

QÜESTIÓ, vol. 22, 1, p. 175-194, 1998

MODELIZACIÓN DE LA DISTRIBUCIÓN ESPACIAL DE QUISTES EN EL ESTÓMAGO DE LA MARSOPA MEDIANTE UN PROCESO DE GIBBS*

JORGE MATEU*

FRANCISCO MONTES**

En un trabajo previo (Balaguer et al., 1992) los autores estudiaron la distribución espacial de los quistes en el estómago de la marsopa para contrastar su compatibilidad con la hipótesis de aleatoriedad espacial completa (AEC). El resultado condujo a rechazar la hipótesis señalando en la dirección de un patrón agregado, confirmando la sospecha inicial de los zoólogos.

Este trabajo completa aquel estudio en la dirección que parece más coherente a sus autores: la de conjeturar un modelo puntual espacial con interacciones dos a dos (un proceso de Strauss) y estimar sus parámetros. Previamente se lleva a cabo el imprescindible desarrollo teórico poniendo el énfasis en los distintos métodos de estimación máximo-verosímil.

Modelling the spatial distribution of cysts placed in a porpoise stomach using a Gibbs process.

Palabras clave: Procesos puntuales finitos, procesos de Gibbs, procesos de Strauss.

* Este trabajo ha sido financiado parcialmente con la ayuda de la Generalitat Valenciana GV-2221-94 y la Fundació Bancaixa P1A94-25.

* Jorge Mateu. Departamento de Matemáticas de la Universitat Jaume I. 12071 Castellón.

** Francisco Montes. Departamento de Estadística e I.O. de la Universitat de València. 46100 Burjassot. València.

– Recibido en mayo de 1996.

– Aceptado en septiembre de 1997.

1. INTRODUCCIÓN

En el estómago principal de algunos cetáceos como la *marsopa común* y el *delfín listado*, aparecen quistes producidos por unos parásitos, los *trematodos*, especie de gusanos planos que se fijan en la pared del estómago y producen los quistes. Esta investigación comenzó con una única imagen digitalizada de una sección plana del estómago de una marsopa en la que aparecían 10 quistes y se disponía, por tanto, de un patrón puntual con 10 sucesos. Los biólogos involucrados en el estudio conjeturaban una alineación y concentración de los quistes cerca del conducto que conecta el estómago principal con el pilórico. El análisis de este problema intenta dar una solución a las hipótesis iniciales planteadas (Balaguer *et al.*, 1992).

Desde el punto de vista estadístico, el problema puede plantearse como el estudio del patrón espacial que siguen n sucesos $\{s_i = (x_i, y_i) : i = 1, \dots, n\}$ situados en una región del plano. En la *figura 1*, digitalización de una imagen real, se esquematiza la situación: se observan 10 pequeños cuadrados blancos que localizan los quistes en el estómago extendido de una marsopa, que en la figura aparece como una región irregular del plano.

La hipótesis marco a contrastar es la de la *Aleatoriedad Espacial Completa* (AEC) en la que se verifican las dos condiciones siguientes:

- (i) El número de sucesos en una región X de área $S(X)$ sigue una distribución de Poisson de parámetro $\lambda S(X)$.
- (ii) Dados n sucesos en una región X , éstos se distribuyen *independiente y uniformemente* en X .

Ambas condiciones configuran lo que se denomina *proceso puntual de Poisson*.

La alternativa a la AEC no es obviamente única, pero una aproximación prudente y realista contempla como alternativas naturales aquellas en las que la presencia de algún mecanismo, inherente al fenómeno, alienta o inhibe la presencia de sucesos en el entorno de algún suceso ya existente, situaciones que se conocen con el nombre de *proceso cluster de Poisson* o *proceso hard-core* respectivamente. En (i), la constante λ es la intensidad o el número medio de sucesos por unidad de área. De acuerdo con (i), la AEC implica que la intensidad de los sucesos no varía en el plano. La condición (ii) impide la existencia de interacciones entre los sucesos. Por ejemplo, la independencia de (ii) sería violada si la existencia de un suceso en s inhibiera o primara la ocurrencia de otros sucesos en la vecindad de s .

En (Balaguer *et al.*, 1992) se contrasta la AEC mediante métodos de distancias llegando a la conclusión que dicho patrón puntual no es aleatorio (patrón de Poisson)

destacando una mayor presencia de distancias pequeñas lo que se traduce en una tendencia a la agregación y formación, por tanto, de un patrón cluster.

Teniendo en cuenta que nuestro patrón puede ser considerado como un proceso puntual en un dominio acotado en el que el número total de sucesos es finito con probabilidad 1, podemos analizarlo como un proceso puntual finito. En este sentido, los procesos de Gibbs nos sirven de marco para modelizar nuestro patrón. Este patrón envuelve una cierta dependencia local o markoviana. Los biólogos conjeturan que las interacciones entre los sucesos son un pilar fundamental para la modelización de dicho patrón. Esto nos lleva a considerar los procesos puntuales Pairwise Interaction como el modelo más adecuado en nuestro caso.

Una hipótesis fundamental en el análisis es que el patrón puntual puede ser considerado como una realización parcial de un proceso puntual estocástico (Cox *et al.*, 1980). Llamamos al conjunto de datos anterior, un **patrón puntual espacial** y nos referiremos a las localizaciones del patrón como **sucesos** para distinguirlas de los puntos arbitrarios de la región de estudio. Un patrón puntual contiene información sobre el fenómeno en estudio, es por ello que su análisis puede ser usado para describir y modelizar propiedades de dicho fenómeno (Mateu y Montes, 1994; Mateu *et al.*, 1995).

