Statistics & Probability Letters 60 (2002) 313–320

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

# Testing the independence of Poisson variates under the Holgate bivariate distribution: the power of a new evidence test

J.M. Stern[a,b,1], S. Zacks[a,*]

[a] Department of Mathematical Sciences, Binghamton University, P.O. Box 6000, Binghamton, NY 13902-6000, USA

[b] Bioinfo, Nopef, and IME-USP, University of Sao Paulo, Brazil

Received May 2002; received in revised form August 2002

**Abstract**

A new Evidence Test is applied to the problem of testing whether two Poisson random variables are dependent. The dependence structure is that of Holgate's bivariate distribution. These bivariate distribution depends on three parameters, $0 < \theta_1, \theta_2 < \infty$, and $0 \leqslant \theta_3 \leqslant \min(\theta_1, \theta_2)$.

The Evidence Test was originally developed as a Bayesian test, but in the present paper it is compared to the best known test of the hypothesis of independence in a frequentist framework. It is shown that the Evidence Test is considerably more powerful when the correlation is not too close to zero, even for small samples.



*MSC:* 62F03; 62F04

*Keywords:* Correlation and dependence; Evidence test; Holgate distribution; Test of hypotheses; Numerical integration and optimization

## 1. Introduction

Multivariate discrete random variables are important in reliability, spatial statistics, applied stochastic processes, and other areas of statistics. A good introduction and motivation for the Holgate and other discrete bivariate processes can be found in Barlow and Prochan (1981, Chapter 5, Multivariate distributions for dependent components), and Kocherlakota and Kocherlakota (1992). For testing whether Poisson random variables are independent, one needs a dependence structure as an

* Corresponding author. Fax: +1-607-777-2450.

*E-mail addresses:* jmstern@hotmail.com (J.M. Stern), shelly@math.binghamton.edu (S. Zacks).

[1] Grant Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP 2001-03484-1.



PII: S0167-7152(02)00314-0

alternative to the null hypothesis. In the present paper we consider the Holgate bivariate distribution of $W=(X,Y)$, see Holgate (1964), Irwin (1963), Karilis and Ntzoufras (1998), McKendrick (1926), and Teicher (1954). In this bivariate distribution, $X$ and $Y$ have marginal Poisson distributions with means $\theta_1$ and $\theta_2$. Let $p(j;\theta)$ denote the pdf of a Poisson distribution with mean $\theta$. The pdf of the Holgate bivariate Poisson distribution can be derived by considering independent Poisson random variables $J_1$, $J_2$, $J_3$ with means, $\lambda_1$, $\lambda_2$, $\lambda_3$. Let $X=J_1+J_3$, $Y=J_2+J_3$ and $Z=J_3$. The marginal bivariate distribution of $W=(X,Y)$ is the Holgate distribution with parameters $\theta_1=\lambda_1+\lambda_3$, $\theta_2=\lambda_2+\lambda_3$, $\theta_3=\lambda_3$; $0<\theta_1<\infty$, $0<\theta_2<\infty$, $0\leqslant\theta_3<\min(\theta_1,\theta_2)$.

Thus, the pdf of the Holgate distribution is

$$f(x,y;\theta_1,\theta_2,\theta_3)=\sum_{l=0}^{\min(x,y)} p(x-l;\theta_1-\theta_3)p(y-l;\theta_2-\theta_3)p(l;\theta_3). \tag{1.1}$$

Notice that $p(l;0)=I\{l=0\}$, where $I\{\ \}$ is the indicator function of the set. Thus, (1.1) yields immediately that $f(x,y;\theta_1,\theta_2,0)=p(x;\theta_1)p(y;\theta_2)$. On the other hand, since the marginal pdf's of $X$ and $Y$ are $p(x;\theta_1)$ and $p(y;\theta_2)$, respectively, writing (1.1) as

$$\begin{aligned} f(x,y;\theta_1,\theta_2,\theta_3) &= \mathrm{e}^{-\theta_3}p(x;\theta_1-\theta_3)p(y;\theta_2-\theta_3) \\ &\quad + I\{\min(x,y)\geqslant 1\}\mathrm{e}^{-(\theta_1+\theta_2-\theta_3)}\sum_{l=0}^{\min(x,y)}\frac{(\theta_1-\theta_3)^{x-l}(\theta_2-\theta_3)^{y-l}\theta_3^l}{(x-l)!(y-l)!l!}, \end{aligned} \tag{1.2}$$

we immediately obtain that $f(x,y;\theta)=p(x;\theta_1)p(y;\theta_2)$ for *all* $x,y\in\{0,1,\ldots\}$ only if $\theta_3=0$. Thus, $X$ and $Y$ are *independent* if and only if $\theta_3=0$.

