SUFFICIENCY AND THE NUMBER OF LEVEL CROSSINGS BY A STATIONARY PROCESS

by

Benjamin Kedem

DEPARTMENT OF MATHEMATICS
University of Maryland
College Park, Maryland
20742

MD77-52-BK
TR77-41
August 1977

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is approved for public release IAW AFR 150-12 (7b), distribution is unlimited.
A. D. RECE
Technical Information Officer
SUFFICIENCY AND THE NUMBER OF LEVEL CROSSINGS BY A STATIONARY PROCESS

Benjamin Kedem
University of Maryland, College Park

Summary. It is shown how to derive the exact distribution of the number of axis crossings by a stationary process when the binary process obtained by clipping the original process is a pth-order Markov chain. The same method is used in deriving the asymptotic distribution of the number of upcrossings of a high level by a stationary process.

Key words and phrases: binary, Markov chain, level crossings, symbol changes, upcrossings, high level

1. Introduction. Let \( Z_t, t=1,...,n, \) be a strictly stationary time series, and let \( X_t, t=1,...,n, \) be the binary time series which takes the values 1 whenever \( Z_t \geq a \) and 0 otherwise. \( X_t \) as well as quantities defined by it should be indexed by the level \( a \), but except in one case we shall avoid this indexing for the sake of simplified notation.

Associated with \( X_t \) are the statistics

\[
D(n) = 2 \sum_{t=1}^{n} X_t X_{t-1} - \frac{1}{2} \sum_{t=2}^{n} X_t X_{t-1} - (X_1 + X_n) \quad \text{and} \quad D_a(n) = \sum_{t=1}^{n} X_t - \sum_{t=2}^{n} X_t X_{t-1}.
\]

\( D(n) \) counts the number of symbol changes in the binary series and hence it counts the number of crossings of level \( a \) by \( Z_t \). When \( X_1 + X_n = 0 \), \( D_a(n) \) counts the number of upcrossings of level \( a \) by \( Z_t \). We shall find the distribution of \( D(n), n \) fixed, for level \( a = 0 \) and the asymptotic distribution of \( D_a(n), a, n \to \infty \) in a suitable manner, when \( X_t \) is either a first or second-order Markov chain. The same technique applies to higher order chains.

We shall make use of the results in Kedem (1976,a). Consequently we define
\[ p = \Pr(Z_t \geq a), \quad \lambda_k = \Pr(Z_t \geq a | Z_{t-k} \geq a), \quad k=1,2 \]

\[ \mu = \Pr(Z_t \geq a | Z_{t-1} \geq a, Z_{t-2} \geq a), \]

\[ S = \sum X_t, \quad R_1 = \sum X_t X_{t-1}, \quad R_2 = \sum X_t X_{t-2}, \quad C = \sum X_t X_{t-1} X_{t-2}, \quad H = X_1 + X_n, \]

\[ U = X_2 + X_{n-1}, \quad V = X_1 X_2 + X_{n-1} X_n. \]

For a review of level crossings problems and an extensive bibliography see Leadbetter (1972).

2. The number of axis crossings. In this section \( a = 0 \) and \( p = 1/2 \). That is \( \Pr(Z_t = 0) = 1/2 \).

**Theorem 1.** If \( X_t \) is a first-order Markov chain, then the number of axis crossings by \( Z_t, t=1, \ldots, n \), has a binomial distribution \( b(n-1, 1-\lambda_1) \).

**Proof.** The probability of a 0-1 series for which \( D(n) = d \) is given by

\[ \Pr(X_1 = x_1, \ldots, X_n = x_n) = \frac{1}{2} (1-\lambda_1)^d \lambda_1^{(n-1)-d} \]  \( (1) \)

and there are \( 2^{(n-1)} \) such sequences. Multiply this number by \( (1) \) to obtain the desired binomial distribution.

Observe that under the conditions of the theorem \( D(n) \) is minimal sufficient for \( \lambda_1 \) and the maximum likelihood estimate of \( \lambda_1 \) is \( \hat{\lambda}_1 = \{(n-1)-D(n)\}/(n-1) \) while \( \sqrt{n}(\hat{\lambda}_1-\lambda) \) is asymptotically \( N(0, \lambda_1(1-\lambda_1)) \).