Figura 1. (a) Imagen real del estómago de la marsopa.
(b) Simulaciones bajo AEC del patrón.

2. PROCESOS PUNTUALES FINITOS: PROCESOS DE GIBBS

2.1. Procesos puntuales finitos

De manera informal, un proceso puntual es un modelo estocástico que genera localizaciones de sucesos $\{s_i\}$ en algún conjunto X . Puesto que nuestro interés se centra en los *procesos puntuales espaciales*, X será un subconjunto de \mathbb{R}^d , aunque X podría ser cualquier espacio Hausdorff localmente compacto cuya topología tiene una base numerable (Cressie, 1991).

En lo que sigue, χ denotará la σ -álgebra de Borel de $X \subset \mathbb{R}^d$ y ν la medida de Lebesgue en X , siendo $\nu(B) = |B|$ el volumen (área en \mathbb{R}^2) de B . La forma más natural de definir un *patrón puntual espacial* (una realización de un proceso puntual espacial), es a través de las localizaciones de los sucesos s_1, s_2, \dots en X . Sin embargo, a menudo es matemáticamente más conveniente definir un patrón puntual mediante *medidas de conteo* Φ en X . Para cada conjunto de Borel B , $\Phi(B)$ es el número de sucesos en B ; así $\Phi(B) \in \{0, 1, 2, \dots\} \forall B \in \chi$. Asumimos que la medida de conteo Φ es localmente finita, es decir, $\Phi(B) < \infty, \forall B_{acotado} \in \chi$. Conocer $\Phi(B) \forall B \in \chi$ es equivalente a conocer las localizaciones espaciales de todos los sucesos s_1, s_2, \dots en X .

Recuérdese que, para un proceso puntual de Poisson, condicionando al número total de sucesos en una región acotada, éstos individualmente pueden ser tratados como independientes e idénticamente distribuidos en la región. Esto sugiere un enfoque alternativo para especificar la estructura de un proceso puntual en un dominio acotado, o más generalmente, la de *cualquier proceso puntual en el que el número total de sucesos es finito con probabilidad 1*. A tales procesos se les llama *procesos puntuales finitos (PPF)*. La alternativa sugerida consiste en:

1. Especificar la distribución del número total de sucesos $N = n$,
2. Dado $N = n$, especificar la distribución conjunta de los sucesos en la región.

Tales procesos puntuales surgen de forma natural como modelos para poblaciones de animales, insectos y plantas en el campo ecológico y como modelos para procesos de partículas en Física.

Definición 1:

Las condiciones que fundamentan un proceso puntual finito son:

1. *Los sucesos se localizan en un espacio métrico completo y separable (e.m.c.s.) X , como por ejemplo $X = \mathbb{R}^d$.*

2. Se considera una distribución $\{p_n\}$ ($n = 0, 1, \dots$) que determina el número total de sucesos en la población, cumpliendo $\sum_{n=0}^{\infty} p_n = 1$.
3. Para cada entero $n \geq 1$, se tiene una distribución de probabilidad $\prod_n(\cdot)$ sobre los conjuntos de Borel de $X^n = X \times \dots \times X$ que determina la distribución conjunta de las posiciones de los sucesos del proceso dado que el número total de sucesos es n .

Las condiciones anteriores proporcionan una guía para la *simulación de los procesos puntuales finitos*. En primer lugar, generar un número aleatorio N según la distribución $\{p_n\}$ (señalar que $\Pr\{0 \leq N < \infty\} = 1$). Posteriormente, suponiendo que $N = n$ (si $n = 0$ no hay nada que hacer), generar un vector aleatorio (s_1, \dots, s_n) según la distribución $\prod_n(\cdot)$.

Cuando se habla de procesos puntuales estocásticos, se trabaja inherentemente con conjuntos de sucesos *sin orden*. Por tanto, para ser consistentes con el tratamiento de los procesos puntuales como una teoría de conjuntos sin orden, consideramos que las distribuciones $\prod_n(\cdot)$ deben asignar el mismo peso a cada una de las $n!$ permutaciones de las coordenadas de (s_1, \dots, s_n) , o en otras palabras, $\prod_n(\cdot)$ debe ser *simétrica*. Si no lo es, podemos simetrizarla de la siguiente forma. Sea (A_1, \dots, A_n) una partición de X , entonces,

$$\prod_n^{sim}(A_1 \times \dots \times A_n) = (n!)^{-1} \sum_{perm} \prod_n(A_{i_1} \times \dots \times A_{i_n})$$

en la que \sum_{perm} se calcula sobre todas las $n!$ permutaciones (i_1, \dots, i_n) de los enteros $(1, \dots, n)$.

