Stochastic Processes and their Applications 37 (1991) 275-279
North-Holland

Independence of partial autocorrelations for a classical immigration branching process

T.M. Mills

Department of Mathematics, Bendigo College of Advanced Education, P.O. Box 199, Bendigo, Vic. 3550, Australia

E. Seneta

Department of Mathematical Statistics, University of Sydney, N.S.W. 2006, Australia

Received 30 April 1990
Revised 11 June 1990

It is shown that for a data set from a branching process with immigration, where the offspring distribution is Bernoulli and the immigration distribution is Poisson, the normed sample partial autocorrelations are asymptotically independent. This makes possible a goodness-of-fit test of known (Quenouille) form. The underlying process is a classical model in statistical mechanics.

autoregression * sample partial autocorrelation * Quenouille's test * subcritical * Galton–Watson * residual autocorrelation

1. Introduction

One of the oldest stochastic processes to be fitted by parameter estimation from data is the subcritical branching (Galton–Watson) process with immigration (BPI); Section 5 of Heyde and Seneta (1972) gives the background. The early context is statistical mechanics. Yet tests of goodness-of-fit of data to such a fitted process have been examined only recently (Venkataraman, 1982; Mills and Seneta, 1989); the approach has come from time series analysis, because of certain similarities of the BPI to the classical AR(1) process.

The special case of particular interest in statistical mechanics is the simple one where the offspring distribution (p.g.f. $F(s)=q+ps$, $0<p<1$) is Bernoulli, and the immigration distribution (p.g.f. $B(s)=\exp\lambda(s-1)$) is Poisson. In our earlier paper (Mills and Seneta, 1989) we developed a test on the basis of the sample partial autocorrelations for a *general* subcritical BPI, and this was applied to a classical data set from a purported Bernoulli–Poisson situation. In Mills (1988) a number of other classical data sets are reproduced and goodness-of-fit investigated.

In the present note we point out the rather startling fact that under the null hypothesis that data comes from a Bernoulli–Poisson BPI, the normed sample partial autocorrelations at lag $\geqslant 2$ are not only asymptotically jointly Gaussian (as is the case in general), but *independent*, permitting an alternative test (in this simple case of most interest) which is almost identical with Quenouille's test for a stationary AR(1) process.

2. General results

We give a formal definition of the BPI $\{X(t)\}$ in order to introduce notation: $X(t)$, $t=0,1,2,\dots$, is defined by

$$X(t)=\begin{cases}\sum_{r=1}^{X(t-1)} Z(t,r)+Y(t) & \text{if } X(t-1)>0,\\ Y(t) & \text{if } X(t-1)=0,\end{cases}$$

where $X(0)$, $Z(t,r)$, $Y(s)$, t, r, $s\geq 1$, are independent non-negative integer-valued random variables, with the $Z(t,r)$ $(t,r\geq 1)$ identically distributed (like a non-degenerate r.v. Z, say, with p.g.f. $F(s)$) and $Y(s)$, $s=1,2,\dots$, identically distributed (like an r.v. Y say, with p.g.f. $B(s)$). The subcritical case has $EZ<1$ and we make this assumption here. For results about the general process the additional assumptions $E(Z^p)<\infty$, $p=2,3$, $E(Y^p)<\infty$, $p=1,2,3$, are needed (these clearly hold in the Bernoulli-Poisson case); and $E(X^2(0))<\infty$. We then need the notation

$$m=EZ,\quad \sigma_1^2=\operatorname{Var} Z,\qquad \lambda=EY,\quad \sigma_2^2=\operatorname{Var} Y,$$

$$\mu=(1-m)^{-1}\lambda,\qquad \sigma_0^2=\mu\sigma_1^2+\sigma_2^2.$$

We assume that we have an observed data sequence $X(0), X(1),\dots,X(N)$, and

$$R(k)=\sum_{t=1}^{N-k}(X(t)-\bar{X})(X(t+k)-\bar{X})\Big/\sum_{t=1}^{N}(X(t)-\bar{X})^2,$$

$$\bar{X}=N^{-1}\sum_{t=1}^{N}X(t),$$

so that $R(k)$, $k\geq 1$, are the sample autocorrelations of the process. The sample partial autocorrelations $\hat{\beta}_k$, $k\geq 1$, are defined in a standard way in terms of them; see Mills and Seneta (1989). According to Venkataraman (1982, Theorem 4.3, part (b), and Theorem 1.1), $N^{1/2}(R(k)-m^k)$, $k=1,2,\dots,H$, converges in distribution to a normal vector $\{V(k),k=1,2,\dots,H\}$ where