Just when may we expect the above binomial distribution to be a reasonable approximation to the actual distribution of the number of axis crossings? So, consider a stationary AR(1) process \( Z_t = \phi Z_{t-1} + u_t, \quad |\phi| < 1, \quad u_t \) are independent \( N(0,1) \) variates. For each of 19 values of \( \phi \) 1000 time series of size 5 were generated. The size was fixed at 5 to allow the expected number of successes in each cell in a multinomial experiment to exceed 1 in 1000 repetitions. We wish to test \( H_0: D(n) \sim b(n-1, 1-\lambda_1) \) where now \( n = 5 \) and \( \lambda_1 = \frac{1}{2} + \frac{1}{\pi} \sin^{-1}(\phi) \). The
results of the 19 chi-square goodness of fit tests are summarized in Table 1. It is seen that the results are very satisfactory for $-0.588 < \phi < 0.600$ where $H_0$ is accepted at level of significance 0.01. This example indicates that the above binomial distribution is reasonable when neighboring observations in the $Z_t$ series are at most moderately correlated.

Table 1: Observed (expected) frequencies of the number of axis crossings by $Z_t = \phi Z_{t-1} + \mu_t$, $t = 1, \ldots, 5$, $\mu_t \sim N(0,1)$, in 1000 independent realizations.