Una formulación alternativa, más ventajosa, utiliza las *medidas de Janossy* (J_n) (no son de probabilidad) introducidas por Janossy (1950),

$$J_n(A_1 \times \dots \times A_n) = p_n \sum_{perm} \prod_n(A_{i_1} \times \dots \times A_{i_n}) = n! p_n \prod_n^{sim}(A_1 \times \dots \times A_n).$$

Una de las ventajas de estas medidas es su simple interpretación. Si, por ejemplo, $X = \mathbb{R}^d$ y $j_n(s_1, \dots, s_n)$ es la densidad de $J_n(\cdot)$ con respecto a la medida de Lebesgue en $(\mathbb{R}^d)^{(n)}$ con $s_i \neq s_j$ para $i \neq j$, entonces, $j_n(s_1, \dots, s_n) ds_1 \dots ds_n = \Pr\{\text{haya exactamente } n \text{ sucesos en el proceso, cada uno en las } n \text{ regiones infinitésimas distintas } (s_i, s_i + ds_i)\}$.

La condición normalizadora $\sum p_n = 1$ toma la forma

$$(1) \quad \sum_{n=0}^{\infty} (n!)^{-1} J_n(X^n) = 1$$

interpretando $J_0(X^0) = p_0$ y para $n \geq 1$ se tiene que

$$J_n(X^n) = p_n \sum_{perm} \prod_n(X^n) = n!p_n.$$

La siguiente proposición establece formas equivalentes de definir un PPF.

Proposición 1:

Sea X un e.m.c.s. y sea $\chi^{(n)}$ la σ -álgebra producto de X^n . Entonces los siguientes apartados son equivalentes y cada uno es suficiente para definir un proceso puntual finito en X :

1. Una distribución de probabilidad $\{p_n\}$ sobre los enteros no negativos y una familia de distribuciones de probabilidad simétricas $\prod_n^{sim}(\cdot)$ en $\chi^{(n)}$, $n \geq 1$, con la convención de que $\prod_n^{sim}(X^0) = p_0$ y el conjunto X^0 denota un punto aislado de forma que $X^0 \times X = X = X \times X^0$.
2. Una familia de medidas simétricas no negativas $J_n(\cdot)$ en $\chi^{(n)}$, $n \geq 1$, satisfaciendo una condición normalizadora (1) con $J_0(X^0) = p_0$, un escalar no negativo.
3. Una medida de probabilidad P sobre los conjuntos de Borel de la unión contable $X^\cup = X^\cup = X^0 \cup X^1 \cup X^2 \cup \dots$. En Moyal (1962) se define (X^\cup, P) como el espacio de probabilidad canónico de un proceso puntual finito.

Para una visión más completa de los PPF pueden consultarse las referencias Daley y Vere-Jones (1988) y Cressie (1991).

2.2. Procesos de Gibbs

Una clase fundamental de procesos puntuales que surgió de la Física Estadística, es la clase de los procesos de Gibbs. Éstos forman una familia natural de modelos para patrones puntuales en los que la formación del patrón es una consecuencia de las interacciones entre los sucesos.

Un proceso de Gibbs adecuado a nuestro problema y de fácil interpretación es el modelo propuesto por Strauss (1975) en el que la densidad depende sólo del número de pares de sucesos vecinos. La densidad de Janossy de un proceso de Strauss (o equivalentemente la verosimilitud $l(\cdot; \theta)$, siendo $\theta = (\beta, \gamma)$), toma la forma (Kelly y Ripley, 1976)

$$(2) \quad l(s_1, \dots, s_n; \theta) = j_n(s_1, \dots, s_n; \theta) = e^{-|X|} \alpha^{-1} \beta^n \gamma^n$$

siendo β y γ los parámetros no negativos del modelo, α una constante normalizadora, $|X|$ corresponde al área de la región donde se observan los sucesos e Y_n representa el número de pares de sucesos vecinos definido por

$$Y_n(s_1, \dots, s_n) = \sum_{1 \leq i < j \leq n} I[\|s_i - s_j\| \leq r].$$

siendo r un valor prefijado que delimita el radio de vecindad e I la función indicatriz que toma el valor 1 si se cumple la condición establecida y 0 en otro caso.

Figura 2. Simulaciones de un proceso de Strauss: 50 puntos en un cuadrado unidad con $r = 0.1$.

(a) $\gamma = 0$

(b) $\gamma = 1/2$

(c) $\gamma = 1$

El proceso de Strauss es un modelo puntual comúnmente utilizado en aplicaciones espaciales, pues dependiendo del valor del parámetro γ podemos modelizar cualquier alternativa a un proceso de Poisson. De esta forma, el caso $\gamma = 1$ se corresponde con un proceso de Poisson con intensidad β ; $\gamma < 1$ indica regularidad en el patrón; $\gamma = 0$ proporciona un proceso de inhibición simple, *proceso hard-core* que no contiene ningún suceso a distancia menor o igual que r y finalmente, $\gamma > 1$ da lugar a un *proceso de cluster o agregación*. En este último caso la constante α no es finita por lo que no podemos hablar de densidad en (2) como se señala en Kelly y Ripley (1976). En la *figura 2* se representan simulaciones de un proceso de Strauss bajo tres valores distintos del parámetro.

Una forma de encontrar un proceso cluster de Strauss es condicionando el proceso al número n de sucesos observados en X . El proceso condicional será estable. La verosimilitud (2) condicionada a n toma la forma $l_n(s_1, \dots, s_n; \gamma) = n! \gamma^n C_n(\gamma)^{-1}$ siendo $C_n(\gamma)$ la constante normalizadora.

