n > 0$, the failure times are denoted by $0 < T_1 < T_2 < \cdots < T_n < T^*$. We compare several of the tests that are available for this time-truncated framework. One of the earliest is attributed to Laplace and is based on the statistic $L = \sum_{i=1}^n T_i/T^*$. Under H_0 and conditional on $N = n$, the random variables T_i/T^* are distributed as the order statistics from a uniform distribution on (0, 1), so L is distributed as the sum of n uniform random variables. H_0 is rejected for large values of L, and an exact expression for $P[L \geq l]$ can be obtained from equation (19) of Johnson and Kotz (1970, p. 64). Since the distribution of the sum of independent uniform random variables approaches normality quite rapidly as n increases, an approximate α-level test is obtained, for moderate or large n, by rejecting H_0 if $L \geq n/2 + z_{1-\alpha}(n/12)^{1/2}$, where $z_{1-\alpha}$ is the $1-\alpha$th quantile of the standard normal distribution. Cox (1955) discussed the use of this test for testing $\beta = 0$ versus $\beta > 0$ in $\lambda(t) = \alpha e^{\beta t}$. Applying theorem 3 of Lehmann (1959, p. 136), this test can be shown to be uniformly most powerful unbiased (UMPU) in this parametric setting. Bartholomew (1956) gave an expression for its power (also see Bates 1955) and showed that it compares favorably with the one-sided Kolmogorov–Smirnov test. Ascher and Feingold (1978) further discussed its use in the study of repairable systems.

Another family of intensity functions, which is quite flexible, is $\lambda(t) = (\beta/\theta)(t/\theta)^{\beta-1}$ for $\beta, \theta > 0$. Because this is the failure rate for the Weibull distribution, the corresponding process has been called the Weibull Poisson Process (WPP). Inferences for a WPP are discussed in Crow (1974, 1982), Saw (1975), Finkelstein (1976), Lee and Lee (1978), Engelhardt and Bain (1978), and Bain and Engelhardt (1980). Crow (1974) gave tests for β with θ a nuisance parameter. In testing $\beta = 1$ versus $\beta > 1$, which is equivalent to testing H_0 versus H_1, $\beta = 1$ is rejected for small values of $Z = 2\sum_{i=1}^n \log(T^*/T_i)$, which conditional on $N = n$, has a chi-squared distribution with $2n$ degrees of freedom under H_0. The test based on Z is UMPU in this WPP setting.

Boswell (1966) developed the likelihood ratio test (LRT), conditional on n failures, for H_0 versus H_1 for an arbitrary NHPP. The maximum likelihood estimator of $\lambda(\cdot)$ under $H_0 \cup H_1$ is shown to be zero on $[0, T_1)$ and constant on $[T_k, T_{k+1})$ for $k = 1, 2, \ldots, n$ with $T_{n+1} = T^*$ and

$$\hat{\lambda}(T_k) = \max_{1\leq\gamma\leq k}\min_{k\leq\delta\leq n}(\delta - \gamma + 1)/(T_{\delta+1} - T_\gamma).$$

A computation algorithm is also given there (compare p. 1567). The LRT rejects for large values of

$$W = 2\left\{\sum_{k=1}^n \log(\hat{\lambda}(T_k)) + n\log(T^*/n)\right\},$$

and letting $\chi^2(k)$ denote a chi-squared variable with k degrees of freedom,

$$P[W \geq w] \approx \sum_{k=1}^n P(k, n)P[\chi^2(k+1) \geq w]$$

for large n, where the $P(k, n)$ are related to the Stirling numbers of the first kind. The $P(k, n)$ are discussed in detail in chapter 3 of Barlow et al. (1972), and they are tabulated in their table A.5.

If the intensity is increasing, the interfailure times $T_k - T_{k-1}$, $k = 1, 2, \ldots, n$ $(T_0 \equiv 0)$, should tend to decrease. Hence nonparametric tests for trends using either Kendall's tau or

* Lee J. Bain, Max Engelhardt, and F. T. Wright are Professors, Department of Mathematics and Statistics, University of Missouri, Rolla, MO 65401. The research for this article was partially supported by Office of Naval Research Contract N00014-80-C-0322 and by Air Force Office of Scientific Research Grant AFOSR-84-0164. The authors thank Tuan Tran for writing the computer program for the Monte Carlo Study. They also thank the referees and the associate editor for their helpful comments.