<table>
<thead>
<tr>
<th>$\phi$</th>
<th>$\lambda$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\chi^2(4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.809</td>
<td>.800</td>
<td>457</td>
<td>322</td>
<td>171</td>
<td>43</td>
<td>7</td>
<td>56.235*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.409.6)</td>
<td>(.409.6)</td>
<td>(153.6)</td>
<td>(.25.6)</td>
<td>(1.6)</td>
<td></td>
</tr>
<tr>
<td>.707</td>
<td>.750</td>
<td>358</td>
<td>366</td>
<td>190</td>
<td>73</td>
<td>13</td>
<td>50.716*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.316.4)</td>
<td>(.421.87)</td>
<td>(.210.94)</td>
<td>(.46.87)</td>
<td>(3.91)</td>
<td></td>
</tr>
<tr>
<td>.600</td>
<td>.705</td>
<td>259</td>
<td>405</td>
<td>243</td>
<td>80</td>
<td>13</td>
<td>6.505</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.247)</td>
<td>(.413.4)</td>
<td>(.260)</td>
<td>(.72.3)</td>
<td>(7.6)</td>
<td></td>
</tr>
<tr>
<td>.588</td>
<td>.700</td>
<td>253</td>
<td>400</td>
<td>252</td>
<td>81</td>
<td>14</td>
<td>6.303</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.240.1)</td>
<td>(.411.6)</td>
<td>(.264.6)</td>
<td>(.75.6)</td>
<td>(8.1)</td>
<td></td>
</tr>
<tr>
<td>.500</td>
<td>.666</td>
<td>209</td>
<td>391</td>
<td>290</td>
<td>87</td>
<td>23</td>
<td>11.549</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.197.4)</td>
<td>(.394.9)</td>
<td>(.296.2)</td>
<td>(.98.72)</td>
<td>(12.3)</td>
<td></td>
</tr>
<tr>
<td>.400</td>
<td>.631</td>
<td>169</td>
<td>365</td>
<td>318</td>
<td>119</td>
<td>29</td>
<td>7.387</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.158.5)</td>
<td>(.370.8)</td>
<td>(.325.2)</td>
<td>(.126.8)</td>
<td>(18.5)</td>
<td></td>
</tr>
<tr>
<td>.309</td>
<td>.600</td>
<td>143</td>
<td>331</td>
<td>336</td>
<td>157</td>
<td>33</td>
<td>4.483</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.129.6)</td>
<td>(.345.6)</td>
<td>(.345.6)</td>
<td>(.153.6)</td>
<td>(25.6)</td>
<td></td>
</tr>
<tr>
<td>.250</td>
<td>.580</td>
<td>127</td>
<td>315</td>
<td>344</td>
<td>173</td>
<td>41</td>
<td>5.783</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.113.1)</td>
<td>(.327.8)</td>
<td>(.356)</td>
<td>(.171.2)</td>
<td>(31.1)</td>
<td></td>
</tr>
<tr>
<td>.200</td>
<td>.532</td>
<td>85</td>
<td>281</td>
<td>363</td>
<td>215</td>
<td>56</td>
<td>1.908</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.80)</td>
<td>(.282)</td>
<td>(.372)</td>
<td>(.218)</td>
<td>(48)</td>
<td></td>
</tr>
<tr>
<td>.100</td>
<td>.500</td>
<td>64</td>
<td>254</td>
<td>378</td>
<td>233</td>
<td>71</td>
<td>2.436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.62.5)</td>
<td>(.250)</td>
<td>(.375)</td>
<td>(.250)</td>
<td>(.62.5)</td>
<td></td>
</tr>
<tr>
<td>-.100</td>
<td>.468</td>
<td>53</td>
<td>217</td>
<td>366</td>
<td>374</td>
<td>90</td>
<td>2.099</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.48)</td>
<td>(.218)</td>
<td>(.372)</td>
<td>(.282)</td>
<td>(80)</td>
<td></td>
</tr>
<tr>
<td>-.250</td>
<td>.420</td>
<td>37</td>
<td>175</td>
<td>345</td>
<td>315</td>
<td>128</td>
<td>4.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.31.1)</td>
<td>(.171.2)</td>
<td>(.356)</td>
<td>(.327.8)</td>
<td>(.113.1)</td>
<td></td>
</tr>
<tr>
<td>-.309</td>
<td>.400</td>
<td>30</td>
<td>160</td>
<td>341</td>
<td>326</td>
<td>143</td>
<td>3.580</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.25.6)</td>
<td>(.153.6)</td>
<td>(.345.6)</td>
<td>(.345.6)</td>
<td>(.129.6)</td>
<td></td>
</tr>
<tr>
<td>-.400</td>
<td>.369</td>
<td>22</td>
<td>133</td>
<td>335</td>
<td>335</td>
<td>175</td>
<td>6.433</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.18.5)</td>
<td>(.126.8)</td>
<td>(.325.2)</td>
<td>(.370.8)</td>
<td>(.158.5)</td>
<td></td>
</tr>
<tr>
<td>-.500</td>
<td>.333</td>
<td>18</td>
<td>113</td>
<td>294</td>
<td>365</td>
<td>210</td>
<td>7.791</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.12.3)</td>
<td>(.98.72)</td>
<td>(.296.2)</td>
<td>(.394.9)</td>
<td>(.197.4)</td>
<td></td>
</tr>
<tr>
<td>-.588</td>
<td>.300</td>
<td>18</td>
<td>97</td>
<td>258</td>
<td>358</td>
<td>269</td>
<td>28.779*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.8.1)</td>
<td>(.75.6)</td>
<td>(.264.6)</td>
<td>(.411.6)</td>
<td>(.240.1)</td>
<td></td>
</tr>
<tr>
<td>-.600</td>
<td>.295</td>
<td>18</td>
<td>93</td>
<td>259</td>
<td>358</td>
<td>272</td>
<td>30.116*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.7.6)</td>
<td>(.72.3)</td>
<td>(.260)</td>
<td>(.413.4)</td>
<td>(.247)</td>
<td></td>
</tr>
<tr>
<td>-.707</td>
<td>.250</td>
<td>12</td>
<td>71</td>
<td>213</td>
<td>343</td>
<td>361</td>
<td>50.268*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.3.91)</td>
<td>(.46.87)</td>
<td>(.210.94)</td>
<td>(.421.87)</td>
<td>(.316.4)</td>
<td></td>
</tr>
<tr>
<td>-.809</td>
<td>.200</td>
<td>10</td>
<td>37</td>
<td>157</td>
<td>305</td>
<td>491</td>
<td>92.191*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.1.6)</td>
<td>(.25.6)</td>
<td>(.153.6)</td>
<td>(.409.6)</td>
<td>(.409.6)</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates that the hypothesis $H_0$ is rejected at level of significance 0.01.
A more realistic assumption is that $X_t$ displays a higher order dependence. The extension of Theorem 1 to the case when $X_t$ is a $k$th-order Markov chain is somewhat more involved but straightforward. For this purpose let us consider the second-order case in detail; the $k$th-order case follows an identical argument.

When $X_t$ is a second-order chain it was shown in Kedem (1976,a) that \{S,R_1,R_2,C,H,U,V\} is a set of sufficient statistics for $\lambda_1, \lambda_2, \mu$, and their joint distribution is given there. An equivalent but a more convenient set of sufficient statistics is \{S,D(n),F,Z',H,U,V\} where $D(n) = 2S - 2R_1 - H$, $F = R_1 - C$ is the number of 1-runs in the $X_t$ series with two or more 1's and $Z' = R_2 - C$ is the number of 0-runs between the first and last 1 with exactly one 0. It follows that the joint distribution of the last set can be obtained from that of the first one. We have

