

SUBSAMPLING A RANDOM FIELD

Antonio Possolo

ABSTRACT

Subsampling techniques developed by Hartigan (for independent observations) and by Carlstein (for mixing processes on the integers) are extended to estimate variances and covariances of statistics of spatial processes.

If the process is stationary, and dependence weakens rapidly with increasing distance, then the procedure is consistent in the sense that, as the region available for observation grows large, both the bias and the variance, of the estimator of the second moments, converge to zero.

The technique is applied to estimate the variance of the sample intensity of a binary Markov random field, and the variance of an index of clumping for spatial point processes studied by quadrat methods.

Antonio Possolo is *Senior Statistician*, The Boeing Company (Seattle, Washington), and *Affiliate Assistant Professor*, Department of Statistics, University of Washington (Seattle, Washington). This work was carried out under Office of Naval Research Contract N00014-85-K-0422. Edward Carlstein (University of North Carolina) kindly made the contents of his doctoral dissertation available to the author prior to publication.

Keywords and Phrases. Mixing, Spatial, Point Process, Quadrat, Binary, Cluster.

AMS 1980 subject classification. Primary 62M05; Secondary 62M30, 60G60, 60G55.

1. Random fields

A random field on $T \equiv \mathbf{Z}^\nu$, for some positive integer ν , is a collection \mathbf{X} of random elements $\{X_t : t \in T\}$, taking values in some measurable space χ . Let P denote its joint distribution. The random field is *uniformly mixing* at an exponential rate if there exists a positive constant α such that, if I and J are finite subsets of T , then

$$\sup |P(E|F) - P(E)| \leq |I| \exp[-\alpha d(I, J)], \quad (1.1)$$

where the supremum is over all $\{X_t : t \in I\}$ -measurable $E \subset T^\chi$, and $\{X_t : t \in J\}$ -measurable $F \subset T^\chi$ with $P(F) > 0$, $|I|$ denotes the cardinality of I , and $d(I, J) = \min\{d(s, t) : s \in I, t \in J\}$ for some metric d on T .

The property is called “uniform regularity” by Dobrushin (1968), who suggests that it is unrealistic to assume, as does Deo (1975), that the upper bound in (1.1) does not depend on the cardinality of I . The measure of dependence between events is stronger than using the supremum of $|P(EF) - P(E)P(F)|$, which can be small just because $P(F)$ is small.

Slower rates of weakening of dependence should suffice to validate subsampling methods (Central Limit theorems, etc.), but Gibbs random fields that satisfy Dobrushin’s (1968) uniqueness condition invariably display exponentially fast decay of covariances with distance (Künsch 1982, Föhlmer 1982). This is implied by uniform mixing at an exponential rate (Ibragimov and Linnik 1971, Theorem 17.2.3): if U and V are square integrable random variables, the former $\{X_t : t \in I\}$ -measurable, and the latter $\{X_t : t \in J\}$ -measurable, then

$$|\text{Cov}(U, V)| \leq 2(|I|EU^2EV^2)^{1/2} \exp[-1/2\alpha d(I, J)]. \quad (1.2)$$

In practice, a random field may be the object of primary interest, or a random field may arise as a statistic of another stochastic structure. Quadrat methods (Greig-Smith 1952) to analyze point processes can be interpreted in this manner. Suppose \mathbf{M} is a point process on \mathbf{R}^ν , and consider a measurable partition thereof into sets $\{B_t : t \in T\}$ that are all translates of one another, and such that $B_t \cap T = \{t\}$. If f is a function (not necessarily real-valued) defined on the non-negative integers, then $\{f[\mathbf{M}(B_t)] : t \in T\}$ is a random field on T .

2. Subsampling

To estimate the standard error of a real-valued statistic of \mathbf{X} , based upon a single realization observed on a finite *sampling window* $S \subset T$, subdivide S into blocks, all of the same shape and cardinality, and compute the value of the statistic on each of them. If the process satisfies (1.2), then these subsample values may be regarded as (approximately) uncorrelated replicates, and the standard error of the statistic is estimated as the standard deviation of the subsample values multiplied by the square root of the cardinality of the subsamples. Hartigan (1969) established conditions for the validity of this procedure for processes of independent random variables, and Carlstein (1984) extended it to α -mixing, stationary processes on \mathbf{Z} .