Journal of the American Statistical Association
June 1985, Vol. 80, No. 390, Theory and Methods

Spearman's rho could be considered for testing H_0 versus H_1. (For a discussion of these tests as tests of trend, see Hollander and Wolfe 1973, p. 190.)

Barlow et al. (1972, p. 197) observed that the failure times could be divided into the two parts, the first d and the last $n - d$, and the ratio $F = (n - d)T_d/(d(T_n - T_d))$ used as a test statistic for trend. Of course, an increasing intensity should correspond to a larger value of this ratio. If the intensity is constant, then conditional on n failures, T_k can be expressed as $\sum_{j=1}^{k} Y_j / \sum_{j=1}^{n+1} Y_j$, where the Y_j are independent exponential variables with a common mean. So F can be written as $\sum_{j=1}^{d} Y_j / \sum_{j=d+1}^{n} Y_j$, and hence F follows Snedecor's F distribution with $2d$ and $2(n - d)$ degrees of freedom. We only consider the test with $d = [n/2]$.

Section 2 contains the results of a study that compares the powers of the tests just described, and in Section 3 our recommendations based on this study are given.

2. POWER COMPARISONS

Some of the tests discussed in the last section are optimal for certain families of intensities, and it is of interest to study their power functions over a broad range of intensities. Some of the other tests are designed to be more omnibus, and information concerning the loss of power incurred when using one of these, in a particular situation, instead of a test that is optimal in that situation, would be useful.

Because of the complex nature of some of the power functions involved, a Monte Carlo study was conducted. The tests based on L, Z, W, F, Kendall's tau (which we denote by K), and Spearman's rho (which we denote by S) are compared. Intensity functions that are exponential, of the Weibull type, logarithmic, and nondecreasing step functions with one jump are considered.

A NHPP with intensity $\lambda(\cdot)$ can be expressed as a function of a HPP. In particular, if $\Lambda(t) = \int_0^t \Lambda(s)\,ds$, the expected number of failures in $[0, t]$, is strictly increasing, and if $\{S_k\}$ denotes the failure times for a HPP with $\lambda = 1$, then $\{\Lambda^{-1}(S_k)\}$ are the failure times for a NHPP with intensity $\lambda(\cdot)$. If $\lambda(\cdot)$ is an intensity function with corresponding mean function $\Lambda(\cdot)$, and $\lambda_\theta(t) = \lambda(t/\theta)/\theta$, then $\Lambda_\theta(t) = \Lambda(t/\theta)$. Hence the failure times for the process with intensity $\lambda_\theta(\cdot)$ can be expressed as $\theta\Lambda^{-1}(S_k)$. All of the statistics considered are invariant under time rescalings ($T_i \to T_i/\theta$ and $T^* \to T^*/\theta$, some $\theta > 0$), and their conditional distributions given $N = n$ do not depend on $E(N) = \Lambda(T^*)$. Hence the conditional powers for a particular alternative rate $\lambda(\cdot)$ and observation time T^* are equal to the conditional powers under rate $\lambda_{\beta,\theta}(t) = \beta\lambda(t/\theta)$ and observation time T^*/θ, for any $\beta > 0$ and $\theta > 0$.

Since it is possible that the choice of test that is to be used may depend on the number of failures observed, we study the powers of the tests conditional on $N = n$, and in particular we consider $n = 10$, 20, and 40. For a NHPP, given that there are n failures in $[0, T^*]$, the failure times are distributed as the order statistics from a random sample of size n from the density $f(t) = \lambda(t)/\Lambda(T^*)$, $0 < t < T^*$, and $f(t) = 0$, otherwise. If $\lambda_\eta(T) = \eta\lambda(t)$ for some fixed $\lambda(\cdot)$, then it is clear from the last statement that N is sufficient for η, so we need only consider one value of η, say $\eta = 1$. In the Monte Carlo study, the order statistics, $U_{(1)} < U_{(2)} < \cdots < U_{(n)}$, from a random sample of size n from the $U(0, 1)$ distribution are generated and T_k is obtained by solving