En realidad, un proceso de Strauss es un caso particular de una clase más general de procesos llamados *procesos puntuales de Gibbs Pairwise Interaction* (Ripley, 1977) los cuales vienen caracterizados por una densidad conjunta

$$(3) \quad l(s_1, \dots, s_n; \theta) = C^{-1} \beta^n \exp \left\{ - \sum_{i=1}^n \sum_{j>i} \Phi(\|s_i - s_j\|; \gamma) \right\} / n!$$

siendo $\|\cdot\|$ la distancia euclídea, $\Phi(\cdot)$ función potencial dependiente del parámetro $\theta = (\beta, \gamma)$, β el parámetro que determina la intensidad del proceso y C la constante normalizadora. Llamamos *energía potencial total* a $U_n(s_1, \dots, s_n; \gamma) = \sum_{i=1}^n \sum_{j>i} \Phi(\|s_i - s_j\|; \gamma)$. A menudo se trabaja con $e(t) = \exp(-\Phi(t))$ llamada *función de interacción*. Se necesitan algunas restricciones en la forma de $\Phi(\cdot)$ para asegurar la finitud de la constante normalizadora C (Ripley, 1977; Baddeley y Moller, 1989).

En general, la *energía potencial total* para una configuración de n sucesos en (s_1, \dots, s_n) viene dada por

$$U_n(s_1, \dots, s_n; \gamma) = \sum_{r=1}^n \sum_{1 \leq i_1 < \dots < i_r \leq n} \Phi_r(s_{i_1}, \dots, s_{i_r}; \gamma)$$

siendo $\Phi_r(\cdot)$ la *función potencial con interacciones de orden r* .

Normalmente sólo son considerados los términos de primer y segundo orden dando lugar a las potenciales para pares de sucesos de la forma

$$U_n(s_1, \dots, s_n; \gamma) = \sum_{i=1}^n \Phi_1(s_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \Phi_2(s_i, s_j; \gamma).$$

En la práctica, $\Phi_1(\cdot)$ viene dado por la intensidad del proceso, β , y $\Phi_2(\cdot)$ suele denotarse por simplemente $\Phi(\cdot)$ (ver (3)).

Uno de los principios fundamentales de la mecánica estadística es que en un estado de equilibrio la densidad de probabilidad de una configuración particular es inversamente proporcional a la exponencial de la energía potencial. En términos de la densidad de Janossy,

$$j_n(s_1, \dots, s_n; \gamma) = C(\gamma) \exp(-\gamma U_n(s_1, \dots, s_n; \gamma))$$

siendo $C(\cdot)$ una constante de proporcionalidad y γ el parámetro de la distribución. El mayor problema de este tipo de procesos reside en el cálculo de la constante C , elegida de forma que satisface:

$$[C(\gamma)]^{-1} = \int_{X^n} \exp(-\gamma U_n(s_1, \dots, s_n; \gamma)) ds_1 \dots ds_n.$$

En el caso particular del proceso de Strauss la constante normalizadora C de (3) toma la expresión $C = \alpha(e^{-|X|n!})^{-1}$ (ver (2)) y la función potencial de (3) viene dada por

$$(4) \quad \Phi(\|s_i - s_j\|; \gamma) = \begin{cases} -\log \gamma, & \text{si } \|s_i - s_j\| \leq r \\ 0, & \text{si } \|s_i - s_j\| > r. \end{cases}$$

3. ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO

Aunque en la literatura (Cressie, 1991; Ogata-Tanemura, 1984; Penttinen, 1984 o Ripley, 1988) podemos encontrar tres métodos de estimación del modelo, máxima verosimilitud, pseudo-máxima verosimilitud y función K , en este trabajo nos concentramos ampliamente en el primero de ellos.

3.1. Estimación Máximo Verosímil

Supongamos que el proceso consiste en n sucesos con localizaciones s_1, s_2, \dots, s_n en una región acotada $X \subset \mathbb{R}^d$. La *función de verosimilitud condicionada* al número n de sucesos en A , *densidad de Janossy condicionada*, para un proceso puntual Gibbs Pairwise Interaction se obtiene a partir de (3)

$$(5) \quad l_n(s_1, \dots, s_n; \gamma) = \frac{n!}{C_n(\gamma)} \exp \left\{ - \sum_{1 \leq i < j \leq n} \Phi(\|s_i - s_j\|; \gamma) \right\},$$

siendo la constante normalizadora,

$$(6) \quad C_n(\gamma) = \int_{X^n} \exp \left\{ - \sum_{1 \leq i < j \leq n} \Phi(\|s_i - s_j\|; \gamma) \right\} ds_1 \dots ds_n.$$

El estimador máximo verosímil $\hat{\gamma}$ de γ se obtiene maximizando (5). Se requiere por tanto evaluar la constante (6), pero habitualmente no puede evaluarse o es muy complicado pudiendo recurrir a aproximaciones.

Los estimadores obtenidos sin condicionar y condicionando el modelo son muy similares (Geyer y Moller, 1994) lo que nos permite elegir el procedimiento más conveniente. Veamos algunos *métodos* basados en aproximaciones de (6) que llamaremos de *estimación máximo verosímil aproximada*.