The *sphere*, of radius $0 \leq r < \infty$ and centered at $t \in T$, is the set $\{s \in T : d(s, t) \leq r\}$ for some metric d on T . $\{S_n : n = 1, 2, \dots\}$ is a sequence of finite subsets of T increasing *regularly* to $T : S_n \subset S_{n+1}; \cup_{n=1}^\infty S_n = T$; and there exist sequences

$\{I_n\}, \{J_n\}$ of spheres in T , all centered at the same point, such that $I_n \subset I_{n+1}, J_n \subset J_{n+1}, \cup_{n=1}^\infty I_n = T, \cup_{n=1}^\infty J_n = T, I_n \subset S_n \subset J_n$, and $\sup\{|J_n|/|I_n|\} < \infty$.

A *regular covering* of $\{S_n\}$ is a sequence of collections of *blocks* $\{\{I_{nj} : j \in T, I_{nj} \subset S_n\} : n = 1, 2, \dots\}$, defined, for each $j \equiv (j_1, \dots, j_\nu)$, as

$$I_{n,j} = \{(i_1, \dots, i_\nu) \in T : (j_1 - 1)m_{n,1} + 1 \leq i_1 \leq (j_1 - 1)m_{n,1} + l_{n,1}, \dots, (j_\nu - 1)m_{n,\nu} + 1 \leq i_\nu \leq (j_\nu - 1)m_{n,\nu} + l_{n,\nu}\},$$

with $l_{n,1} = \lambda_1 m_{n,1}, \dots, l_{n,\nu} = \lambda_\nu m_{n,\nu}$, for some $\lambda_1, \dots, \lambda_\nu \geq 1$, and $m_{n,1} \rightarrow \infty, \dots, m_{n,\nu} \rightarrow \infty$ all at the same rate, and the cardinality of the blocks $a_n = l_{n,1} \dots l_{n,\nu} = o(|S_n|)$, as $n \rightarrow \infty$. The number of blocks b_n contained in S_n grows to infinity with n , and the blocks all increase towards T . The diameter of the blocks $\{I_{n,j}\}$ is $\delta_n = \max\{d(s, t) : s, t \in I_{n,j}\}$. Suppose that the metric d is such that δ_n grows to infinity with n faster than $\log a_n$, and that the number of blocks less than δ_n away from each block is bounded by some $\kappa > 0$.

Example: If $S_n = \{(i_1, i_2) \in \mathbb{Z}^2 : 1 \leq i_1, i_2 \leq n\}$, then setting $\lambda_1 = \lambda_2 = 1$ defines a disjoint covering of S_n with $\kappa = 8$. Otherwise, a typical choice is $m_{n,1} = m_{n,2} = \sqrt{n}$, and $\lambda_1 = \lambda_2 = 2$, so that the covering comprises square blocks each pair of which shares $1/2a_n$ sites at most. In general, if $\lambda_1 = \dots = \lambda_\nu = \lambda$, and $m_{n,1} = \dots = m_{n,\nu}$, then two blocks share $a_n(1 - \lambda^{-1})$ at most. \square

$Q_n \equiv g(\mathbf{X}_{S_n})$ is a real-valued statistic, and $Q_{n,j} \equiv g(\mathbf{X}_{I_{n,j}})$ are its subsample values corresponding to a regular covering of S_n . The non-negative weights $\{w_n(j) : j = 1, \dots, b_n\}$ add to 1, and are such that $w_n(j) = O(b_n^{-1})$ as $n \rightarrow \infty$. The weighted average of the subsample values is $\bar{Q}_n = \sum_{j=1}^{b_n} Q_{n,j} w_n(j)$. Propositions 2.1 and 2.2 extend, to the spatial processes under consideration, Carlstein’s (1984) Theorems 3.2 and 3.3.