$$\Lambda(T_k)/\Lambda(T^*) = U_{(k)}. \tag{1}$$

In the WPP setting, $\lambda(t) = \beta t^{\beta-1}$ is considered for $\beta = 1$, 2, 4. Since all of the tests are functions of T_k/T^* and $T_k/T^* = (U_{(k)})^{1/\beta}$ in the case, we need only consider one value of T^*, say $T^* = 1$. In the exponential case, we consider $\lambda(t) = e^t$ with $T^* = 1$, 2, 3, and in the logarithmic case we consider $\lambda(t) = \log(t + 1)$ with $T^* = 10$, 15, 25. Intensities of the form $\lambda(t) = 1$, $0 \le t \le \tau\cdot T^*$, and $\lambda(t) = 2$, $\tau\cdot T^* < t < T^*$ are also considered with $\tau = \frac{1}{3}, \frac{1}{2}, \frac{2}{3}$. Again it is easily shown that the distribution of T_k/T^*, given $N = n$, does not depend on T^*, so we set $T^* = 1$.

Tables 1–4 contain Monte Carlo estimates of the powers for the tests and alternatives previously described. These estimates are based on 5,000 replications. Tests with a nominal significance level of $\alpha = .05$ are considered; however, Boswell and Brunk (1969) observed empirically that the large sample approximation for the LRT yields a true significance level somewhat larger than the nominal level. By experimenting, we found that a nominal level of .04 yields an estimated level of approximately .05 for the LRT for the range of sample sizes being considered. Clearly, this is a disadvantage of the LRT in this situation. As can be seen from the tables, the significance levels for the other tests are very close to .05, with the exception of Kendall's tau at $n = 10$. In this case the largest significance level less than or equal to .05 is .036, without using random-

Table 1. Estimated Powers for Testing H_0 Versus H_1: Weibull Intensity, $\lambda(t) = \beta t^{\beta-1}$

	β								
	1.0			2.0			4.0		
Test	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$
Z (exact)	.0500	.0500	.0500	.6431	.9185	.9978	.9982	.9999	.9999
Z (estimate)	.050	.051	.050	.648	.920	.999	.998	1.000	1.000
W*	.050	.050	.050	.509	.816	.986	.986	1.000	1.000
L (approximate)	.0500	.0500	.0500	.5876	.8744	.993	.9980	1.000	1.000
L (estimate)	.050	.049	.053	.602	.879	.993	.996	1.000	1.000
S	.053	.050	.053	.340	.558	.779	.553	.796	.958
K	.036	.050	.049	.273	.561	.772	.495	.796	.957
F	.052	.049	.052	.418	.678	.921	.935	.999	1.000

* The normal level for W is .04.

Table 2. Estimated Powers for Testing H_0 Versus H_1: Exponential Intensity, $\lambda(t) = e^t$

	T^*								
	1.0			2.0			4.0		
Test	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$
Z	.210	.321	.498	.503	.738	.932	.777	.949	.998
W*	.185	.278	.437	.463	.714	.923	.741	.942	.999
L	.224	.366	.575	.547	.816	.972	.832	.978	1.000
S	.161	.267	.433	.338	.596	.861	.501	.793	.994
K	.119	.259	.424	.267	.591	.856	.417	.790	.991
F	.174	.270	.456	.409	.668	.916	.667	.920	.996

* The nominal level for W is .04.

ization. With a nominal level of .05, the IMSL subroutine that conducts the test based on Kendall's tau uses the critical value for $n = 10$, which gives $\alpha = .036$. Of course, the Z and F tests are exact. The normal approximation is used for the L test, but as we have seen, it can be made exact also.

To give some idea of the accuracy in these estimated powers, we consider the power of Z for $\beta > 1$. Conditional on $N = n$, we see from (1) that T_k/T^* is distributed as $(U_{(k)})^{1/\beta}$, so βZ has a conditional distribution that is chi-squared with $2n$ degrees of freedom. The conditional power of Z is $P[\chi^2(2n) \leq \beta\chi^2_\alpha(2n)]$, where $\chi^2_\alpha(2n)$ is the αth quantile of a chi-squared distribution with $2n$ degrees of freedom. The first row in Table 1 gives these exact powers. The largest discrepancy between the exact and estimated powers is .005. Similar comparisons can be made for the L test, since it is based on approximate normal distributions. Since, conditional on $N = n$, T^*L can be expressed as a sum of the unordered observations from the density $f(t)$ given earlier, the conditional mean and variance are given by, respectively,

$$\mu_n = E(L \mid N = n) = n \int_0^{T^*} t\lambda(t)\, dt/(T^*\Lambda(T^*))$$

and

$$\sigma_n^2 = V(L \mid N = n) = n \int_0^{T^*} t^2\lambda(t)\, dt/(T^{*2}\Lambda(T^*)) - \mu_n^2/n.$$