The problem studied in the present paper is to test the composite hypothesis $\mathrm{H}_0:\theta_3=0$; $\theta_1$, $\theta_2$, arbitrary (positive), against the alternative composite hypothesis, $\mathrm{H}_1:\theta_3>0$; $\theta_1$, $\theta_2$, arbitrary.

Let $\rho$ denote the correlation between $X$ and $Y$. Since $\rho=\theta_3/\sqrt{\theta_1\theta_2}$, a test of independence of $X$ and $Y$ is equivalent to a test whether $\rho=0$ versus $\rho>0$.

Paul and Ho (1989) compared the power of six alternative tests statistics of $\mathrm{H}_0$ against $\mathrm{H}_1$. They established by numerical simulations that a modified-F test (MF) based on the sample correlation, $r_n$, is more powerful than the other tests.

We show that the Evidence Test is considerably more powerful than the MF test of Paul and Ho in certain parts of the parameter space. The general definition of Evidence Against a Hypothesis, and respective test is given in Section 2. This test is based on the developments of Irony et al. (2001), Madruga et al. (2001), Pereira and Stern (1999, 2001) and Stern (2001). The Evidence Test is called in these papers Full Bayesian Significance Test (FBST). We use the acronym FBST to designate this test.

As will be shown, the determination of the FBST is computationally intensive. In particular, one has to determine the $(1-\alpha)$ quantiles of the null distributions of the evidence tests. This problem and some of its computational aspects are discussed in Section 3. Finally, we present in Section 4 the power of the FBST, computed at the same points of the parameter space, as those in Paul and Ho (1989). As shown in Tables 2 and 3 of Section 4, if the sample of bivariate observations is of size 20 or more, the power of the FBST is greater than that of the MF test, whenever the parameter $\theta_3$ is not too close to zero. We see that the difference in power is more than 0.2 around a correlation of $\rho = 0.45$ for a sample of size $n = 20$. Even in small samples, $n = 10$, the difference in power is considerable if the correlation $\rho > 0.45$.

## 2. The evidence test

The FBST was developed to test a precise hypothesis $H_0$ against the negation of $H_0$. The FBST was originally developed as a Bayesian test, but in the present paper it is compared to the best known test of the hypothesis of independence in a frequentist framework. A hypothesis is called precise if the points satisfying $H_0$ belong to a manifold

$$\Theta_0 = \{\theta \in \Theta \mid h(\theta) = 0\}. \tag{2.1}$$

The manifold $\Theta_0$ is specified by a constraint function, $h(\theta)$, so that the dimension of $\Theta_0$ is smaller than that of $\Theta$. In the present problem of testing independence, $\Theta_0 = \{\theta \in \Theta \mid \theta_3 = 0\}$, is a two-dimensional manifold corresponding to a precise hypothesis, $H_0$, in the original three-dimensional parameter space.

Let $X_1, X_2, \ldots$ be i.i.d. random variables (or vectors) having a common density $f(X;\theta)$, with respect to a $\sigma$-finite measure $\mu$. Let $X_1, X_2, \ldots, X_n$ be a random sample, and let $L(\theta; X) = \tilde{p}(T(X); \theta)$ be a version of the likelihood function of $\theta$ on $\Theta$. $X$ is the vector (matrix) of i.i.d. random variables (vectors) $X_1, X_2, \ldots, X_n$, and $T(X)$ is the likelihood statistics. Let $p(\theta)$ be a *prior* density of $\theta$ in $\Theta$, with respect to a $\sigma$-finite measure $\lambda$. The *posterior* density of $\theta$, given $T(X)$, is

$$p(\theta \mid T(X)) = \frac{\tilde{p}(T(x); \theta)\, p(\theta)}{\int_{\Theta} \tilde{p}(T(X); \theta)\, p(\theta)\, \mathrm{d}\lambda(\theta)}. \tag{2.2}$$