Proposition 2.1. *If (i) \mathbf{X} is stationary, and uniformly mixing at an exponential rate, (ii) $\lim \mathbf{E}Q_n = \theta \in \mathbf{R}$, and (iii) $\{\mathbf{E}Q_n^2\}$ are uniformly bounded, then \bar{Q}_n converges to θ in mean square, as $n \rightarrow \infty$.*

Proof. Suppose the blocks are labeled in such a manner that, for each $n, d(I_{n,1}, I_{n,2})$ is a smallest of the inter-block distances to $I_{n,1}$ that the greater than, or equal to δ_n .

$$\sqrt{\bar{Q}_n} \leq O(b_n^{-2})[b_n \mathbf{V}Q_{n,1} + 2\kappa b_n \mathbf{V}Q_{n,1} + 2b_n^2 |\text{Cov}(Q_{n,1}, Q_{n,2})|],$$

where the second summand on the right-hand side arises from using the bound $\mathbf{V}Q_{n,1}$ for the covariance between $Q_{n,1}$ and each $Q_{n,j}$ such that $d(I_{n,1}, I_{n,j}) < \delta_n$. If A_2 denotes the uniform bound for the second moments of the $\{Q_{n,j}\}$, then

$$\begin{aligned} \sqrt{\bar{Q}_n} &\leq (2\kappa + 1)A_2 b_n^{-1} + 2|\text{Cov}(Q_{n,1}, Q_{n,2})| \\ &\leq (2\kappa + 1)A_2 b_n^{-1} + 4\sqrt{a_n} A_2 \exp(-1/2\alpha\delta_n) \rightarrow 0. \quad \square \end{aligned}$$

$Y_n \equiv \phi(\mathbf{X}_{S_n})$ and $Z_n \equiv \psi(\mathbf{X}_{S_n})$ are real-valued statistics. $Y_{n,j} \equiv \phi(\mathbf{X}_{I_{n,j}})$ and $Z_{n,j} \equiv \psi(\mathbf{X}_{I_{n,j}})$ are their subsample values corresponding to a regular covering of

$\{S_n\}$. If $\{(Y_{n,j}, Z_{n,j}) : j = 1, \dots, b_n\}$ were independent, identically distributed pairs of random variables, then $1/2E[(Y_{n,j_1} - Y_{n,j_2})(Z_{n,j_1} - Z_{n,j_2})] = \text{Cov}(Y_{n,j}, Z_{n,j})$. Hence, the subsampling estimator of the covariance between $\sqrt{|S|}Y_n$, and $\sqrt{|S_n|}Z_n$ is

$$\Gamma_n = 1/2a_n \sum_{1 \leq j_1 < j_2 \leq b_n} (Y_{n,j_1} - Y_{n,j_2})(Z_{n,j_1} - Z_{n,j_2})\omega_n(j_1, j_2), \quad (2.1a)$$

for any non-negative, symmetrical, finite weights $\{\omega_n(j_1, j_2) : 1 \leq j_1 < j_2 \leq b_n\}$ adding to 1 and such that $\omega_n(j_1, j_2) = O(b_n^{-2})$, for example

$$\omega_n(j_1, j_2) = \frac{\{I_{n,j_1} \cap I_{n,j_2} = \emptyset\}}{\sum_{1 \leq k_1 < k_2 \leq b_n} \{I_{n,k_1} \cap I_{n,k_2} = \emptyset\}}. \quad (2.1b)$$

In particular, if $\phi \equiv \psi$, then Γ_n estimates the variance of $\sqrt{|S_n|}Y_n$.