A) Método de Ogata y Tanemura

Ogata y Tanemura (1981) propusieron un método de aproximación de (6). Supongamos que los *sucesos del proceso puntual están distribuidos de forma dispersa y poco densa* de forma que las integrales cluster de órdenes superiores al tercero no proporcionan apenas información. Entonces la constante normalizadora es aproximadamente

$$C_n(\gamma) = (|X|^n) \left\{ 1 - \frac{b(\gamma)}{|X|} \right\}^{n(n-1)/2}$$

donde

$$b(\gamma) = \frac{d\pi^{d/2}}{\Gamma(1+d/2)} \int_0^\infty t^{d-1} \{1 - \exp(-\Phi(t; \gamma))\} dt.$$

Esta aproximación sólo es válida para procesos Gibbs *estables*, es decir, con constante normalizadora finita. Para procesos inestables las integrales cluster de órdenes mayores al tercero son importantes y no pueden obviarse.

- En el caso de un **proceso de Strauss**,

$$b(\gamma) = \pi(1 - \gamma)r^2.$$

B) Método de Penttinen

Penttinen (1984) propuso un método basado en aproximaciones por Monte-Carlo mediante un algoritmo de Newton-Raphson. Supongamos que la función potencial $\Phi(\cdot)$ y la energía potencial total $U_n(\cdot)$ son dos veces diferenciables respecto de γ . Consideremos la función («score function»)

$$(7) \quad \beta(\gamma) = \partial \log \{l_n(\cdot; \gamma)\} / \partial \gamma = -\partial U_n(s_1, \dots, s_n; \gamma) / \partial \gamma - \partial \log \{C_n(\gamma)\} / \partial \gamma.$$

Entonces, la media dada por

$$\bar{B}_T(\gamma) = \frac{1}{T} \sum_{t=1}^T \beta_t(\gamma; \phi_n(t)),$$

con $\beta_t(\gamma; \phi_n(t)) = -\partial U_n(s_1, \dots, s_n; \gamma) / \partial \gamma + \partial U_n(\phi_n(t); \gamma) / \partial \gamma$, debe ser una buena aproximación a la score function (7). Sea $\Gamma(\gamma)$ la matriz cuadrada cuyo elemento (i, j) -ésimo es $\partial^2 \log\{l_n(\cdot; \gamma)\} / \partial \gamma_i \partial \gamma_j$. La matriz $\Gamma(\gamma)$ puede ser aproximada por

$$\bar{\Gamma}_T(\gamma) = \frac{1}{T} \sum_{t=1}^T [\Gamma_t(\gamma; \phi_n(t))],$$

donde $\Gamma_t(\gamma; \phi_n(t))$ tiene como elemento (i, j) -ésimo la diferencia entre

$$\partial^2 U_n(\phi_n(t); \gamma) / \partial \gamma_i \partial \gamma_j - \partial^2 U_n(s_1, \dots, s_n; \gamma) / \partial \gamma_i \partial \gamma_j$$

y el elemento (i, j) -ésimo de

$$\{\beta_t(\gamma; \phi_n(t)) - \bar{B}_T(\gamma)\} \{\beta_t(\gamma; \phi_n(t)) - \bar{B}_T(\gamma)\}'.$$

El algoritmo de Newton-Raphson basado en Monte-Carlo, que arrancará con $\hat{\gamma} = \hat{\gamma}_0$, viene dado por

$$\hat{\gamma}_{k+1} = \hat{\gamma}_k - (\bar{\Gamma}_T(\hat{\gamma}_k))^{-1} \bar{B}_T(\hat{\gamma}_k)$$

para $k = 1, 2, \dots$, los procesos $\phi_n(1), \dots, \phi_n(T)$ son procesos de Gibbs simulados con parámetro $\hat{\gamma}_k$ y (s_1, \dots, s_n) es el proceso original.

A $\{\beta_t(\cdot), t = 1, \dots, T\}$ se le llama *proceso estocástico «efficient score»* y al valor medio $\bar{B}_T(\cdot)$ *estadístico «efficient score»*. Una buena aproximación al estimador máximo verosímil es aquella con \bar{B}_T próxima al 0. Una medida de la variación de este proceso estocástico viene dada por la varianza muestral $S_{\beta}^2 = 1/T \sum_{t=1}^T [\beta_t(\gamma; \phi_n(t)) - \bar{B}_T(\gamma)]^2$.

- En el caso de un **proceso de Strauss**, se tiene que

$$U_n(\phi_n(t); \gamma) = \sum_{1 \leq i < j \leq n} \Phi(\|\cdot\|; \gamma) = -\log(\gamma^{Y_n})$$

y su derivada $\partial U_n(\cdot) / \partial \gamma = -Y_n / \gamma$. El estimador de la score function, estadístico efficient score, toma la forma $\bar{B}_T(\hat{\gamma}_k) = 1/T \sum_{t=1}^T (1/\hat{\gamma}_k) [Y_n^{emp} - Y_n(t)]$, donde *emp* hace referencia al patrón original empírico e $Y_n(t)$ al número de sucesos vecinos en el patrón simulado. Finalmente, $\partial^2 U_n(\cdot) / \partial \gamma^2 = Y_n / \gamma^2$, lo que da lugar a una forma sencilla de $\bar{\Gamma}_T(\gamma)$.

