On a Test of Randomness of Spatial Point Patterns

L. Heinrich

Summary. For an estimator of the reduced second moment measure (suggested by Ohsen & Stoyan) asymptotic normality is proved, if the underlying stationary point process is Poissonian and the area of the rectangular sampling window is very large. The variance of the estimator and a rate of convergence are given. Basing on it, a rejection region for testing the hypothesis of an underlying Poisson process is constructed.

Key words: Point pattern, reduced second moment measure, asymptotic normality, test for randomness.

1. Introduction

In many biological problems planar random point patterns are occurring (see e.g. Pielou [7] and Ripley [9]). When it is possible to assume that such a pattern is a (piece of a) stationary isotropic point process, the second order analysis is an important step in its statistical analysis. It consists mainly in determining and studying the so-called reduced second moment measure $K(r)$ (see Ripley [8], [9] and Ohsen & Stoyan [4]). For fixed $r > 0$, $\lambda K(r)$ is the mean number of random points in a circle with radius $r$ centred in a "typical point" of the pattern, and $\lambda$ denotes the intensity, the mean number of points per unit area. If the point pattern is a Poisson process, then $K(r) = \pi r^2$. Consequently, if a curve $\hat{K}(r)$ estimated from a sample differs from $\pi r^2$, one should reject the hypothesis that the point pattern under consideration is a Poisson process. For doing this test in a suitable manner, the distribution of an estimator of $K(r)$ in the case of a Poisson process is needed. By simulation Ripley [8], p. 185, determined a symmetric critical region (of size $\alpha = 0.01$) for his unbiased estimator $\hat{K}(r)$, whereas Silverman [11] proved its asymptotic normality. He assumed that the number of points in the sampling window tends to infinity. Consequently, for a large number of points in the sampling window the distribution of $\hat{K}(r)$ is theoretically known and when using an approximation of the variance of $\hat{K}(r)$ (see Ripley [9], p. 163), this result can be used for testing the hypothesis of an underlying Poisson process.

In this paper, another approach to the asymptotic normality of an estimator

1 Bergakademie Freiberg, Sektion Mathematik, DDR - 9200 Freiberg, Bernhard-von-Cotta-Str. 2.
for \( K(r) \) is given. The estimator is that suggested by Ohser & Stoyan [4] and Ohser [5],

\[
\hat{K}_F(r) = \frac{1}{\lambda^2} \sum_{x,y \in F} \frac{1_{b(0,r)}(x-y) \ 1_F(x) \ 1_F(y)}{v(F \cap F_{x-y})}
\]

where the notations are listed below. Here, \( F \) is the sampling window, which is assumed to be a "very large" rectangle. Furthermore, the variance of \( \hat{K}_F(r) \) is approximately calculated. Finally, the bounds of the rejection region (of size \( z = 0.01 \)) obtained by this method are compared with Ripley's simulated values.

**Notations:**

\( \sum \forall x \neq y \): summation over pairwise different points,

\( v \): Lebesgue measure on \( \mathbb{R}^2 \),

\( b(x, r) \): (open) sphere in \( \mathbb{R}^2 \) with centre \( x \) and radius \( r \),

\( \mathcal{P} \): stationary random point process on \( \mathbb{R}^2 \),

\( 1_A(x) = 1 \) if \( x \in A \), \( = 0 \) otherwise,

\( A_x = A + x, \quad A \subset \mathbb{R}^2, \quad x \in \mathbb{R}^2 \),

\( \Phi(x) = (2\pi)^{-1/2} \int \frac{e^{-u^2/2}}{\sqrt{2\pi}} \, du \) standard normal distribution,

\( \gamma, \Phi^{-1}(\gamma) \): quantile of order \( \gamma \),

\( c_1, c_2, \ldots \): positive constants.

---

### 2. Results and proofs

At first we compute and estimate the variance of \( \hat{K}_F(r) \).