Proposition 2.2. *If (i) \mathbf{X} is stationary and uniformly mixing at an exponential rate, (ii) $\lim |S_n| \text{Cov}(Y_n, Z_n) = \gamma \in \mathbf{R}$, and (iii) the fourth central moments of $\{\sqrt{|S_n|}Y_n\}$, and $\{\sqrt{|S_n|}Z_n\}$ are uniformly bounded, then Γ_n converges to γ in mean square, as $n \rightarrow \infty$.*

Proof. Define centered, $\sqrt{a_n}$ -standardized subsample valued $Y_{n,j}^* \equiv \sqrt{a_n}(Y_{n,j} - EY_{n,j})$, and $Z_{n,j}^* \equiv \sqrt{a_n}(Z_{n,j} - EZ_{n,j})$.

$$\begin{aligned} \Gamma_n &= 1/2a_n \sum_{1 \leq j_1 < j_2 \leq b_n} (Y_{n,j_1}^* - Y_{n,j_2}^*)(Z_{n,j_1}^* - Z_{n,j_2}^*)\omega_n(j_1, j_2) \\ &= \sum_{j_1=1}^{b_n} Y_{n,j_1}^* Z_{n,j_1}^* \left[\sum_{j_2=1}^{b_n} 1/2\omega_n(j_1, j_2)\{j_1 \neq j_2\} \right] \end{aligned} \quad (2.2a)$$

$$- 1/2 \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} Y_{n,j_1}^* Z_{n,j_2}^* \omega_n(j_1, j_2)\{j_1 \neq j_2\}. \quad (2.2b)$$

Since

$$\sum_{j_1=1}^{b_n} [1/2 \sum_{j_2=1}^{b_n} \omega_n(j_1, j_2)] = \sum_{1 \leq j_1 < j_2 \leq b_n} \omega_n(j_1, j_2) = 1,$$

(2.2a) is a weighted average that converges to γ in mean square by Proposition 2.1. Suppose the blocks are labeled in such a manner that, for each n , $d(I_{n,1}, J_{n,2})$ is the smallest of the inter-block distances to $I_{n,1}$ that are greater than, or equal to δ_n . The absolute value of the mean value of (2.2b) is bounded by

$$\begin{aligned} &O(b_n^{-2}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} |\text{Cov}(Y_{n,j_1}^*, Z_{n,j_2}^*)|\{j_1 \neq j_2\} \\ &\leq O(b_n^{-1})[\kappa\sqrt{\mathbf{V}Y_{n,1}^* \mathbf{V}Z_{n,1}^*} + b_n |\text{Cov}(Y_{n,1}^*, Z_{n,2}^*)|] \\ &\leq O(b_n^{-1})\kappa A_2 B_2 + O(1)\sqrt{a_n} \exp(-1/2\alpha\delta_n) \rightarrow 0, \end{aligned}$$

where A_2 and B_2 are finite constants that bound uniformly the second moments of $\{Y_{n,j}^*\}$ and $\{Z_{n,j}^*\}$. The variance of (2.2b) is bounded by

$$O(b_n^{-4}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} \mathbf{V}(Y_{n,j_1}^* Z_{n,j_2}^*) \tag{2.3a}$$

$$+O(b_n^{-4}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} \sum_{j_3=1}^{b_n} |\text{Cov}(Y_{n,j_1}^* Z_{n,j_2}^*, Y_{n,j_1}^* Z_{n,j_3}^*)| \{j_2 \neq j_3\} \tag{2.3b}$$

$$+O(b_n^{-4}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} \sum_{j_3=1}^{b_n} |\text{Cov}(Y_{n,j_1}^* Z_{n,j_2}^*, Y_{n,j_3}^* Z_{n,j_2}^*)| \{j_1 \neq j_3\} \tag{2.3c}$$

$$+O(b_n^{-4}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} \sum_{j_3=1}^{b_n} \sum_{j_4=1}^{b_n} |\text{Cov}(Y_{n,j_1}^* Z_{n,j_2}^*, Y_{n,j_3}^* Z_{n,j_4}^*)| \{j_1 \neq j_2 \neq j_3 \neq j_4\}. \tag{2.3d}$$

If A_4 and B_4 are finite constants that bound uniformly the fourth moments of $\{Y_{n,j}^*\}$ and $\{Z_{n,j}^*\}$, then (2.3a) is bounded by

$$O(b_n^{-4}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} \mathbf{E}(Y_{n,j_1}^* Z_{n,j_2}^*)^2 \leq O(b_n^{-4}) b_n^2 \sqrt{A_4 B_4} \rightarrow 0.$$