Let

$$\theta^* = \arg\max_{\theta \in \Theta_0} p(\theta \mid T(X)), \tag{2.3}$$

$$p^* = p(\theta^* \mid T(X)). \tag{2.4}$$

The *evidence against* $H_0$ is defined as the *credibility of the highest probability density set* (HPDS) on which $p(\theta \mid T(X)) \geqslant p^*$. We denote this evidence by $\mathrm{Ev}(T(X))$. In other words, the evidence against $H_0$ is the posterior probability of the HPDS $\Theta^*$, where

$$\Theta^* = \{\theta \in \Theta \mid p(\theta \mid T(X)) \geqslant p^*\}, \tag{2.5}$$

$$\mathrm{Ev}(T(X)) = \int_{\theta \in \Theta^*} p(\theta \mid T(X))\, \mathrm{d}\lambda(\theta). \tag{2.6}$$

$\mathrm{Ev}(T(X))$ is a measurable function of $T(X)$ that can be used as a test statistic for $H_0$. Common Bayesian tests, see Box and Tiao (1973), Good (1983), Lindley (1978), and Zellner (1971), accept $H_0$ if the HDPS of a given credibility intersects $\Theta_0$. Notice that according to the above definition, the HPDS $\Theta^*$ is *tangential* to $\Theta_0$ at $\theta^*$. Thus, a test based on $\mathrm{Ev}(T(X))$ rejects $H_0$ if $\mathrm{Ev}(T(X) > \xi$ for some $0 < \xi < 1$. We determine $\xi$ from a frequentist point of view. Thus, for $\theta \in \Theta_0$, the probability of type I error is

$$\alpha(\theta; \xi) = \int I\{\mathrm{Ev}(T(X)) > \xi\} \prod_{i=1}^{n} f(X_i; \theta)\, \mathrm{d}\mu(X_i). \tag{2.7}$$

For a given level of significance, $\alpha$, $\alpha(\theta;\xi) \leqslant \alpha$ if $\xi \geqslant q_{\alpha,n}(\theta)$, where $q_{\alpha,n}(\theta)$ is the $(1-\alpha)$ quantile of the distribution of $\mathrm{Ev}(T(X))$ under $\theta$. Accordingly, the test is of size $\alpha$ if $\xi = \xi_\alpha$, where,

$$\xi_\alpha = \sup_{\theta \in \Theta_0} q_{\alpha,n}(\theta). \tag{2.8}$$

The power function of this test, for $\theta \in \Theta \setminus \Theta_0$, is

$$\psi_n(\theta) = \Pr_\theta\{\mathrm{Ev}(T(X)) > \xi_\alpha\} = \int I\{\mathrm{Ev}(T(X)) > \xi_\alpha\} \prod_{i=1}^{n} f(X_i;\theta)\,\mathrm{d}\mu(X_i). \tag{2.9}$$

In order to make the FBST free, as much as possible, of subjective choice of prior distributions, we shall try to determine $p(\theta \mid T(X))$ as a normalized likelihood function. This is possible if we have a normalizing constant

$$c = \int_\Theta \tilde{p}(T(X);\theta)\,\mathrm{d}\lambda(\theta) < \infty. \tag{2.10}$$

If we do not have a finite normalizing constant, we have to choose a "regulating" prior density $p(\theta)$, so that the denominator of (2.2) is finite. In many testing problems a uniformly most powerful test does not exist. If a prudent choice of prior increase the power of the FBST in a certain subset of the parameter space, so much the better.

## 3. The evidence test of independence for the Holgate bivariate Poisson

The composite hypothesis $\mathrm{H}_0{:}\, \theta_3 = 0;\ \theta_1,\theta_2 \geqslant 0$ is a precise two-dimensional hypothesis in the three-dimensional parameter space, $\Theta$. Let $W_i = (X_i, Y_i)$, $i = 1 \ldots n$ be i.i.d. vectors, having a common joint Holgate pdf, as in Section 1. Let $h_i = \min(X_i, Y_i)$. On the parameter space $\Theta$, define the functions

$$L_i(\theta; W_i) = f(W_i \mid \theta), \quad i = 1 \ldots n, \tag{3.1}$$

where the parameter vector is $\theta = (\theta_1, \theta_2, \theta_3)$.