**Lemma.** Let \( \mathcal{P} \) be a stationary Poisson process with intensity \( \lambda \) and let \( F = [0, a] \times [0, b] \). Then, for \( r < \min(a, b) \), the variance of \( \hat{K}_F(r) \) is equal to

\[
\frac{2}{\lambda^2} \int_{R^2} \frac{1_{b(0,r)}(u)}{v(F \cap F_u)} \, du + \frac{4}{\lambda^2} \int_{(R^2)^2} \frac{1_{b(0,r)}(u) \ 1_{b(0,r)}(v)}{v(F \cap F_u \cap F_v)} \, du \ (F \cap F_{u,v})
\]

and for \( r < \frac{1}{2} \min(a, b) \), the estimate

\[
\text{Var} \hat{K}_F(r) = \frac{2\pi r^2}{\lambda^2 ab} (1 + 2\pi r^2 \lambda) \left( 1 + \frac{\Theta_1 r}{a - r} \right) \left( 1 + \frac{\Theta_2 r}{b - r} \right)
\]

holds with appropriate \( \Theta_1, \Theta_2 \in [-1, 1] \).

We direct the reader's attention to the remarkable fact that the variance of \( \hat{K}_F(r) \) is not proportional to \( \lambda^{-2} \) but the order of \( \text{Var} \hat{K}_F(r) \) is indirect proportional to the area of the sampling window. Using the fact that the coefficients of \( \lambda^{-2} \) and \( \lambda^{-1} \) only depend on the ratios \( r/\sqrt{a^2 + b^2} \) and \( a/b \), Ohser [5] has given a table for them.

**Proof.** Let

\[
f(x, y) = f(y, x) = \frac{1_{b(0,r)}(x-y) \ 1_F(x) \ 1_F(y)}{v(F \cap F_{x-y})}
\]
We have

\[(\lambda^2 K_F(r))^2 = 2 \sum_{x_1, x_2 \in W} f^2(x_1, x_2) + 4 \sum_{x_1, x_2, x_3 \in W} f(x_1, x_2) f(x_2, x_3) \]

Using the well-known fact that

\[E \sum_{x_1, \ldots, x_n \in W} g(x_1, \ldots, x_n) = \lambda^n \int_{(R^2)^n} g(x_1, \ldots, x_n) \, dx \]

we obtain

\[(\lambda^2 E K_F(r))^2 = E \sum_{x_1, x_2, x_3 \in W} f(x_1, x_2) f(x_2, x_3)\]

and

\[\lambda^4 \text{Var} K_F(r) = 2E \sum_{x_1, x_2 \in W} f^2(x_1, x_2) + 4E \sum_{x_1, x_2, x_3, x_4 \in W} f(x_1, x_2) f(x_2, x_3) \]

Substituting \( u = u(x) = x - y \) yields

\[I_1 = \int_{(R^2)^2} \frac{1_{(0, \infty)}(x-y) f(x, y)}{v(F \cap F_y - y)} \, dx \, dy = \int_{R^2} \frac{1_{(0, \infty)}(u)}{v(F \cap y)} \, du \]

Analogously, we get

\[I_2 = \int_{(R^2)^2} \frac{1_{(0, \infty)}(u) 1_{(0, \infty)}(v) v(F \cap F_u 

This proves the first part of the lemma.

Taking \( u = (u_1, u_2) \) and \( v = (v_1, v_2) \) we obtain

\[v(F \cap F_u) = (a - |u_1|) (b - |u_2|)\]

and

\[1 \equiv \frac{v(F \cap F_u)}{v(F \cap F_u)} \equiv \frac{(a - |u_1|) - |u_2|}{(a - |u_1|) (b - |u_2|)} \equiv \left( 1 - \frac{r}{a - r} \right) \left( 1 - \frac{r}{b - r} \right)\]

for \( a > 2r \) and \( b > 2r \).

The latter inequality implies

\[v(b(0, r)) I_1 \left( 1 - \frac{r}{a - r} \right) \left( 1 - \frac{r}{b - r} \right) \equiv I_2 \equiv v(b(0, r)) I_1 \]

and

\[v(b(0, r)) \equiv I_1 = \int_u \int \frac{du_1 du_2}{(a - |u_1|) (b - |u_2|)} \equiv \frac{v(b(0, r))}{(a - r) (b - r)}.

