

# **Modelización estadística en Neurofisiología**

**Carmen María Cadarso Suárez**

Departamento de Estadística e IO

Universidade de Santiago de Compostela

## **1. INTRODUCCIÓN.**

La Neurofisiología estudia el comportamiento en condiciones normales (es decir, la fisiología) del sistema nervioso de los animales, desde niveles celulares a otros más complejos. La Neurofisiología utiliza diferentes abordajes técnicos, por ejemplo, la biología molecular, condicionamiento de Pavlov, o la Electrofisiología. La Electrofisiología, en particular, registra la actividad eléctrica producida por las neuronas y los estímulos sensoriales (tacto, visión, audición, etc.) se codifican en la frecuencia de producción de potenciales de acción. Estos estudios pueden realizarse en cualquier parte del sistema nervioso: corteza cerebral (visual, motora, tacto, etc.), músculo, etc., generalmente en primates subhumanos.

Las tareas de investigación en Neurofisiología son complicadas y arduas. Ello se debe, fundamentalmente, a que:

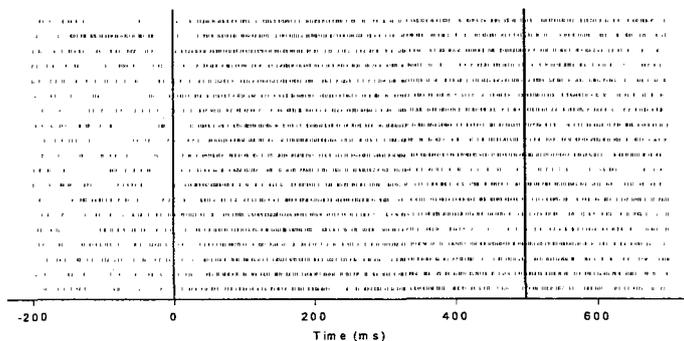
- a) Requiere el entrenamiento del animal para la tarea apropiada de conducta, lo que suele conseguirse pasados unos meses.
- b) Cada experimento diario (aún con éxito), registra un nº muy bajo de neuronas (4 a 7) ha de repetirse un número suficiente T de veces (pruebas, trials).
- c) Para obtener conclusiones científicamente válidas, ha de registrarse un gran nº de neuronas (lo que lleva días, o incluso meses) y el experimento ha de repetirse con, al menos, 2 animales.

Veamos, en primer lugar, cómo se registra la actividad eléctrica en una neurona particular.

### **1.1. Registro de la actividad eléctrica neuronal.**

Cuando la neurona se activa por un estímulo visual (p.ej. una barra) se produce una diferencia transitoria (1,5 ms) de potencial eléctrico (micro volts) entre el interior de la célula y el exterior: es el **potencial de acción**, que es captado por el microelectrodo. El sistema nervioso utiliza un código de frecuencias: las características de los estímulos (la barra) se señalan con un número de potenciales de acción separados por un intervalo. A mayor frecuencia, menor intervalo. Cada potencial de acción

se representa como un **spike**. La **Figura 1** muestra la representación de la actividad eléctrica de una neurona particular, repetida un número  $T=19$  de trials (o repeticiones del mismo experimento). En cada prueba del experimento, cada spike se representa gráficamente mediante un segmento y todos ellos forman un *raster* (cada línea horizontal de la **Figura 1**). El neurofisiólogo entiende que gráfico de todos los *rasters* es una réplica de la actividad neuronal.



**Figura 1.** Representación gráfica de la actividad neuronal (“raster plot”), basada en un total de  $T=19$  trials.

Finalmente se almacenan los intervalos (tiempos) entre potenciales de acción, y estos datos se utilizan para los análisis estadísticos posteriores.

### 1.2. Objetivos.

Tomando como base los registros de la actividad neuronal, en este trabajo nos centraremos en el estudio de la actividad eléctrica neuronal en la corteza visual en primates subhumanos, centrándonos en tres cuestiones de gran interés en Neurofisiología. En la sección 2, expondremos los distintos métodos de estimación de la tasa de disparo de una neurona particular. En la sección 3, modelaremos las posibles asociaciones entre los estímulos visuales (en forma de *barras*) y la respuesta neuronal, y por último en la sección 4, estudiaremos el concepto de sincronía neuronal, que permita cuantificar la actividad eléctrica simultánea de dos (o más) neuronas. Para la resolución de todas estas cuestiones, utilizaremos una serie de técnicas estadísticas basadas en las funciones de densidad de probabilidad y regresión.