C) Método de las expansiones viriales

Otro método de aproximación de la constante normalizadora es via expansiones viriales. La aproximación viene dada por

$$n^{-1} \log(C_n) \approx (b_n/2) \int_{\mathbb{R}^2} f_{12} ds_2 + (b_n^2/4) \int_{\mathbb{R}^4} f_{12} f_{13} f_{23} ds_2 ds_3 \\ + (b_n^3/8) \int_{\mathbb{R}^6} (f_{12} f_{13} f_{14} f_{23} f_{24} f_{34} + 6f_{12} f_{13} f_{14} f_{23} f_{24} + 3f_{12} f_{14} f_{23} f_{34}) ds_2 ds_3 ds_4 \dots$$

siendo $b_n = n/|X|$ y $f_{ij} = \exp\{-\Phi(\|s_i - s_j\|; \gamma)\} - 1$ (Ripley, 1988).

Esta aproximación es válida para procesos cercanos al Poisson en el sentido de que $\Phi(t)$ tiende rápidamente a 0 con el crecimiento de t. En particular, el rango de $\Phi(t)$ debe ser mucho menor que la media de las distancias entre sucesos. En principio estas expansiones pueden usarse para cualquier orden de integral pero su complejidad aumenta considerablemente con él, lo que limita su uso.

- En el caso del **proceso de Strauss** la expansión de cuarto orden es

$$\log(C_n) \approx -\pi n(n-1)\psi r^2/(2|X|) - 0.2932516\pi^2 n(n-1)(n-2)\psi^3 r^4/(6|X|^2) \\ - \pi^3 n(n-1)(n-2)(n-3)\{-0.2743278\psi^6 + 2.1854207\psi^5 \\ - 1.3788611\psi^4\}r^6/(24|X|^3),$$

siendo $\psi = 1 - \gamma$.

4. ANÁLISIS DEL PATRÓN ESPACIAL DE LOS QUISTES EN EL ESTÓMAGO DE LA MARSOPA

Tal como ya se comentó en la introducción, tenemos que modelizar un proceso de agregación o cluster. Para ello podemos recurrir a un modelo de Strauss condicionando al número de sucesos observados, lo que nos lleva a una verosimilitud condicionada de la forma $l_n(\gamma) = n! \gamma^n C_n(\gamma)^{-1}$ con γ el parámetro a estimar.

La imagen digitalizada está formada por 512x512 pixels y la región en estudio, el estómago de la marsopa, tiene una superficie de $|X| = 56612$ pixels, lo que supone una superficie realtiva de 0.216.

La elección del parámetro de escala r viene dada por el sentido biológico de vecindad, se trata por tanto de un valor a determinar por los expertos en el problema. Al no disponer de una respuesta clara por parte de ellos, consideraremos diferentes valores de r . Una descripción del número de vecinos según el parámetro de escala viene dada en la *tabla 1*.

Tabla 1. Número de sucesos vecinos Y_n

Par. escala(r)	41	51.2	76.8	100	125	150
Y_n	5	9	23	30	38	41

Para la estimación del parámetro utilizamos el método de máxima verosimilitud aproximada. Consideramos las aproximaciones de Ogata y Tanemura (1981), el desarrollo en expansiones viriales y el método de Monte-Carlo sugerido por Penttinen (1984). Las aproximaciones a la constante normalizadora $C_n(\gamma)$ proporcionadas por los dos primeros métodos han sido obtenidas para un proceso de Strauss, por lo que maximizando la función de verosimilitud aproximada se han obtenido las estimaciones máximo-verosímiles (EMV) de la *tabla 2*.

Tabla 2. EMV bajo las aprox. de Ogata-Tanemura y Expansiones viriales

Par. escala(r)	Ogata-Tanemura	Expan. viriales
41	1.215	1.182
51.2	1.468	1.302
76.8	2.148	1.303
100	1.604	1.141
125	0.835	0.972
150	-2.04	0.796

Los resultados muestran estimaciones del parámetro superiores al 1 para aquellos parámetros de escala inferiores a $r = 125$ indicando una clara agregación en el patrón puntual. Para los valores de $r = 125$ y $r = 150$ las estimaciones no son tan reveladoras si bien ello puede ser debido a que con estos parámetros de escala prácticamente todos los sucesos son vecinos entre ellos, eliminándose así la posible exclusión de la verosimilitud por no vecindad. En cualquier caso, ambos métodos proporcionan estimaciones parecidas y coinciden con la tendencia de agregación del patrón puntual.

Las simulaciones necesarias en el método de Monte-Carlo propuesto por Penttinen (1984), han sido realizadas mediante el algoritmo de Ripley (1979). Se han considerado dos valores distintos como criterios de parada en el algoritmo de Newton-Raphson, $E = 0.01$ y $E = 0.005$ y otros dos valores distintos para T , número de iteraciones en Newton-Raphson para el cálculo de cada estimador $\hat{\gamma}$, $T = 25$ y $T = 1000$, por tratarse de valores de T extremos cuya influencia sobre los resultados que aparecen en la *tabla 3* queremos estudiar.