The estimates (2.2) and (2.3) yield

\[\frac{2 \pi^2}{\lambda^4 ab} \left[ 1 + 2 \pi^2 \lambda \left( 1 - \frac{r}{a - r} \right) \left( 1 - \frac{r}{b - r} \right) \right] \equiv \text{Var} K_F(r) \equiv \frac{2 \pi^2 (1 + 2 \pi^2 \lambda)}{\lambda^2 \left( a - r \right) \left( b - r \right)}.

Since \( a > 2r \) and \( b > 2r \), we get immediately (2.1) and thus the lemma is proved.

The main result of this paper is the following theorem.
Theorem 1. Let $\mathcal{P}$ be a stationary Poisson process with intensity $\lambda$ and let $F=[0, a] \times [0, b]$. If $\lambda$ and $r$ are constant and $r=\frac{1}{2} \min (a, b)$, then

$$\left| \mathbb{P} \left( \frac{R_P(r) - K(r)}{\sqrt{\text{Var} R_P(r)}} < x \right) - \Phi(x) \right| \leq \frac{c_1(\lambda, r)}{\sqrt{\text{max} (a, b)} \left( 1 + |x| \right)^3}$$

holds for all real $x$. Here, $c_1(\lambda, r)$ is a constant not depending on $a, b$ and $x$.

Proof. Without loss of generality we assume $a \geq b$ and put $m=\left\lfloor \frac{a}{br} \right\rfloor$ and $n=\left\lfloor \frac{b}{2r} \right\rfloor$, where $[x]$ denotes the integer part of $x$. Let us decompose $F=[0, a] \times [0, b]$ in the following way:

$$F = \bigcup_{i=1}^{m+1} \bigcup_{j=1}^{n+1} F_{ij},$$

where

$$F_{ij} = \left[ 2 \left( i-1 \right) r, 2ir \right) \times \left[ 2 \left( j-1 \right) r, 2jr \right), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n,$$

$$F_{i+1} = \left[ 2 \left( i-1 \right) r, 2ir \right) \times \left[ 2nr, b \right], \quad 1 \leq i \leq m,$$

$$F_{m+1,j} = \left[ 2mr, a \right] \times \left[ 2 \left( j-1 \right) r, 2jr \right), \quad 1 \leq j \leq n,$$

$$F_{m+1,n+1} = \left[ 2mr, a \right] \times \left[ 2nr, b \right].$$

Elementary calculations yield

$$K_P(r) = \sum_{i=1}^{m+1} \sum_{j=1}^{n+1} X_{ij},$$

where

$$X_{ij} = \frac{1}{\lambda^2} \sum_{z \in \mathcal{P}} \frac{1}{v(E \cap E_{x-y})} \left( 1_{F_{ij}}(x) 1_{F_{ij}}(y) + 2 \left( 1_{F_{ij}}(x) 1_{F_{ij+1}}(y) + 1_{F_{ij}}(x) 1_{F_{ij+1}}(y) + 1_{F_{ij}}(x) 1_{F_{ij+1}}(y) + 1_{F_{ij}}(x) 1_{F_{ij+1}}(y) \right) \right).$$

The properties of the Poisson process ensure that the random vectors $(X_{kl})_{1 \leq k \leq i \leq m+1, 1 \leq l \leq n+1}$ and $(X_{kl})_{k \leq i \leq m+1, l \leq j \leq n+1}$ are independent, if $k=1$ or $l=1$. Therefore, the random variables

$$Y_i = \sum_{j=1}^{n+1} X_{ij}, \quad i = 1, ..., m+1,$$

are 1-dependent, i.e. the vectors $(Y_1, ..., Y_i)$ and $(Y_{i+1}, ..., Y_{m+1})$ are independent for $i=1, ..., m+1$. In order to derive (2.4) we use some results of the summa-
tion theory for sums of 1-dependent random variables. We quote the following non-uniform estimate in the central limit theorem given by the author [1]:

Let \( U_1, \ldots, U_n \) be 1-dependent random variables with \( \mathbb{E}U_k = 0 \) and \( \mathbb{E}|U_k|^3 < \infty \), \( S_n = U_1 + \ldots + U_n \). Suppose that \( \mathbb{V}ar \ S_n/n \max_{1 \leq k \leq n} \mathbb{V}ar \ U_k \equiv c_2 \). Then

\[
\mathbb{P} \left( S_n < x \left| \frac{1}{n} \sum_{k=1}^{n} \mathbb{E} U_k \right| \right) = \Phi(x) + \frac{c_2 n}{(1 + |x|)^{3/2} \mathbb{V}ar \ S_n^{3/2}}
\]

(2.6)

for every real \( x \).

A corresponding uniform estimate was given by Shergin [10].

Next, we shall estimate \( \mathbb{V}ar \ Y_i \) and \( \mathbb{E}|Y_i|^3 \), \( i = 1, \ldots, m + 1 \). The 1-dependence of \( X_{i1}, \ldots, X_{im+1} \) implies

\[
\mathbb{V}ar \ Y_i = \sum_{j=1}^{m+1} \mathbb{V}ar \ X_{ij} + 2 \sum_{j=1}^{m+1} \mathbb{E} (X_{ij} - \mathbb{E}X_{ij}) (X_{ij+1} - \mathbb{E}X_{ij+1})
\]

(2.7)

In view of (2.5) it is obvious that

\[
X_{ij} \leq \frac{1}{\lambda^2(a-r)(b-r)} \left[ \Psi^2(F_{ij}) + 2\Psi(F_{ij}) (\Psi(F_{ij+1}) + \Psi(F_{i+1,j})
\]

\[
+ \Psi(F_{i+1,j+1}) + 2\Psi(F_{i+1,j}) \Psi(F_{i,j+1}) \right].
\]

This estimate and the stationarity of \( \Psi \) yield

\[
\mathbb{E}X^k_{ij} \leq \frac{3^k}{(\lambda^2(a-r)(b-r))^k} \mathbb{E}\Psi^k(F_{11}), \quad k = 1, 2, \ldots
\]

(2.8)

Because \( \Psi(F_{11}) \) is Poisson distributed with parameter \( 4r^2\lambda \) we obtain after a short calculation

\[
\mathbb{E}\Psi^k(F_{11}) = c_4(k) \max (4r^2\lambda, (4r^2\lambda)^{2k}),
\]

(2.9)

where \( c_4(2) = 15 \) and \( c_4(3) = 203 \).

Applying the inequality \( |x_1 + x_2|^3 \leq 4 (|x_1|^3 + |x_2|^3) \) and the relation

\[
\mathbb{E}|Z_1 + \ldots + Z_n|^3 \leq c_5 n^{3/2} \max_{1 \leq k \leq n} \mathbb{E}|Z_k|^3
\]

for independent random variables \( Z_1, \ldots, Z_n \) with \( \mathbb{E}Z_k = 0 \) (see Petrov [6], p. 60), we get

\[
\mathbb{E}|Y_i - \mathbb{E}Y_i|^3 \leq 4 (\mathbb{E}|(X_{i1} - \mathbb{E}X_{i1}) + (X_{i2} - \mathbb{E}X_{i2}) + \ldots|^3
\]

(2.10)

\[
+ \mathbb{E}|(X_{i2} - \mathbb{E}X_{i2}) + (X_{i4} - \mathbb{E}X_{i4}) + \ldots|^3 \leq c_6 (n + 1)^{3/2} \max_{1 \leq j \leq m+1} \mathbb{E}|X_{ij}|^3.
\]

Using (2.7)–(2.10) we see that

\[
\mathbb{V}ar \ Y_i \leq 15 \left( \frac{9}{\lambda^2(a-r)(b-r)} \right)^2 \max (4r^2\lambda, (4r^2\lambda)^4) (n + 1)
\]