$$V(k)=\sigma_0^{-2}\sum_{u=1}^{\infty}(m^{|u-k|}-m^{u+k})\xi(u),\quad k=1,\dots,H, \tag{1}$$

where $\xi(s)$, $s\geq 1$, is a zero-mean Gaussian process, with

$$E((\xi(s))^2)=\sigma_0^4+\sigma_1^2 m^{s-1}\theta,$$

where

$$\theta=\{(1-m^2)^{-1}m\sigma_0^2\sigma_1^2+B_1\mu+B_2\},\quad B_1=E((Z-m)^3),\ B_2=E((Y-\lambda)^3),$$

and

$$E(\xi(r)\xi(s)) = \sigma_0^2\sigma_1^4 m^{r-2}, \quad r > s \geq 1.$$

Then by Theorem 1 of Mills and Seneta (1989) the vector $\{N^{1/2}\hat{\beta}_k, k = 2, \dots, H\}$ converges in distribution as $N \to \infty$ to the vector $\{W(k), k = 2, \dots, H\}$ where

$$W(k) = (V(k) - 2mV(k-1) + m^2 V(k-2))/(1 - m^2) \tag{2}$$

with $V(0) = 0$ by definition.

Clearly $\{W(k), k = 2, \dots, H\}$ has a multivariate normal distribution, $EW(k) = 0$, and (1) and (2) imply that

$$W(k) = \sigma_0^{-2}\left(-m\xi(k-1) + (1 - m^2)\sum_{u=k}^{\infty} m^{u-k}\xi(u)\right), \quad k \geq 2. \tag{3}$$

From (3) after tedious calculations checked numerous times by each of us individually, and independently by Dr N. Weber, to whom our thanks are due, we obtain

$$E(W^2(k)) = 1 + \frac{m^{k-1}}{1 + m + m^2}\left[\frac{\sigma_1^2\theta(1 + 2m)}{\sigma_0^4} - \frac{2\sigma_1^4}{\sigma_0^2}\right] \tag{4}$$

for $k \geq 2$, while

$$E(W(k)W(s)) = \frac{m^{2s-k-3}}{1 + m + m^2}\{\sigma_0^{-2}\sigma_1^4(1 + m - m^2) - \sigma_0^{-4}\sigma_1^2\theta m(1 - m^2)\} \tag{5}$$

for $2 \leq k < s \leq H$.

3. Bernoulli offspring distribution and Poisson immigration distribution

In this case $\sigma_1^2 = m(1-m)$, $B_1 = m(1-m)(1-2m)$, $\sigma_2^2 = \lambda$, $B_2 = \lambda$, $\mu = \lambda/(1-m)$, $\sigma_0^2 = \lambda(1+m)$, $\theta = \lambda(1 + m - m^2)$.

These formulae, when substituted into (4), (5) yield the result that for this special Galton-Watson process with immigration

$$\sigma^2(k) \equiv E(W^2(k)) = 1 + \frac{m^k(1-m)(1-m-m^2+2m^3)}{\lambda(1+m)^2(1+m+m^2)}, \quad k \geq 2, \tag{6}$$

and the centrally important and surprising result that for $2 \leq k < s$,

$$E(W(k)W(s)) = 0.$$

Hence for this special case, $N^{1/2}\hat{\beta}_k$ and $N^{1/2}\hat{\beta}_s$ for $2 \leq k < s$ are asymptotically uncorrelated, and hence (by normality) *independent*, in complete analogy to the stationary AR(1) model. From this observation we can develop a goodness-of-fit test for this special progress on the basis of sample partial autocorrelations completely analogous to Quenouille's test. A further parallel follows from the fact that for moderate m, clearly $\sigma^2(k) \simeq 1$ for most $k \geq 2$; this will be seen in the example below. Recall that for the general Galton–Watson process with immigration (Mills and Seneta, 1989, Theorem 2), the best that can be done is independence of linear forms of two successive $\hat{\beta}_k$, at lags ≥ 2.

In the present case on the basis of a data set $X(0), \ldots, X(N)$, approximately

$$N\hat{\beta}^2_{k+1}/\hat{\sigma}^2(k+1) \sim \chi^2_1, \qquad N \sum_{k=1}^{T} \hat{\beta}^2_{k+1}/\hat{\sigma}^2(k+1) \sim \chi^2_T \tag{7}$$

for any fixed T, $2 \leq T+1 \leq H$, where $\hat{\sigma}^2(k)$ is a consistent estimator of $\sigma^2(k)$, which can be obtained by replacing m and λ in (6) by their consistent estimators as $N \to \infty$, such as the least squares estimators $\hat{m}$, $\hat{\lambda}$ as in Venkataraman (1982, p. 3) as used in Mills and Seneta (1989).