The contribution from (2.3b) is bounded by

$$O(b_n^{-4}) \sum_{j_1=1}^{b_n} \sum_{j_2=1}^{b_n} \sum_{j_3=1}^{b_n} [\mathbf{E}(Y_{n,j_1}^* Z_{n,j_2}^*)^2 \mathbf{E}(Y_{n,j_1}^* Z_{n,j_3}^*)^2]^{1/2} \\ \leq O(b_n^{-4}) b_n^3 (A_4^2 B_4^2)^{1/4} \rightarrow 0,$$

and the same applies to (2.3c).

Since the number of terms in (2.3d) is proportional to b_n^4 , bounds sharper than those used before are called for. Classify the relevant quadruplets (j_1, j_2, j_3, j_4) according to whether (a) the distance between $I_{n,j_1} \cup I_{n,j_2}$ and I_{n,j_3} is less than δ_n ; or (b) the distance is greater than, or equal to δ_n . The contribution from the former is bounded by

$$O(b_n^{-4}) b_n^3 \kappa (A_4^2 B_4^2)^{1/4};$$

the contribution from the latter is bounded by

$$O(b_n^{-4}) b_n^4 2\sqrt{a_n} (A_4^2 B_4^2)^{1/4} \exp(-1/2\alpha\delta_n) :$$

both go to zero as $n \rightarrow \infty$. The conclusion is that (2.2b) converges to zero in mean square. \square

3. Examples

3.1. Intensity of a binary random field

Verhagen's (1977) process on \mathbf{Z}^2 is a stationary, isotropic, binary Markov random field that allows a unilateral construction (Pickard (1977)), and whose distribution is determined by three parameters:

$$\theta = \Pr \{X_{(i,j)} = 1\},$$

$$\alpha = \Pr \{X_{(i,j+1)} = 1 | X_{(i,j)} = 1\},$$

$$\delta = \Pr \{X_{(i+1,j+1)} = 1 | X_{(i,j)} = 1, X_{(i,j+1)} = 1, X_{(i+1,j)} = 1\},$$

for all $(i, j) \in \mathbf{Z}^2$. The correlation between $X_{(i_1, j_1)}$ and $X_{(i_2, j_2)}$ is $\exp[(-|i_1 - i_2| + |j_1 - j_2|) \log \rho]$, with $\rho = (\alpha - \theta)/(1 - \theta)$.

N_n denotes the cardinality of $S_n \equiv \{(i, j) \in \mathbf{Z}^2: 1 \leq i, j \leq n\}$. $Y_n \equiv N_n^{-1} \sum_{s \in S_n} X_s$ is the average number of sites in S_n where the field has assumed the value 1. Partition \mathbf{Z}^2 into shells of sites around $(0, 0)$, such that the k th shell comprises $4k$ sites all at (Manhattan) distance k from $(0, 0)$. The limiting variance of $\sqrt{N_n} Y_n$ is

$$\gamma_\infty^2 = \theta(1 - \theta) \left[1 + \sum_{k=1}^{\infty} 4k\rho^k \right] = \theta(1 - \theta) [1 + 4\rho(1 - \rho)^{-2}]. \quad (3.1)$$

If $\rho > 0$, then the covariance function is positive, and it may be anticipated that Γ_n will typically be negatively biased for γ_∞^2 . In general, the larger the blocks the better the assessment of the dependence structure, hence the more accurate the subsampling estimator: a scheme with overlapping blocks tries to strike a compromise between the conflicting interests of accuracy (large blocks) and stability (many blocks), within the limits imposed by a sampling window of finite extent.

Three sampling situations are studied empirically by means of one thousand independent replicates, with $n = 100$, $N_n = 10,000$. The subsamples are square blocks of cardinality a_n either (A) $\sqrt{N_n} = 100$ (one hundred non-overlapping blocks), or (B) $2\sqrt{N_n}$ rounded to 196 (eighty-one overlapping blocks) with weights as in (2.1b). The first estimator is based upon 4,950, and the second upon 2,968, squared differences between pairs of subsample values.