(2.11)

and

\[
\mathbb{E}|Y_i - \mathbb{E}Y_i|^3 \leq 203 c_6 \left( \frac{9}{\lambda^2(a-r)(b-r)} \right)^3 \max (4r^2\lambda, (4r^2\lambda)^6) (n + 1)^{3/2}
\]

(2.12)

for \( i = 1, \ldots, m + 1 \).
In order to apply relation (2.6) it suffices to show the uniform positivity of
\( \text{Var } K_F(r)/(m+1) \max Y_i \), where \( K_F(r) = Y_1 + \ldots + Y_{m+1} \). This is easily done
by using our lemma and (2.11).

Finally, we have
\[
\frac{\max Y_i \text{Var } Y_i}{\text{Var } (Y_1 + \ldots + Y_{m+1})} = c_f(\lambda, r) \frac{(n+1)^2 (m+1)^{1/2}}{(2m+1)^{3/2} (2n+1)^3}.
\]

This implies the assertion of Theorem 1.

**Theorem 2.** Under the assumptions of Theorem 1 the estimator is (uniformly)
strong consistent as max \((a, b) \to \infty\), i.e.
\[
\lim \sup_{m \to \infty} |K_F(r) - K(r)| = 0, \quad 0 < r_0 < \infty,
\]
with probability 1.

**Proof.** It is well-known that for an arbitrary sequence \((V_n)_{n=1}^{\infty}\) of random
variables the convergence \(V_n \to 0\) as \(n \to \infty\) with probability 1 is equivalent to
\[
P(\sup_{n \geq m} |V_n| \leq \varepsilon) \to 0, \quad n \to \infty,
\]
for every \(\varepsilon > 0\) (see Petrov [6], p. 264). Put
\[
V_m = \sum_{i=1}^{m+1} (Y_i - E Y_i) = K_F(r) - K(r) \quad \text{with} \quad m = \left[ \frac{n}{2b} \right].
\]
Then
\[
P(\sup_{n \geq m} |V_n| \leq \varepsilon) \leq \sum_{n \geq m} (1 - \Phi(\varepsilon \text{Var } V_n^{-1/2}))
\]
\[+ \sum_{n \geq m} |P(|V_n| \leq \varepsilon) - (1 - \Phi(\varepsilon \text{Var } V_n^{-1/2}) + \Phi(-\varepsilon \text{Var } V_n^{-1/2}))|.
\]
Using the inequality \(1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \) for \(x \geq 1\), the estimate \(\text{Var } V_n \leq c_0/n\)
and applying Theorem 1 we get for sufficiently large \(m\) and every \(\varepsilon > 0\)
\[
P(\sup_{n \geq m} |V_n| \leq \varepsilon) \leq c_0(\varepsilon) m^{-1/2 - \delta},
\]
where \(\delta > 0\) is arbitrary small.

This ensures that \(K_F(r) - K(r)\) as \(m \to \infty\) with probability 1 for every real \(r\).
The uniform convergence
\[
\sup_{0 \leq r \leq r_0} |K_F(r) - K(r)| \to 0 \quad \text{as} \quad m \to \infty
\]
with probability 1 can be verified in the same way as Glivenko's theorem. This
completes the proof of Theorem 2.

3. A test for randomness

We consider as a particular case the square \(F = [0, a] \times [0, a]\) and assume that \(a/r\)
is sufficiently large. The results of Section 2 permit us to consider \(K_F(r)\) as an
approximately normally distributed random variable with mean \(\pi r^2\) and variance

\[
\text{Var } K_F(r) = c_0(\varepsilon) m^{-1/2 - \delta},
\]

where \(\delta > 0\) is arbitrary small.
\[
\left( \frac{r}{\lambda a} \right)^2 \frac{2\pi}{1 + 2\pi r^2 \lambda}. \]

Let there be given a level of significance \( z \). For testing the null hypothesis of an underlying Poisson process (with given intensity) the test statistic is

\[
T(r, \lambda, a) = \frac{\lambda a (K_P(r) - \pi r^2)}{\sqrt{r(1 + 2\pi r^2 \lambda)}}.
\]