## 2. TASA DE DISPARO NEURONAL. APROXIMACIONES ESTADÍSTICAS.

La mayor parte de los modelos de códigos neuronales se cuantifican a través de la “tasa de disparo”. Es un concepto fundamental para la descripción de la actividad temporal de una neurona (y de cualquier proceso en general, Cox e Isham, 1980). En multitud de aplicaciones fisiológicas, la tasa

de disparo indica aspectos relevantes de la implicación de las neuronas en el procesado de la información.

Desde el punto de vista matemático, la tasa de disparo,  $\lambda(t)$ , es una función del tiempo verificando:

1.  $\lambda(t) \geq 0$
2.  $\int_{t_a}^{t_b} \lambda(t) dt = E[n^\circ \text{ spikes en } [t_a, t_b]]$

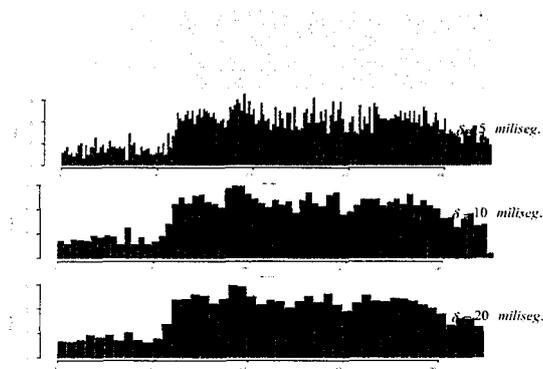
Existen diversas aproximaciones estadísticas a la función  $\lambda(t)$ . Dichas aproximaciones se basan en el proceso  $\{t_1, t_2, \dots, t_j, \dots, t_S\}$  de los tiempos en los cuales hay actividad en (ms.) a través de todos los trials. Dicho proceso puede representarse por un Proceso de Poisson (de conteo) no homogéneo, a través de la v.a.  $Y(t) = n^\circ \text{ de spikes/unidad de tiempo} = \{0, 1, 2, 3, \dots\}$ , o bien por un proceso binario a través de la v.a.  $Y(t) = \{0, 1\}$ , siendo  $Y(t)=1$  si hay spike en  $t$ , ó  $Y(t)=0$  si no hay spike en  $t$ .

### 2.1. Estimación empírica.

Tradicionalmente, la tasa de disparo  $\lambda(t)$  suele estimarse de forma empírica a través del llamado Histograma Temporal Peri-estímulo (Peri-Stimulus Time Histogram ó PSTH), propuesto por Gerstein y Kiang (1960). Este histograma se construye como sigue:

- Se divide el intervalo de tiempos en subintervalos disjuntos (bins) de precisión  $\delta$  (p.e., de 2, 5, ó 10 miliseqs.).
- Se contabiliza el  $n^\circ$  veces que ocurre un spike en cada bin.
- Se divide por el intervalo de tiempo (por ejemplo, 1 seg) y el número de trials para así obtener una tasa  $\hat{\lambda}(t)$ .

La **Figura 2** representa el *raster plot* de la neurona anterior, junto con tres estimaciones PSTH de la tasa media de disparo por segundo. Como puede observarse en la figura, el PSTH presenta un perfil más “suave” a medida que  $\delta$  aumenta.



**Figura 2.** Raster plot de la actividad eléctrica de una neurona, y diferentes aproximaciones PSTH a la tasa promedio de disparo/seg. variando la precisión ( $\delta=5, 10, 20$  miliseqs.).