The stronger the dependence (the larger the ρ), the poorer the performance of Γ_n , for a fixed block size. In each sampling situation, the larger the blocks the smaller the

| Situation | Parameters | | | ρ | γ_∞^2 |
|-----------|------------|----------|----------|--------|-------------------|
| | θ | α | δ | | |
| I | 0.25 | 0.325 | 0.5 | 0.10 | 0.28 |
| II | 0.25 | 0.5 | 0.75 | 0.33 | 0.75 |
| III | 0.25 | 0.75 | 0.975 | 0.67 | 4.69 |

| | | Bias | Variance | MSE |
|------|---------------|---------|----------|--------|
| | $\rho = 0.10$ | | | |
| IA | | -0.0124 | 0.0014 | 0.0016 |
| IB | | -0.0068 | 0.0026 | 0.0027 |
| | $\rho = 0.33$ | | | |
| IIA | | -0.157 | 0.009 | 0.0335 |
| IIB | | -0.133 | 0.016 | 0.0338 |
| | $\rho = 0.67$ | | | |
| IIIA | | -1.87 | 0.23 | 3.73 |
| IIIB | | -1.36 | 0.53 | 2.38 |

bias in absolute value, and the fewer the usable pairs of subsample values the larger the variance of Γ_n . In situations II and III the bias squared accounts for more than half of the mean squared error. Scheme B (larger, overlapping blocks, squared differences between subsample values corresponding to pairs of non-overlapping blocks only) minimizes the contribution of the bias to the mean squared error, uniformly over all sampling situations.

3.2. Clusters in spatial point processes

A stationary point process \mathbf{M} on \mathbf{R}^2 is observed on a square sampling window $A_n \equiv \{(u_1, u_2) \in \mathbf{R}^2: 1 \leq u_1, u_2 < n + 1\}$. Partition A_n into square quadrats $\{B_{(i_1, j_2)}: 1 \leq i_1, j_2 \leq n\}$, of the form $B_{(i_1, j_2)} = \{(u_1, u_2) \in \mathbf{R}^2: i_1 \leq u_1 < i_1 + 1, i_2 \leq u_2 < i_2 + 1\}$. Define $X_{(i_1, j_2)} \equiv \mathbf{M}(B_{(i_1, j_2)})$ as the number of points allocated by \mathbf{M} to the quadrat $B_{(i_1, j_2)}$. Put $N = n^2$.

David and Moore (1954) suggest $Y_N \equiv (S_N^2/\bar{X}_N) - 1$ as an index of clumping, where \bar{X}_N and S_N^2 are the average and sample variance of the counts $\{X_{(i_1, j_2)}\}$. If \mathbf{M} is Poisson, then this index should be close to zero. Y_N positive indicates that the points tend to coalesce into clusters, and Y_N negative indicates that the presence of a point at a site tends to inhibit the occurrence of other points nearby.

If the quadrat counts are independent, and the number N of quadrats is large, then

$$\begin{aligned}
 \mathbf{V}Y_N &\approx \frac{\sigma^4}{\mu^2} \left[\frac{\mathbf{V}S_N^2}{\sigma^4} - \frac{2 \text{Cov}(\bar{X}_N, S_N^2)}{\mu\sigma^2} + \frac{\mathbf{V}\bar{X}_N}{\mu^2} \right] \\
 &\approx \frac{\sigma^4}{N\mu^2} \left[\frac{\mu_4}{\sigma^4} - \frac{2\mu_3}{\mu\sigma^2} + \frac{\sigma^2}{\mu^2} - 1 \right], \tag{3.2}
 \end{aligned}$$

where $\mu, \sigma^2, \mu_3, \mu_4$ denote the mean, variance, and third and fourth central moments of the counts (Kendall & Stuart 1977, 10.17). For point processes other than Poisson the quadrat counts generally will not be independent, and (3.2) may fail to provide realistic assessments of the variability of Y_N : typically, (3.2) will underestimate $\mathbf{V}Y_N$ in cooperative (clustering) processes. However, if \mathbf{M} is such that the associated quadrat process \mathbf{X} satisfies (1.2), then $\mathbf{V}Y_N$ can be estimated via subsampling.