Hence the conditional power of L is approximately $1 - \Phi((n/2 - \mu_n + z_{1-\alpha}(n/12)^{1/2})/\sigma_n)$. Note that in the WPP case, the formulas for μ_n and σ_n^2 simplify to

$$\mu_n = n\beta/(\beta + 1)$$

and

$$\sigma_n^2 = n[\beta/(\beta + 2) - (\beta/(\beta + 1))^2].$$

The fourth row in Table 1, which is labeled L (approximate), gives the approximate powers computed using these formulas, and the fifth row [labeled L (estimate)] gives the Monte Carlo estimates of these powers. The largest discrepancy between the approximate and Monte Carlo values occurs for $n = 10$ and $\beta = 2$, and the difference in this case seems to be due primarily to the normal approximation. One could obtain a further check on accuracy. Bartholomew (1956) gave the power of L for exponential intensities, conditional on n failures, but due to the complex nature of this power function, these exact values have not been computed.

The Z test has the largest estimated powers in each case considered with logarithmic (Table 3) or Weibull (Table 1) alternatives. Of course, the L test is most powerful for the cases with exponential intensities (Table 2). For the step function alternatives (Table 4), Z has the largest estimated power if the single jump occurs at $\tau = \frac{1}{3}$, but if $\tau = \frac{1}{2}$ or $\frac{2}{3}$, L is preferred. (For $\tau = \frac{2}{3}$ and $n = 20$ and 40, the F test has larger estimated power than L, but the differences are small enough to be due to Monte Carlo error, and L clearly outperforms F for $\tau = \frac{1}{2}$.)

For the situations considered here, the nonparametric tests cannot be recommended. While the test based on Spearman's rho would seem to be preferred over that based on Kendall's tau, the power of S drops to 53%–55% of the largest power in some cases (i.e., WPP with $\beta = 2$ and $n = 10$, WPP with $\beta = 4$ and $n = 10$, the logarithmic intensity with $T^* = 10$, 15, 25 and $n = 40$, and the step function intensity with $\tau = \frac{1}{3}$ and $n = 40$. If the distribution assumption on inter-failure times is

Table 3. Estimated Powers for Testing H_0 Versus H_1: Logarithmic Intensity, $\lambda(t) = \log(t + 1)$

	T^*								
	10			15			25		
Test	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$
Z	.322	.590	.876	.282	,519	.819	.240	.443	.733
W*	.233	.422	.726	.207	.367	.649	.179	.306	.567
L	.285	.488	.762	.250	.430	.695	.218	.321	.604
S	.203	.322	.480	.187	.288	.439	.171	.260	.393
K	.150	.312	.461	.136	.284	.426	.124	.257	.382
F	.196	.314	.524	.177	.276	.467	.157	.237	.397

* The nominal level for W is .04.

Table 4. Estimated Powers for Testing H_0 Versus H_1: Step Function Intensity, $\lambda(t) = 1, 0 \le t \le \tau, \lambda(t) = 2, \tau < t \le 1$

	τ								
	1/3			1/2			2/3		
Test	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$
Z	.185	.288	.456	.217	.326	.493	.198	.277	.417
W*	.145	.207	.335	.177	.262	.426	.189	.278	.445
L	.158	.259	.431	.227	.372	.588	.241	.379	.585
S	.115	.167	.244	.149	.243	.397	.170	.289	.476
K	.083	.166	.239	.111	.235	.377	.119	.271	.459
F	.125	.172	.266	.201	.313	.531	.235	.389	.589

* The nominal level for W is .04.

in question, however, one might wish to consider S. Overall, the F test outperforms the nonparametric tests, but for the logarithmic intensity, with each choice of $T^*_\cdot$ and $n = 20$, its power is only 53% of that for Z. If one wishes to guard against slowly increasing intensities, the F test should not be used.