## 2.2. Estimación no paramétrica a través de la función de densidad.

En general el PSTH no es un verdadero histograma, ya que no se observan datos "independientes e idénticamente distribuidos" (i.i.d.). Sin embargo, para un proceso de **Poisson no homogéneo**, condicionalmente al número total de spikes, los tiempos donde éstos se producen son i.i.d, con **función de densidad** de probabilidad :

$$f(t) = \frac{\lambda(t)}{\int \lambda(t) dt}$$

Por tanto, si a partir de un *raster plot* obtenemos un total de S spikes, en T trials, dado un estimador de la densidad,  $\hat{f}(t)$ , podríamos construir el siguiente estimador para la tasa de disparo  $\hat{\lambda}(t) = S \times \hat{f}(t)$ . Tanto en la literatura estadística (Silverman (1986)) como en la neurofisiológica (Nawrot, Arresten y Rotter (1999); Süzcs (1998)) podemos encontrar diversos métodos no paramétricos tipo núcleo para la estimación de la función de densidad.

Supongamos que se realizan T trials sobre una neurona particular, y obtenemos el proceso puntual ("agregado") de tiempos de disparo:  $\{t_1, t_2, \dots, t_j, \dots, t_S\}$ .

Se define el **suavizador tipo kernel** de la tasa de disparo como:

$$\hat{\lambda}(t) = T^{-1} h^{-1} \sum_{i=1}^S K\left(\frac{(t-t_i)}{h}\right)$$

donde  $h$  = parámetro **ventana** y  $K$  = función **núcleo** (*kernel*) verificando

$$K(t) \geq 0, \int K(t) dt = 1, \quad K(t) = K(-t)$$

Como es bien sabido en estimación paramétrica tipo núcleo, el suavizador  $\hat{\lambda}(t)$  depende de la elección de la función  $K$  pero sobretodo de la ventana  $h$ . En la actualidad, existen diversos criterios para la selección automática de este parámetro, como por ejemplo el método de *validación cruzada* (Silverman (1986)).

Las estrategias de estimación expuestas en las secciones 2.1.y 2.2., permiten promediar la actividad neuronal sobre T trials de una manera sencilla, ofreciendo un "resumen temporal" del experimento de interés. Sin embargo, estas metodologías presentan una serie de desventajas: (a) No todos los experimentos en Fisiología pueden acoplarse a un diseño de "pruebas repetidas"; (b) Promediar a través de T trials requiere como restricción la **estacionariedad** a través de los trials, lo que en general no puede garantizarse (los trials se someten a distintas condiciones específicas). Para evitar estos inconvenientes, algunos autores (Nawrot, Aersten, Rotter(1999)) recomiendan la estimación "trial-a-trial", que consiste en realizar una estimación tipo núcleo individual en cada uno de los trials. Esta estrategia, sin embargo, supone una enorme carga computacional cuando T es grande.

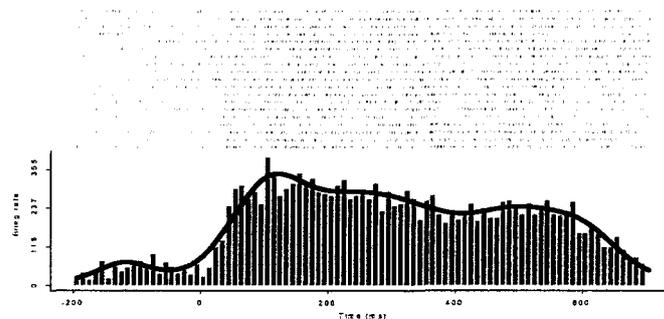
### 2.3. Modelización a través de la función de regresión.

Supone una alternativa muy sugerente y elegante a la estimación a través de la densidad, y recientemente fue propuesta por Ventura, Carta, Kass et al (2002) en el ámbito de la neurofisiología. Este abordaje produce estimaciones similares a las basadas en la densidad y permite: (a) La inclusión de covariables representando condiciones específicas de cada trial (análisis de regresión multivariante); y (b) La modelización de una población de neuronas “similares” (“population-based” estimation). Para una estimación flexible de la tasa temporal de disparo podemos utilizar el modelo lineal generalizado (GLM) de Poisson (McCullagh y Nelder (1989)) en combinación con splines de regresión naturales con  $k$  de knots ( $k_i$ )

$$\text{Ln}\lambda(t) = \beta_0 + ns(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \sum_{i=1}^k (t - k_i)_+^3$$

donde la tasa temporal de disparo  $\lambda(t)$ , representa el valor esperado de la v.a.  $Y(t) = n^\circ$  de spikes/unidad de tiempo.