Tabla 3. EMV por Monte-Carlo: algoritmo de Newton-Raphson

Par. escala(r)	T = 25		T = 1000	
	E=0.01	E=0.005	E=0.01	E=0.005
41	1.236	1.236	1.259	1.259
51.2	1.352	1.361	1.388	1.399
76.8	1.461	1.461	1.472	1.472
100	1.442	1.442	1.424	1.424
125	1.546	1.546	1.580	1.580
150	1.681	1.681	1.701	1.679

Nuevamente los valores estimados son mayores que 1, confirmando la tendencia de agregación. Apenas hay diferencias entre ambos valores de E y entre los dos valores de T . Sin embargo, cabe señalar que la convergencia al EMV es mucho más rápida con $T=1000$ (sólo 3 iteraciones) que con $T=25$, para el que se necesitaron 23 iteraciones, consiguiéndose estimaciones ligeramente más eficientes para el primer valor de T . En la *tabla 4* se presenta una comparación del estadístico \bar{B}_T «efficient score» y de la variación del proceso estocástico del mismo nombre para $r = 41$, $T = 25$ y $T = 1000$ y en la *figura 3* se representan los procesos «efficient score» en las dos últimas iteraciones para ambos valores de T .

Tabla 4. Estudio con $r=41$ de la var. y estadístico effi. score para valores de T

Número de Iteración	T = 1000			T = 25			
	1	2	3	1	2	22	23
Varianza	3.44	4.38	4.12	3.24	5.41	3.63	3.59
Estadístico «effic. score»	1.219	-0.0071	0.101	0.72	0.236	-0.032	0.549

Los resultados anteriores nos dan una estimación para el parámetro que indica agregación espacial. La variación del valor del parámetro es debida a las diferencias entre unos y otros métodos que, recordemos una vez más, son métodos de aproximación. Ningún método es por sí solo concluyente y es conveniente utilizar varios de ellos como en este caso. Sin embargo, en la literatura el método de Penttinen es el más aplicado. La respuesta a los zoólogos debería ser que existe un mecanismo interno al estómago que produce una agregación espacial de los quistes en el mismo, pudiéndose predecir sus posiciones mediante un proceso cluster de Strauss. Como nota final decir que en este trabajo no se han considerado los tamaños de los quistes. Esto podría ser tratado desde el prisma de los procesos puntuales marcados.

Figura 3. Proceso estocástico «efficient score»:

(a) $T = 1000$, Iteración 2.

(b) $T = 1000$, Iteración 3.

(c) $T = 25$, Iteración 22.

(d) $T = 25$, Iteración 23.

AGRADECIMIENTOS

Los autores quieren agradecer las sugerencias y comentarios del referee que han contribuido a la mejora del presente trabajo.

REFERENCIAS

- [1] **Baddeley, A.J.** y **Moller, J.** (1989). «Nearest neighbour Markov point processes and random sets». *International Statistical Review*, **57**, **2**, 89-121.
- [2] **Balaguer, A., Mateu, J.** y **Montes, F.** (1992). «Análisis de la distribución espacial de quistes en el estómago de la marsopa». *XX Congreso de la SEIO*, Cáceres, España.
- [3] **Cox, D.R.** y **Isham, V.** (1980). *Point Processes*. Chapman and Hall, London.
- [4] **Cressie, N.A.C.** (1991). *Statistics for Spatial Data*. Wiley, New York.
- [5] **Daley, D.J.** y **Vere-Jones, D.** (1988). *Introduction to the theory of point processes*. Springer, New York.
- [6] **Geyer, C.J.** y **Moller, J.** (1994). «Simulation and likelihood inference for spatial point processes». *Scand. Journal of Statistics*. (Por aparecer)
- [7] **Janossy, L.** (1950). «On the absorption of a nucleon cascade». *Proc. R. Irish Acad. Sci. Sec. A*, **53**, 181-188.
- [8] **Kelly, F.P.** y **Ripley, B.D.** (1976). «A note on Strauss model for clustering». *Biometrika*, **63**, 357-360.
- [9] **Mateu, J.** y **Montes, F.** (1994). «Modelling pairwise interaction point processes». *9th International Workshop on Statistical Modelling*. Exeter, UK.
- [10] **Mateu, J., Usó, J.L., Montes, F.** y **Antolín, C.** (1995). «Analysis of a bivariate spatial point pattern in Forest vegetation». *Advances in Ecosystems Research*, **2**, 67-71.
- [11] **Moyal, J.E.** (1962). «The general theory of stochastic population processes». *Acta Mathematica*, **108**, 1-31.
- [12] **Ogata, Y.** y **Tanemura, M.** (1981). «Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure». *Ann. Inst. Statist. Math. B*, **33**, 315-338.
- [13] **Ogata, Y.** y **Tanemura, M.** (1984). «Likelihood analysis of spatial point patterns». *Journal of the Royal Statistical Society B*, **46**, 496-518.
- [14] **Penttinen, A.** (1984). «Modelling interaction in spatial point patterns: Parameter estimation by the maximum likelihood method». *Jyvaskyla Studies in Computer Science, Economics and Statistics*, **7**, 1-105.
- [15] **Ripley, B.D.** (1977). «Modelling spatial patterns (with discussion)». *Journal of the Royal Statistical Society, B*, **39**, 172-212.

- [16] **Ripley, B.D.** (1979). «Simulating spatial patterns: Dependent samples from a multivariate density». *Journal of the Royal Statistical Society C*, **28**, 109-112.
- [17] **Ripley, B.D.** (1988). *Statistical inference for spatial processes*. Cambridge University Press, Cambridge, UK.
- [18] **Strauss, D.J.** (1975). «A model for clustering». *Biometrika*, **62**, 467-475.