For each case considered, the power of Z is at least 71% of the largest estimated power in that case, and this minimum relative power of 71% occurs for the step function intensity with $\tau = \frac{2}{3}$ and $n = 20$ and 40. The minimum relative power of W is 69%, and for L it is 82%. The LRT is designed to discriminate against all nondecreasing intensities that are not constant, and if one were concerned with very nonregular intensities, W could be considered. For the alternatives considered here, however, either Z or L would be preferred.

If one wishes to discriminate over the broad range of alternatives considered here, L is recommended. However, if attention is restricted to the "smooth" intensities (i.e., Weibull, exponential, or logarithmic), then the minimum relative power of Z is 87%. This is about 5% above that of L, so Z is recommended for such alternatives.

For some purposes one may be interested in the unconditional powers of these tests. They may be obtained by using the law of total probability, the fact that N has a Poisson distribution with mean $\Lambda(T^*)$ and the formulas for conditional powers, when the latter are available. We obtained Monte Carlo estimates of the unconditional powers for the six tests and some of the alternatives considered here. The study of these estimates of unconditional powers lead to conclusions like those based on the conditional powers.

3. CONCLUSIONS

In testing for an increasing intensity in a Poisson process, the Z test performs quite well for the smooth alternatives studied here—that is, for logarithmic, Weibull, and exponential intensities. In fact, for such alternatives its estimated power is at least 87% as large as that of the other five tests considered, and for the logarithmic and Weibull intensities it is the most powerful of the six. However, if one also wishes to guard against the step function intensities considered, L could be used. For all of the cases considered, the estimated power of L is at least 82% of that of the other tests, and for one of the step function intensities, the estimated power of Z is 71% of that for L.

[*Received April 1983. Revised October 1984.*]

REFERENCES

Ascher, H. E., and Feingold, H. (1978), "Application of Laplace's Test to Repairable System Reliability," in *Proceedings of the International Conference on Reliability and Maintainability* (Société Pour la Diffusion des Sciences et des Arts, France), 219–225.

Bain, L. J., and Engelhardt, M. E. (1980), "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process," *Technometrics,* 22, 421–426.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk, H. D. (1972), *Statistical Inferences Under Order Restrictions,* New York: John Wiley.

Bartholomew, D. J. (1956), "Tests for Randomness in a Series of Events When the Alternative is a Trend," *Journal of the Royal Statistical Society,* Ser. B, 18, 234–239.

Bates, G. E. (1955), "Joint Distributions of Time Intervals for the Occurrence of Successive Accidents in a Generalized Polya Scheme," *Annals of Mathematical Statistics,* 26, 705–720.

Boswell, M. T. (1966), "Estimating and Testing Trend in a Stochastic Process of Poisson Type," *Annals of Mathematical Statistics,* 37, 1564–1573.

Boswell, M. T., and Brunk, H. D. (1969), "Distribution of Likelihood Ratio in Testing Against Trend," *Annals of Mathematical Statistics,* 40, 371–380.

Cox, D. R. (1955), "Some Statistical Methods Connected With Series of Events," *Journal of the Royal Statistical Society,* Ser. B, 17, 129–164.

Crow, L. H. (1974), "Reliability and Analysis for Complex Repairable Systems," in *Reliability and Biometry,* eds. F. Proschan and R. J. Serfling, Philadelphia: Society for Industrial and Applied Mathematics, pp. 379–410.

——— (1982), "Confidence Interval Procedures for the Weibull Process With Applications to Reliability Growth," *Technometrics,* 24, 67–72.

Engelhardt, M. E., and Bain, L. J. (1978), "Prediction Intervals for the Weibull Process," *Technometrics,* 20, 167–169.

Finkelstein, J. (1976), "Confidence Bounds on the Parameters of the Weibull Process," *Technometrics,* 18, 115–117.

Hollander, M., and Wolfe, D. A. (1973), *Nonparametric Statistical Methods,* New York: John Wiley.

Johnson, N. L., and Kotz, Samuel (1970), *Continuous Univariate Distributions—2,* Boston: Houghton-Mifflin.

Lee, L., and Lee, S. K. (1978), "Some Results on Inference for the Weibull Process," *Technometrics,* 20, 41–45.

Lehmann, E. L. (1959), *Testing Statistical Hypotheses,* New York: John Wiley.

Saw, J. G. (1975), "Tests on the Intensity of a Poisson Process," *Communications in Statistics,* 4, 777–782.