La **Figura 3** representa la estimación de la tasa de disparo de una neurona producida por este método de regresión, junto con la obtenida a través del método empírico PSTH correspondiente.



**Figura 3.** Estimación de la tasa de disparo utilizando un modelo de Poisson.

### 3. ASOCIACIÓN ENTRE EL ESTIMULO VISUAL Y LA TASA DE DISPARO.

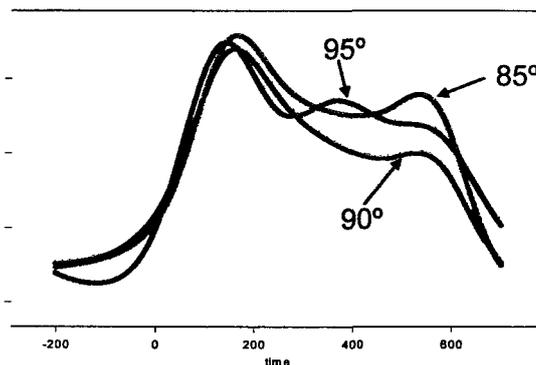
Como se ha comentado anteriormente, no todos los trials se someten a las mismas condiciones experimentales. En estudios neurofisiológicos de la corteza visual, es importante decidir por ejemplo, si la orientación del estímulo visual (en forma de barra), varía la tasa de disparo a lo largo del tiempo. Nuevamente, la función de regresión va a permitir la posible asociación entre la orientación y la tasa de disparo de esta neurona.

En la neurona que estamos estudiando (cuya actividad viene reflejada en la Figura 3), la orientación de la barra fue de 85° en 7 trials, de 90° en 7 trials, y de 95° en los 5 trials restantes. Considerando el “**tiempo**” y la “**orientación**” como covariables, y la v.a.  $Y(t)$ = nº de spikes/unidad de tiempo= {0, 1, 2, 3,...} como respuesta, el siguiente modelo jerárquico de Poisson:

$$\ln(\lambda(t; orient.)) = \beta_0 + ns(t) + orient. + ns(t) \times orient.$$

nos permite estimar la “interacción” entre dichas covariables, es decir, la evolución temporal de la tasa de disparo por cada una de las orientaciones posibles (85°, 90°, 95°).

Como se puede observar en la **Figura 4**, la evolución temporal de la tasa de disparo de la neurona depende claramente de la orientación del estímulo. Las diferencias son debidas fundamentalmente a la orientación 90°.



**Figura 4.** Estimación de la tasa de disparo temporal, según orientación (85°, 90°,95°) del estímulo visual.

#### **4. MODELIZACIÓN DE LA SINCRONÍA NEURONAL.**

Uno de los objetivos actuales de interés en la Neurofisiología del campo visual se centra en la detección (y contraste) de “**sincronía**” entre neuronas. Si existe sincronía, ésta podría explicar la percepción visual de una barra única, a pesar de las variaciones en la tasa de disparo. Desde el punto de vista fisiológico, dos neuronas poseen una “actividad eléctrica sincrónica” si ambas neuronas disparan al mismo tiempo dentro de una ventana temporal que puede estar en el rango del miliseg. La sincronía es transitoria, y puede durar muy poco tiempo (10-100 ms). Además, la sincronía puede depender de ciertas covariables.

##### **4.1. Medidas de sincronía. Razón de dependencia.**

Desde el punto de vista matemático, el concepto de sincronía temporal puede plantearse en términos probabilísticos: Elegidas dos neuronas  $i$  y  $k$ , en el mismo instante  $t$ , si  $Y_i(t)$  e  $Y_k(t)$

representan variables aleatorias binarias que toman el valor 1 si la neurona correspondiente “dispara”, y 0 en otro caso, entonces no existirá sincronía si ambas neuronas disparan “independientemente”, o lo que es equivalente, si

$$p(Y_i(t)=1 \cap Y_k(t)=1) = p(Y_i(t)=1) \times p(Y_k(t)=1) \quad , \quad \forall t$$

Esta definición matemática sugiere una nueva medida de sincronía/asincronía, que llamaremos

**Razón de dependencia:**

$$\tau_{ik}(t) = \frac{p(Y_i(t)=1 \cap Y_k(t)=1)}{p(Y_i(t)=1) \times p(Y_k(t)=1)}$$

Según esta definición, está claro que si  $\tau_{ik}(t) = 1$ , no existirá sincronía en el instante  $t$ .