As an illustration we apply the preceding theory to a data set of the physicist Fürth (1918, 1919). As noted in Heyde and Seneta (1972) Section 5, the Bernoulli–Poisson model was the one Fürth used for purposes of estimation; and as noted in Mills and Seneta (1989, Section 5) the general GWI is found to be a poor fit, which we expect to be even more so the case for the present more restrictive (Bernoulli–Poisson) null hypothesis. Fürth's data ($N = 505$) gives $\hat{m} = 0.665776$, $\hat{\lambda} = 0.532112$. The same simulated data set as in Mills and Seneta (1989) for an actual Bernoulli–Poisson process (with $N = 505$, $m = 0.665776$, $\lambda = 0.532112$) was used for comparison; the least squares estimates $\hat{m} = 0.698434$ and $\hat{\lambda} = 0.492469$ resulted in this case. (See Tables 1 and 2.)

The values of Quenouille's statistic, $N \sum_{k=1}^{T} \hat{\beta}^2_{k+1}$, which for these data sets is clearly almost coincident with our statistic occurring in (7) calculated for $T = 40$ are 70.56 (Fürth's data) and 49.84 (simulated data), while $\chi^2_{40}\,(0.05) = 55.76$. Thus the Bernoulli–Poisson GWI hypothesis is convincingly rejected for Fürth's data by the portmanteau test based on (7), while the simulated data leads to clear acceptance.

Table 1

Goodness-of-fit of Fürth's data to a BPI

k	1	2	3	4	5	6	7	8	9	10
$\hat{\sigma}^2(k+1)$	1.02	1.02	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00
$N\hat{\beta}^2_{k+1}/\hat{\sigma}^2(k+1)$	21.82	4.52	2.07	0.17	3.17	1.39	0.21	6.37	2.07	0.41

Table 2

Goodness-of-fit of simulated data

k	1	2	3	4	5	6	7	8	9	10
$\hat{\sigma}^2(k+1)$	1.02	1.02	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00
$N\hat{\beta}^2_{k+1}/\hat{\sigma}^2(k+1)$	0.05	0.11	3.89	0.71	0.52	0.34	0.44	0.17	0.02	0.14

A portmanteau test statistic value based on $\hat{\beta}_k$, $k=3,\ldots,10$, obtained by summing over columns 2 to 9 in the second row of our Table 1 yields 19.97 (to be compared with $\chi_8^2(0.05)=15.51$), and hence to rejection of the Bernoulli–Poisson hypothesis. A corresponding portmanteau statistic value based on the same β_k's, is obtained by summing over columns 2, 4, 6, 8 in the last row of Table 1 of Seneta and Mills (1989) which gives 3.43 (to be compared with $\chi_4^2(0.05)=9.49$), and hence is far from significant. The arbitrariness inherent in constructing a portmanteau statistic on the basis of the results for a *general* GWI can thus be seen to lead to incorrect results in carelessly using such a statistic, but there are no such problems in the present Bernoulli–Poisson situation.

The quantities $\{N^{1/2}\tilde{r}(k), k=1,\ldots,T\}$, where the $\tilde{r}(k)$ are the sample autocorrelations of residuals $\hat{\varepsilon}(t)=X(t)-\hat{m}X(t-1)-\hat{\lambda}$, $t=1,\ldots,N$, of least squares fit, are asymptotically normal but not asymptotically independent for the Bernoulli–Poisson GWI (Venkataraman, 1982, Theorems 1.1, 4.1, and Lemma 5.1; Klimko and Nelson, 1978, Section 5). They cannot therefore be used in the same way as the $\hat{\beta}_k$'s to construct a portmanteau statistic, although the parallelism exhibited between the $\hat{\beta}_k$'s and $\tilde{r}_k$'s in Mills and Seneta (1989) might suggest this.

References

R. Fürth, Statistik und Wahrscheinlichkeitsnachwirkung, Z. Physik, 19 (1918) 421–426; *ibid.* 20 (1919) 21.

C.C. Heyde and E. Seneta, Estimation theory for growth and immigration rates in a multiplicative process, J. Appl. Probab. 9 (1972) 235–256.

L.A. Klimko and P.I. Nelson, On conditional least squares estimation for stochastic processes, Ann. Statist. 6 (1978) 629–642.

T.M. Mills, Residual analysis for the branching immigration process, MA Thesis, University of Sydney (1988).

T.M. Mills and E. Seneta, Goodness-of-fit for a branching process with immigration using sample partial autocorrelations, Stochastic Process. Appl. 33 (1989) 151–161.

K.N. Venkataraman, A time series approach to the study of the simple subcritical Galton–Watson process with immigration, Adv. in Appl. Probab. 14 (1982) 1–20.