ENGLISH SUMMARY

MODELLING THE SPATIAL DISTRIBUTION OF CYSTS PLACED IN A PORPOISE STOMACH USING A GIBBS PROCESS*

JORGE MATEU**

FRANCISCO MONTES*

The Poisson process can be generalized in many directions. The approach suggested here is first to specify the distribution of the total number N of points and then, given N , to specify the joint distribution of the N points over the region. Such finite point processes arise naturally as models for populations of animals, insects and plants in the ecological field and as models for particle processes in physics.

A fundamental class of finite point processes arising in statistical physics is described by means of forces acting on and between particles. The total potential energy corresponding to a given configuration of particles is assumed to be decomposable into terms representing the interactions between the particles. Such point processes are called Gibbs processes. A particular Gibbs process, the Strauss point process, based only on the number of neighbour pairs of points, is developed.

The paper finishes up with an application to model the spatial pattern of some biological data.

Keywords: Finite point processes, interactions, potential function, Strauss process.

* This work is partly supported by the help of Generalitat Valenciana GV-2221-94 and Fundació Bancaixa P1A94-25.

* Jorge Mateu. Departamento de Matemáticas de la Universitat Jaume I. 12071 Castellón.

** Francisco Montes. Departamento de Estadística e I.O. de la Universitat de València. 46100 Burjassot. València.

– Received May 1996.

– Accepted September 1997.

The Poisson process can be generalized in many directions. In this paper it is examined the generalization in the sense that, for a Poisson process, conditional on the total number of points in a bounded region of time or space, the individual points can be treated as independently and identically distributed over the region. This prompts an alternative approach to specifying the structure of point processes in a bounded domain or, more generally, of any point process in which the total number of points is finite with probability 1. Such a process is called a *finite point process*.

The approach suggested is first to specify the distribution of the total number N of points and then, given N , to specify the joint distribution of the N points over the region. Such finite point processes arise naturally as models for populations of animals, insects and plants in the ecological field and as models for particle processes in physics. To define and simulate a finite point process, let's suppose that the following conditions hold: (a) The points are located in a complete separable metric space X , as, for example, $X = \mathbb{R}^d$. (b) A distribution $\{p_n\}$ ($n = 0, 1, \dots$) is given determining the total number of points in the population, with $\sum_{n=0}^{\infty} p_n = 1$. (c) For each integer $n \geq 1$, a probability distribution $\Pi_n(\cdot)$ is given on the Borel sets of $X^n = X \times \dots \times X$, and it determines the joint distribution of the positions of the points of the process, given that their total number is n .

A fundamental class of point processes arising in statistical physics is described by means of forces acting on and between particles. The total potential energy corresponding to a given configuration of particles is assumed to be decomposable into terms representing the interactions between the particles. Such point processes are called *Gibbs processes*. A particular Gibbs process, the *Strauss point process*, based only on the number of neighbour pairs of points (or events), is developed.

Conditional on n , the likelihood function for a Gibbs model is of the form

$$l_n(s_1, \dots, s_n; \gamma) = \frac{n!}{C_n(\gamma)} \exp \left\{ - \sum_{1 \leq i < j \leq n} \Phi(\|s_i - s_j\|; \gamma) \right\}$$

where the normalizing constant is

$$(1) \quad C_n(\gamma) = \int_{X^n} \exp \left\{ - \sum_{1 \leq i < j \leq n} \Phi(\|s_i - s_j\|; \gamma) \right\} ds_1 \cdots ds_n.$$

Maximum likelihood estimation of γ requires the evaluation of (1). Although it is not usually obtainable in closed form, approximate expressions are sometimes available. Ogata-Tanemura (1981) used the cluster-expansion method of statistical mechanics to obtain an approximation of (1). Using up to the second-order cluster integral, the normalizing constant is approximately

$$C_n(\gamma) = (|X|^n) \left\{ 1 - \frac{b(\gamma)}{|X|} \right\}^{n(n-1)/2}$$

where

$$b(\gamma) = \frac{d\pi^{d/2}}{\Gamma(1+d/2)} \int_0^\infty t^{d-1} \{1 - \exp(-\Phi(t; \gamma))\} dt.$$

Penttinen (1984) suggests a Newton-Raphson-type algorithm for solving the m.l. estimating equation

$$\hat{\gamma}_{k+1} = \hat{\gamma}_k - (\bar{\Gamma}_T(\hat{\gamma}_k))^{-1} \bar{B}_T(\hat{\gamma}_k)$$

where $\bar{B}_T(\theta)$ and $\bar{\Gamma}_T(\theta)$ are given in Section 3.1. Finally, it is studied another method based on virial expansions. The approximation consists of

$$\begin{aligned} n^{-1} \log(C_n) &\approx (b_n/2) \int_{\mathbb{R}^2} f_{12} ds_2 + (b_n^2/4) \int_{\mathbb{R}^4} f_{12} f_{13} f_{23} ds_2 ds_3 \\ &+ (b_n^3/8) \int_{\mathbb{R}^6} (f_{12} f_{13} f_{14} f_{23} f_{24} f_{34} + 6f_{12} f_{13} f_{14} f_{23} f_{24} + 3f_{12} f_{14} f_{23} f_{34}) ds_2 ds_3 ds_4 \dots \end{aligned}$$

where $b_n = n/|X|$ and $f_{ij} = \exp\{-\Phi(\|s_i - s_j\|; \gamma)\} - 1$ (Ripley, 1988).

The paper finishes up with an application of these theoretical techniques to model the spatial pattern of some biological data.