\[
g(s,d,f,z',h,u,v) = P_r(S=s, D(n)=d, F=f, Z'=z', H=h, U=u, V=v)
= N(s,d,f,z',h,u,v)K_n(\xi_1, \xi_2, \xi_3, \xi_4) S(\xi_2, \xi_3, \xi_4)^{-\frac{1}{2}d} (\xi_3, \xi_4)^{-f} \xi_3' \cdot [((\xi_2, \xi_3, \xi_4)^{-\frac{1}{2}\xi_5})^h \xi_6, \xi_7],
\]

where $K_n, \xi_1, \xi_2, \ldots, \xi_7$ are functions of $p = 1/2, \lambda_1, \lambda_2, \mu$ and are given in Kedem (1976) and

\[
N(s,d,f,z',h,u,v)
= \left(\begin{array}{ccc}
2 & \max(h,u) & \frac{1}{2}(d+h)-1 \\
\max(h,u) & v & \frac{n-s-\frac{1}{2}(d-h)-2}{}
\end{array}\right) \cdot \left(\begin{array}{ccc}
\frac{1}{2}(d-h) & s-\frac{1}{2}(d+h)-1 \\
\frac{1}{2}(d-h)-z'-u+v & f-v & f-1
\end{array}\right)
\]

**Theorem 2.** If $X_t$ is a second order Markov chain then the distribution of the number of crossings by $Z_t$, $t=1, \ldots, n$ is given by

\[
P_r(D(n)-d) = \sum_{(h,u,v)} \sum_{s=h+u} \sum_{f=v} \sum_{z'=0} g(s,d,f,z',h,u,v),
\]
where \((h,u,v)\) takes values in \((1,2,1),(1,1,1),(1,1,0),(1,0,0)\) when \(d\) is odd and in \((2,2,2),(2,1,1),(0,2,0),(2,0,0),(0,1,0),(0,0,0)\) when \(d\) is even.

In principle it is possible to extend our method to obtain the distribution of \(D(n)\) when the 0-1 series displays a higher order dependence but the joint distribution of the sufficient statistics becomes messier.

3. Upcrossings of a high level. In this section we shall elicit the Poisson nature of the upcrossings of a high level \(a\) by \(Z_t\), by using the above method of examining the joint distribution of several sufficient statistics. The Poisson nature of these upcrossings\([3]\) has been known for nearly twenty years for continuous parameter Gaussian processes under various moment conditions. \(Z_t\), however, is not necessarily Gaussian.

**Theorem 3.** Assume \(X_t\) is a first order Markov chain. If \(a,n \to \infty\) such that

\[
\begin{align*}
(i) & \quad nP_r(Z_t \geq a) = a, \quad a \text{ remains constant}, \\
(ii) & \quad P_r(Z_t \geq a | Z_{t-1} \geq a) = \lambda_1(a) \to \lambda_1,
\end{align*}
\]

then

\[
\lim_{a \to \infty} P_r(D_a(n)=k) = \frac{e^{-a(1-\lambda_1)}}{\lambda_1^k k!} [a(1-\lambda_1)]^k, \quad k=0,1,\ldots
\]

**Proof.** A simple combinatorial argument shows that

\[
\begin{align*}
P_r(S=s, D_a(n) = k, H=0) &= \binom{s-1}{s-k} \binom{n-s-1}{k} p^s q^{s-n+2} \lambda_1^{s-k} (1-\lambda_1)^{2k} \\
&\quad \cdot (1-2p+\lambda_1)^{n-s-k}.
\end{align*}
\]

Replace \(p\) by \(a/n\) and \(q\) by \(1-a/n\) and note that \(\{H=0\}\) becomes a sure event as \(a \to \infty\). Then
\[
\lim_{a \to \infty} P_r(S=s, D_a(n)=k) = \frac{[a(1-\lambda_1)]^k e^{-a(1-\lambda_1)}}{k!} \left(\frac{s-l}{k-l}\right) (1-\lambda_1)^k \lambda_1^{s-k},
\]
and sum over \( s \).

As consequences we have firstly
\[\lim_{a \to \infty} P_r(\max_{t=1,...,n} Z_t \leq a) = e^{-a(1-\lambda_1)},\]
and secondly, the asymptotic distribution of \( S \), the total time spent above a high level \( a \), is the Polya-Aeppli distribution obtained by summing (5) over \( k \), with mean \( a \) and variance \( a(1+\lambda_1)/(1-\lambda_1) \).

Similar results can be obtained for the second-order case. To simplify matters assume \( Z' \to 0 \) as \( a \to \infty \) with probability one, which happens if and only if \( \lambda_2 - \lambda_1 > 0 \), \( a \to \infty \).