Neyman’s (1939) Type A contagious distribution defines a Poisson mixed Poisson process: cluster centers are outcomes of a homogeneous Poisson process with intensity $\kappa > 0$; the cardinalities of the clusters are independent, Poisson random variables with mean $\lambda > 0$; conditionally upon the location of the cluster center and the cluster

cardinality, the points in each cluster are independent with a common probability distribution Q on \mathbf{R}^2 ‘centered’ at the cluster center. If Q has support with finite diameter, then the counts corresponding to quadrats separated by a distance greater than this diameter are independent, and (2.1) is satisfied.

Three sampling situations are examined empirically: (I) Poisson process with 10 points per unit area, observed on the sampling window A_{100} (one thousand replicates); (II) Neyman’s process with $\kappa = 1$ cluster center per unit area, and $\lambda = 10$ points per cluster, observed on A_{100} (one thousand replicates); (III) Neyman’s process as in II, observed on A_{225} (five hundred replicates). In II and III the points in each cluster are distributed uniformly at random over a square, of sidelength 5, centered at the cluster center.

The subsamples are square blocks comprising a_n quadrats, and a_n is the largest perfect square not exceeding $2\sqrt{N}$. The subsampling estimator (B) uses the weights defined in (2.1b). (C) is (3.2) with the moments there estimated by the corresponding sample moments. The square root of the mean squared error of the subsampling estimator averages about twenty percent of the value of the estimand: this is comparable to the typical error of the sample variance of fifty independent, identically distributed Gaussian observations.

| <i>Situation</i> | <i>Bias</i> | <i>Variance</i> | <i>MSE</i> |
|------------------|-------------|-----------------|------------|
| IB | 0.06 | 0.14 | 0.14 |
| IC | 0.07 | 0.01 | 0.01 |
| IIB | −1.2 | 2.00 | 3.5 |
| IIC | −3.7 | 0.05 | 13.7 |
| IIIB | −0.7 | 0.82 | 1.33 |
| IIIC | −3.5 | 0.01 | 12.5 |

References

- Carlstein, E. G. (1984). *Asymptotic normality and variance estimation for a general statistic from a stationary process*. Ph.D. Thesis, Yale University.
- David, F. N. & Moore, P. G. (1954). Notes on contagious distributions in plant populations. *Annals of Botany New Series* 18:47–53.
- Deo, C. M. (1975). A functional central limit theorem for stationary random fields. *The Annals of Probability* 3:708–715.
- Dobrushin, R. L. (1968). The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of Probability and its Applications* 13:197–224.
- Föllmer, H. (1982). A covariance estimate for Gibbs measures. *Journal of Functional Analysis* 46:387–395.
- Greig-Smith, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany New Series* 16:293–316.

- Hartigan, J. A. (1969). Using subsample values as typical values. *Journal of the American Statistical Association* 64:1303–1317.
- Ibragimov, I. A. & Linnik, Y. V. (1971). *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing. Groningen.
- Kendall, M. & Stuart, A. (1977). *The Advanced Theory of Statistics*, Volume 1. MacMillan Publishing Co., Inc. New York.
- Künsch, H. (1982). Decay of correlations under Dobrushin's uniqueness condition and its applications. *Communications in Mathematical Physics* 84:207–222.
- Neyman, J. (1939). On a new class of 'contagious' distributions applicable in entomology and bacteriology. *The Annals of Mathematical Statistics* 10:35–57.
- Pickard, D. (1977). A curious binary lattice process. *Journal of Applied Probability* 16:12–24.
- Verhagen, A. M. W. (1977). A three parameter isotropic distribution of atoms and the hard-core square lattice gas. *The Journal of Chemical Physics* 67:5060–5065.