The likelihood function, given $W = (W_1, \ldots, W_n)$, is

$$L(\theta; W) = \prod_{i=1}^{n} L_i(\theta; W_i). \tag{3.2}$$

A recursive form of the likelihood function, useful for large values of $n$, can be found in Kocherlakota and Kocherlakota (1992).

We performed the numerical computation of the evidence by Monte Carlo (MC) simulation, see Liu (2001), for the estimation of the ratio

$$\mathrm{Ev}(W) = \frac{\int_{\Theta^*} L(\theta; W)\,\mathrm{d}\theta}{\int_\Theta L(\theta; W)\,\mathrm{d}\theta}. \tag{3.3}$$

Since the space $\Theta$ is unbounded, we randomly chose the values of $\theta = (\theta_1, \theta_2, \theta_3)$ according to an "importance sampling" density $g(\theta)$, which is positive on $\Theta$. The evidence function (3.3) is equivalent to

$$\mathrm{Ev}(W) = \frac{\int_\Theta Z_g^*(\theta; W) g(\theta)\,\mathrm{d}\theta}{\int_\Theta Z_g(\theta; W) g(\theta)\,\mathrm{d}\theta}, \tag{3.4}$$

$$Z_g(\theta; W) = \frac{L(\theta; W)}{g(\theta)} \tag{3.5}$$

and

$$Z_g^*(\theta; W) = I^*(\theta; W) Z_g(\theta; W), \tag{3.6}$$

$$I^*(\theta; W) = I\{L(\theta; W) \geqslant l^*\}. \tag{3.7}$$

Thus, a Monte Carlo estimate of (3.4) is

$$\hat{\mathrm{E}}\mathrm{v}_{g,m}(W) = \frac{\sum_{i=1}^m Z_g^*(\theta_{i,\bullet}; W)}{\sum_{i=1}^m Z_g(\theta_{i,\bullet}; W)}, \tag{3.8}$$

where $\theta_{i,\bullet}$, $i = 1 \dots m$ are i.i.d. and independently chosen in $\Theta$ according to the importance sampling density $g(\theta)$. Thus,

$$\hat{\mathrm{E}}\mathrm{v}_{g,m}(W) \overset{m\to\infty}{\to} \mathrm{Ev}(W) \quad \text{a.s. } [g]. \tag{3.9}$$

We have found, for the present problem, that a random choice of $\theta_1$ and $\theta_2$ from gamma distributions, which are conjugate to the Poisson distributions, DeGroot (1970), and a random choice of $\theta_3$ uniformly in $(0, \min(\theta_1, \theta_2))$, yields good results. More definitely, let $g(\bullet \mid \alpha, \beta)$ denote the gamma density with shape parameter $\alpha$ and scale parameter $\beta$, then we used for the simulation

$$g(\theta) = g(\theta_1 \mid \alpha_1, \beta_1) g(\theta_2 \mid \alpha_2, \beta_2) \frac{I\{\theta_3 \leqslant \min(\theta_1, \theta_2)\}}{\min(\theta_1, \theta_2)}, \tag{3.10}$$

$$g(\theta \mid \alpha, \beta) = \theta^{\alpha-1} \mathrm{e}^{-\beta\theta} \beta^\alpha / \Gamma(\alpha). \tag{3.11}$$

Furthermore, we use the conjugate posterior values

$$\alpha_1 = n\bar{X}, \quad \beta_1 = n, \quad \alpha_2 = n\bar{Y}, \quad \beta_2 = n. \tag{3.12}$$

In order to control the number of points, $m$, used at each MC simulation, we can use the asymptotic variance of the MC evidence estimator given by the delta method, Bickel and Doksum (2001).

## 4. Estimating quantiles under the null hypothesis, and power under the alternative

In Table 1, we present quantile estimates $q_{\alpha,n}(\theta)$ for $\alpha = 0.05$, for $\theta_3 = 0$, $\theta_1 = 1$ and values of $\theta_2$ as in the paper of Paul and Ho (1989). The values in brackets are the attained significance level of the MF test in Paul and Ho (1989). Notice that the type I error for the MF test is usually larger than the predicted $\alpha = 0.05$, so the power tables that follow have a consistent bias favoring the MF test.