#### 4.2. Modelo Plackett-Dale para la estimación de la sincronía.

El modelo de regresión Plackett-Dale (Geys, Molenberghs y Ryan (1999)) ha sido propuesto para la modelización estadística conjunta de variables respuestas bivariantes  $(Y_i, Y_k)$ . Este modelo es flexible y se acopla a cualquier tipo de variables respuesta (binarias, discretas, continuas, o mixturas), y está formado por dos estructuras marginales (una por cada componente, y una estructura de asociación entre ambas componentes. La estimación de las tres estructuras se lleva a cabo por métodos de Pseudo-Verosimilitud.

Recientemente Faes, Cadarso, Molenberghs, y Heys (2003) han propuesto una adaptación de dicho modelo al estudio de sincronía. Supongamos que el par de variables binarias  $(Y_i(t), Y_k(t))$  representan el suceso “disparo” (1=si; 0=no) para el par de neuronas  $(i, k)$ . El modelo Plackett-Dale adaptado a nuestro estudio tiene las siguientes componentes, que vienen expresadas por modelos lineales generalizados logísticos (McCullagh y Nelder (1989)):

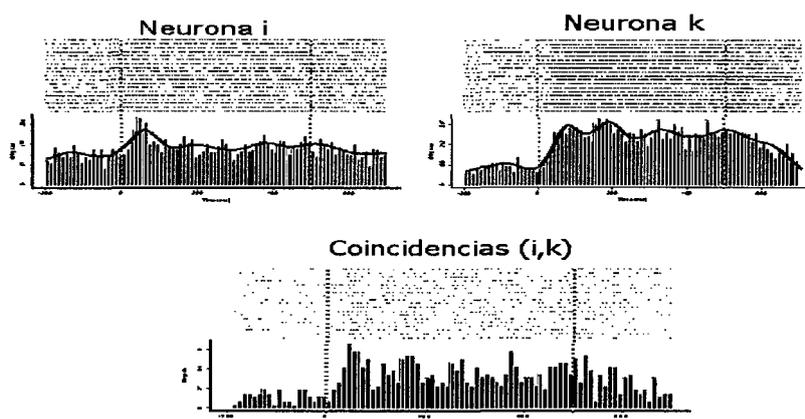
$$\left\{ \begin{array}{l} \text{Ln} \left( \frac{p[Y_i(t)=1]}{1-p[Y_i(t)=1]} \right) = \alpha_0 + ns_i(t) \\ \text{Ln} \left( \frac{p[Y_k(t)=1]}{1-p[Y_k(t)=1]} \right) = \alpha_0 + ns_k(t) \\ \text{Ln} \tau_{ik}(t) = \gamma_0 + \gamma_i + \gamma_k + ns_{ik}(t) \end{array} \right.$$

Aplicando el método de pseudo-verosimilitud, se pueden extraer simultáneamente, las estimaciones de las tasas de disparo de cada una de las neuronas, y una estimación temporal (a través de  $\tau_{ik}(t)$ ) de la sincronía entre ellas. Las posibles ventajas adicionales de esta formulación es que el

modelo permite (a) introducir de forma natural covariables representando condiciones específicas de los trials (como por ejemplo, la orientación del estímulo) y (b) una generalización al caso de sincronía entre un número  $N (>2)$  neuronas.

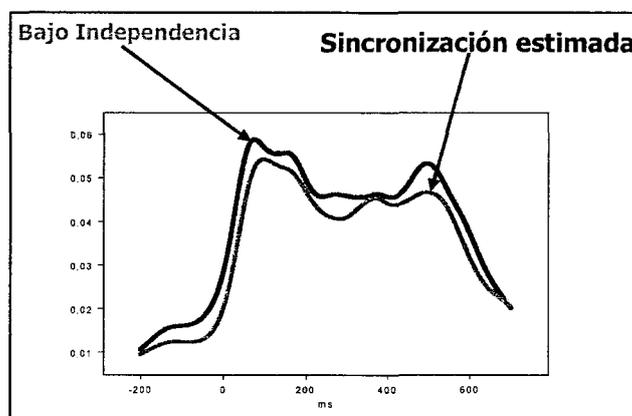
### 4.3. Aplicación a datos reales.