**Theorem 4.** If \( X_t \) is a second-order Markov chain such that (i) and (ii) above hold and (iii) \( P_r(Z_t \geq a | Z_{t-1} \geq a, Z_{t-2} \geq a) = \mu(a) \to \mu \), (iv) \( (\lambda_2 - \lambda_1) Z' \to 1 \) with probability one, as \( a \to \infty \),

then \( D_a(n) \) has an asymptotic Poisson distribution with parameter \( a(1-\lambda_1) \).

**Proof.** From (2) with \( p = a/n \) and the fact that \{H=0, U=0, V=0\} becomes a sure event, it follows that
\[\lim_{a \to \infty} P_r(S=s, D_a(n)=k, F=f) = \frac{a^k e^{-a(1-\lambda_1)}}{k!} \binom{k}{f} \binom{s-k-1}{f-1} \lambda_1^f \mu^{s-k-f} (1-\mu)^{2f} (1-\lambda_1(2-\mu))^{k-f}.\]
But
\[\sum_{s=k+f}^{\infty} \binom{s-k-1}{f-1} \mu^{s-k-f} = (1-\mu)^{-f}\]
and
so that \( P_r(D_a(n) = k) \to e^{-a(1-\lambda_1)} [a(1-\lambda_1)]^k/k! \).

4. Some applications.

When parameters of interest are related in some fashion to the number of axis crossings, Theorem 1 can be used in deriving appropriate estimators and their approximate distributions. We bring two such cases.

Estimation in AR(1). Suppose \( Z_t = \phi Z_{t-1} + u_t \) is a stationary AR(1) process as above, and suppose it is clipped at level \( D \). If the clipped process \( X_t \) approximates a first order Markov chain, then the maximum likelihood estimate of \( \phi \) based on the clipped data is

\[
\hat{\phi} = \phi(\hat{\lambda}_1) = \sin n \left\{ \frac{(n-1) - (\# \text{ of axis crossings})}{n-1} - \frac{1}{2} \right\}
\]

Experience shows \([2]\) that this estimator behaves remarkably well even when \(|\phi|\) is close to 1. When \(|\phi|\) is small so that the binomial approximation to the distribution of the number of axis crossings is adequate, it follows directly that

\[
\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{L} N(0, \pi^2 \lambda_1(1-\lambda_1)), \quad n \to \infty.
\]

Estimation of the mean frequency. Let \( Z(t) = <t < \infty \), be a zero mean stationary Gaussian process with correlation function \( \rho(t) \). Assume that the sample functions are continuous with probability one and that in a sufficiently small time interval, say \( \Delta \), the
probability that \(Z(t)\) has more than one 0 is negligible. Consider the interval \([0,t]\) and partition it into \((n-1)\) subintervals of size \(\Delta\). Then \((n-1)\Delta = T\). We hold \(T\) fixed as \(\Delta \to 0\) and \(n \to \infty\) simultaneously. Let \(X_{i,n}\) take the value 1 when \(Z((i-1)\Delta) \geq 0\), and 0 otherwise, \(i = 1, \ldots, n\), and let \(D(n)\) be the number of symbol changes in the \(X_{i,n}\) series. If \(D\) is the true number of axis crossings in \([0,T]\) then \(D(n) \to D\), \(n \to \infty(\Delta \to 0)\) a.s. As a first approximation to the distribution of \(D(n)\) we take \(D(n) \sim b(n-1, 1-\lambda_{1,n})\). Thus, by l'Hospital's rule

\[
E(D) = \lim_{n \to \infty} E(D(n)) = \lim_{\Delta \to 0}^{n \to \infty} \frac{T}{\Delta} \left[ \frac{1 - \frac{1}{\pi}}{\sin^{-1}(\rho(\Delta))} \right] = \frac{T}{\pi} \gamma,
\]

\[
\gamma = [-\rho''(0)]^{1/2}
\]

provided the derivative exists. \(\gamma\) is called the mean frequency.

A reasonable estimate for \(\gamma\) is then \([4]\)

\[
\hat{\gamma} = \frac{\pi D(n)}{T}
\]

whose approximate distribution is easily obtained from that of \(D(n)\).

References


This research was supported by the U.S. Air Force Office of Scientific Research Grant AFOSR 77-3332.
It is shown how to derive the exact distribution of the number of axis crossings by a stationary process when the binary process obtained by clipping the original process is a \( p \)-th order Markov chain. The same method is used in deriving the asymptotic distribution of the number of upcrossings of a high level by a stationary process.