Table 1
FBST quantiles, $q_{\alpha,n}$, for $\alpha = 0.05$. MF-test significance levels appear in brackets

| $\theta_1 = 1$ | $\theta_2$ | | | |
|---|---|---|---|---|
| $n$ | 0.5 | 1.0 | 1.5 | 2.0 |
| 10 | 0.67 | 0.68 | 0.68 | 0.68 |
| | (0.05) | (0.05) | (0.05) | (0.05) |
| 20 | 0.66 | 0.64 | 0.66 | 0.66 |
| | (0.06) | (0.06) | (0.06) | (0.06) |
| 50 | 0.62 | 0.62 | 0.62 | 0.62 |
| | (0.06) | (0.06) | (0.06) | (0.05) |

Table 2
Power of FBST. MF-test power values appear in brackets

| $\theta_1 = 1$ | | $\theta_3$ | | | | |
|---|---|---|---|---|---|---|
| $\theta_2$ | $n$ | 0.10 | 0.22 | 0.33 | 0.40 | 0.49 |
| 0.5 | 10 | 0.10 | 0.23 | 0.41 | 0.57 | 0.80 |
| | | (0.14) | (0.21) | (0.28) | (0.33) | (0.38) |
| 0.5 | 20 | 0.14 | 0.38 | 0.67 | 0.86 | 0.99 |
| | | (0.17) | (0.32) | (0.44) | (0.52) | (0.62) |
| 0.5 | 50 | 0.25 | 0.69 | 0.96 | 1.00 | 1.00 |
| | | (0.26) | (0.51) | (0.73) | (0.82) | (0.90) |

The power estimates are given in Tables 2 and 3. For a given $\theta = (\theta_1, \theta_2, \theta_3)$, with $\theta_1$ and $\theta_2$ in Table 1 and $\theta_3 > 0$, we had $s = 10^3$ simulation runs. At each run we generated three independent samples of size $n$ from Poissons $p(\theta_1 - \theta_3)$, $p(\theta_2 - \theta_3)$ and $p(\theta_3)$. By adding the first sample to the third one, and the second to the third, we obtain the $X$ and $Y$ samples. Next, we computed $\hat{\text{E}}\text{v}_{g,m}(W)$, initially with $m = m\min = 0.5 \times 10^3$ points.

The proportion of cases, out of $s$ simulation runs in which $\hat{\text{E}}\text{v}_{g,m}(W) > q_{\alpha,n}(\theta)$ is the initial estimated power of the test at $\theta$. As in the quantile estimation, a careful estimation-refinement procedure is necessary to obtain the desired accuracy in reasonable computation time. As shown in Tables 2 and 3, when the sample size is not too small or $\theta_3$ is not too close to zero, the FBST is considerably more powerful than the MF test.

We see from the tables that the quantile and power values change slowly with the parameter $\theta_2$, and smoothly with $\theta_3$. So it is easy to prepare quantile and power tables for a parameter range needed at a given application. Simple interpolation techniques can be used to obtain intermediate values. We also see that the FBST is considerably more efficient than the MF test when the sample size is about 20 and $\rho$ is not too close to zero. Since both tests are consistent, the difference in power diminishes as the sample size $n$ grows. This is shown for the powers when the sample size is $n = 50$.