Como ilustración de la regresión Plackett-Dale al estudio de sincronía neuronal, ofrecemos los resultados obtenidos al aplicar dicho modelo sobre dos neuronas registradas simultáneamente en la corteza visual. En la **Figura 5**, se representan las estimaciones marginales de las tasas de disparo de las dos neuronas, junto con la tasa de “coincidencias” en el disparo.



**Figura 5.** Arriba: Estimación marginal de las tasas de disparo de dos neuronas. Abajo: Raster plot y PSTH de las coincidencias temporales de disparo de ambas neuronas.

Como puede observarse en la **Figura 6**, la curva de probabilidad temporal de disparo conjunto (sincronía) no evoluciona de la misma forma que la obtenida bajo la hipótesis de independencia en el disparo.

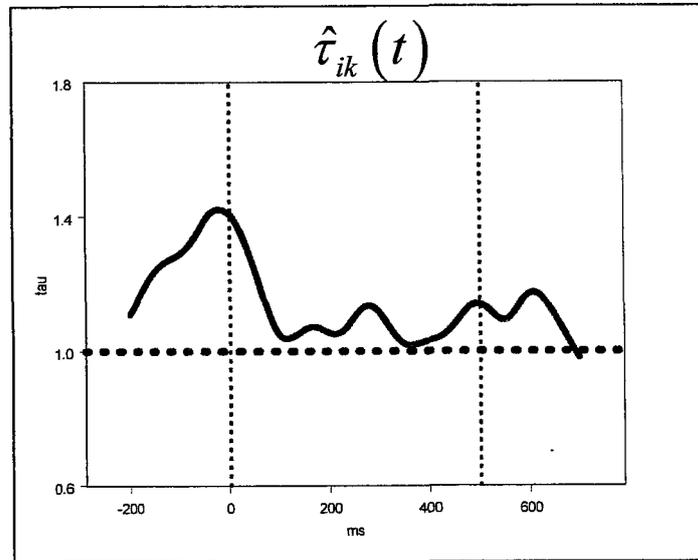


**Figura 6.** Estimación de la probabilidad de disparo de dos neuronas y de la probabilidad bajo la hipótesis de “independencia” en el disparo.

Esta situación viene cuantificada por la estimación de la razón de dependencia

$$\tau_{ik}(t) = \frac{p(Y_i(t)=1 \cap Y_k(t)=1)}{p(Y_i(t)=1) \times p(Y_k(t)=1)}$$

obtenida a través del Modelo Plackett-Dale, y que viene representada en la **Figura 7**.



**Figura 7.** Estimación de la razón de dependencia entre las neuronas  $i$  y  $k$ .

## Referencias bibliográficas

Cox DR e Isham V. (1980). *Point Processes*. London: Chapman and Hall.

Faes C, Cadarso-Suárez C, Molenberghs G y Heys H. (2003). Using the Plackett-Dale model in neuronal synchrony. Manuscrito.

Gerstein GL y Kiang NY-S. (1960). An approach to the quantitative analysis of electrophysiological data from single neurons. *Biophys J*,54,1513-28.

Geys H, Molenberghs G y Ryan L (1999). Pseudo-likelihood modelling of multivariate outcomes in developmental toxicology. *Journal of the American Statistical Association*, 94, 734-745.

McCullagh P y Nelder JA. (1989). *Generalized Linear Models*. Chapman and Hall, London.

Nawrot M, Aertsen A y Rotter S. Single-trial estimation of neuronal firing rates: from single-neuron spike trains to population activity. *J Neurosci Methods* 1999;94:81-92.

- Silverman B. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.
- Szücs A. (1998). Application of the spike density function in analysis of neuronal firing patterns. *J Neurosci Meth*, 81, 159-67.
- Ventura V, Carta R, Kass RE, Gettner SN y Olson CR. (2002). Statistical Analysis of Temporal Evolution in Single-Neuron Firing Rates. *Biostatistics*, 1, 1-20.