The (symmetric) critical region of size \( z \) for \( T(r, \lambda, a) \) is \((-\infty, -z_{1-z/2}) \cup (z_{1-z/2}, \infty)\). Thus, the critical region of size \( z \) for \( K_P(r) \) is \((-\infty, B_l(r, \lambda, a)) \cup (B_u(r, \lambda, a), \infty)\), where

\[
B_l(r, \lambda, a) = \pi r^2 (\pm) z_{1-z/2} \frac{r}{\lambda a} \sqrt{2\pi (1 + 2\pi r^2 \lambda)}.
\]

At the end of this section we consider an example which was discussed in Ripley [8], p. 185, Fig. 10/11). We compare the values \( B_l(r, \lambda, a) \), \( B_u(r, \lambda, a) \) (for \( z = 0.01 \)) for some \( r \) with the lower and upper envelopes \( E_l, E_u \) (measured in Fig. 11) of the plots of \( K(r) \) for 99 simulations of the Poisson process with \( \lambda = 688 \). We have \( a = 0.3 \).

<table>
<thead>
<tr>
<th>( r )</th>
<th>( B_l(r, \lambda, a) )</th>
<th>( B_u(r, \lambda, a) )</th>
<th>( E_l )</th>
<th>( E_u )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03</td>
<td>0.0008</td>
<td>0.0028</td>
<td>0.0003</td>
<td>0.0045</td>
</tr>
<tr>
<td>0.04</td>
<td>0.0015</td>
<td>0.0086</td>
<td>0.0006</td>
<td>0.0079</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0025</td>
<td>0.0132</td>
<td>0.0017</td>
<td>0.0114</td>
</tr>
<tr>
<td>0.06</td>
<td>0.0037</td>
<td>0.0190</td>
<td>0.0057</td>
<td>0.0171</td>
</tr>
<tr>
<td>0.07</td>
<td>0.0044</td>
<td>0.0264</td>
<td>0.0085</td>
<td>0.0228</td>
</tr>
<tr>
<td>0.08</td>
<td>0.0067</td>
<td>0.0335</td>
<td>0.0127</td>
<td>0.0297</td>
</tr>
<tr>
<td>0.09</td>
<td>0.0086</td>
<td>0.0423</td>
<td>0.0171</td>
<td>0.0354</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0106</td>
<td>0.0522</td>
<td>0.0228</td>
<td>0.0417</td>
</tr>
</tbody>
</table>

4. Remarks

In recent papers the asymptotic normality of some other test statistics was shown. Liebetrau [3] proved the asymptotic normality of some estimators of \( \text{Var} \, \hat{\Psi}(F) \) as \( n \to \infty \). Jolivet [2] showed (under certain restrictions) the asymptotic normality of some classes of estimators, not assuming that the underlying stationary ergodic point process is Poissonian.

The results presented in Section 2 evidently extend to higher dimensions. It seems that the following generalizations and improvements of the above results are possible: the rate of convergence in Theorem 1 is \( c_{10}(ab)^{-1/2} \), the discrepancy \( \sup_{0 \leq r \leq r_0} \left| K_P(r) - K(r) \right| \) is approximately Kolmogorov distributed, Theorems 1 and 2 remain valid for certain classes of (strong) mixing point processes. These problems are subject of a future paper.

I am indebted to my colleagues Dr. D. Stoyan, Dr. K.-H. Hanisch and Dr. J. Ohser for their helpful discussions.
References


Zusammenfassung


Резюме

Для одной оценки предварительной второй момент-меры доказывается асимптотическую нормальность, если рассматриваемый точечный процесс является Пуассоновским и площадь прямоугольного окна наблюдения стремится к бесконечности. Дисперсия оценки и скорость сходимости в центральной предельной теореме выведены. Полученные результаты применяются для проверки гипотезы, что рассматриваемая структура точек является Пуассоновским процессом.

Received April 1982.