Table 3
Power of FBST. MF-test power values appear in brackets

| $\theta_1 = 1$ | | $\theta_3$ | | | |
|---|---|---|---|---|---|
| $\theta_2$ | $n$ | 0.11 | 0.22 | 0.45 | 0.67 |
| 1.0 | 10 | 0.09 | 0.15 | 0.38 | 0.73 |
| | | (0.12) | (0.16) | (0.26) | (0.39) |
| 1.0 | 20 | 0.12 | 0.24 | 0.65 | 0.95 |
| | | (0.13) | (0.21) | (0.43) | (0.61) |
| 1.0 | 50 | 0.19 | 0.45 | 0.95 | 1.00 |
| | | (0.19) | (0.35) | (0.72) | (0.90) |
| $\theta_2$ | $n$ | 0.14 | 0.27 | 0.55 | 0.82 |
| 1.5 | 10 | 0.09 | 0.15 | 0.39 | 0.76 |
| | | (0.13) | (0.17) | (0.27) | (0.37) |
| 1.5 | 20 | 0.12 | 0.24 | 0.65 | 0.96 |
| | | (0.13) | (0.23) | (0.43) | (0.61) |
| 1.5 | 50 | 0.20 | 0.45 | 0.95 | 1.00 |
| | | (0.19) | (0.37) | (0.71) | (0.90) |
| $\theta_2$ | $n$ | 0.16 | 0.31 | 0.63 | 0.95 |
| 2.0 | 10 | 0.09 | 0.15 | 0.40 | 0.79 |
| | | (0.12) | (0.17) | (0.27) | (0.38) |
| 2.0 | 20 | 0.12 | 0.24 | 0.65 | 0.98 |
| | | (0.12) | (0.21) | (0.43) | (0.61) |
| 2.0 | 50 | 0.20 | 0.46 | 0.95 | 1.00 |
| | | (0.18) | (0.36) | (0.71) | (0.89) |

## Acknowledgements

The authors are grateful for the support of the Department of Mathematical Sciences at the State University of New York at Binghamton. This work was funded in part by grant FAPESP 2001-03484-1.

## References

Barlow, R.E., Prochan, F., 1981. Statistical Theory of Reliability and Life Testing Probability Models. To Begin With, Silver Spring.

Bickel, P.J., Doksum, K.A., 2001. Mathematical Statistics. Prentice-Hall, London.

Box, G.E.P., Tiao, G.C., 1973. Bayesian Inference in Statistical Analysis. Wiley, New York.

DeGroot, M.H., 1970. Optimal Statistical Decisions. McGraw-Hill, New York.

Good, I.J., 1983. Good Thinking: The Foundations of Probability and its Applications. University of Minnesota Press, Minneapolis, MN.

Holgate, P., 1964. Estimation for the bivariate Poisson distribution. Biometrika 51, 241–245.

Irony, T.Z., Lauretto, M., Pereira, C.A.B., Stern, J.M., 2001. A Weibull wearout test: full Bayesian approach. In: Hayabawa, Y., Irony, T., Xie, M. (Eds.), Systems and Bayesian Reliability. World Scientific, Singapore, pp. 287–300.

Irwin, J.O., 1963. The place of mathematics in medical and biological statistics. J. Royal Statist. Soc. Ser. A 126, 1–44.
Karilis, D., Ntzoufras, I., 1998. Statistical modeling for soccer games. Tech. Rep., Dept. Statistics, Athens University of Economics and Business.
Kocherlakota, S., Kocherlakota, K., 1992. Bivariate Discrete Distributions. Marcel Decker, New York.
Lindley, D.V., 1978. The Bayesian approach. Scand. J. Statist. 5, 1–26.
Liu, J.S., 2001. Monte Carlo Strategies in Scientific Computing. Springer, New York.
Madruga, M.R., Esteves, L.G., Wechsler, S., 2001. On the Bayesianity of Pereira–Stern tests. Test 10, 291–299.
McKendrick, A.G., 1926. Applications of mathematics to medical problems. Proc. Edinburgh Math. Soc. 44, 98–130.
Paul, S.R., Ho, N.I., 1989. Estimation in the bivariate Poisson distribution and hypothesis testing concerning independence. Comm. Statist.—Theory Methods 18, 1123–1133.
Pereira, C.A.B., Stern, J.M., 1999. Evidence and credibility: full Bayesian significance test for precise hypotheses. Entropy 1, 69–80.
Pereira, C.A.B., Stern, J.M., 2001. Model selection: full Bayesian approach. Environmetrics 12 (6), 559–568.
Stern, J.M., 2001. The full Bayesian significance test for the covariance structure problem. ISAS-2001 International Conference on Systems Analysis and Synthesis, Orlando, FL, Vol. 7, pp. 60–65.
Teicher, H., 1954. On the multivariate Poisson distribution. Skand. Aktuar. 37, 1–9.
Zellner, A., 1971. An Introduction to Bayesian Inference in Econometrics. Wiley, New York.