Clase 2

Notación del modelo paramétrico clásico

$$Y_i \overset{iid}{\sim} \text{Bernoulli}(\mu_i)$$

- Y_i y Y_j son independientes.
- Ambas tienen el mismo modelo (f. de probabilidad o densidad)

Modelo "logistico" viene del modelo

Bernoulli

$Y \sim \text{Bernoulli}(\mu)$

$$P(Y=y) = \mu^{y}(1-\mu)^{1-y}, \quad y=1,0.$$

$P(Y=1|X) =$

Suponer que Y es dada

$$Y = \begin{cases} 1, & \text{si } A \text{ ocurre.} \\ 0, & \text{c.c.} \end{cases}$$

(Y, X) un modelo conjunto. Va existir el modelo condicional

$$Y \mid X = x \quad \text{o} \quad Y \mid x$$

Vamos a suponer que

$$F(x) = \frac{e^{x}}{1-e^{x}}$$

$$\mu = P(Y=1|x)$$

Modelo logístico

$$\boxed{\mu = \frac{e^{\beta x}}{1+e^{\beta x}} = P(Y=1|x)}$$

No es único.

$$\mu = h(\beta x)$$

$$h : \mathbb{R} \longrightarrow (0,1).$$

h: logística, $\Phi(z)$, ...

Predictor lineal : βx

La función de link:

$$g(\mu) = \beta x$$

Con una muestra aleatoria

$Y_1|x_1, \ldots, Y_n|x_n$ iid.

$$Y_i|x_i \sim \text{Bernoulli}(\mu_i)$$

en que $\mu_i = \dfrac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}$

$\mu_i = f_i(\beta)$. $\beta = (1, \beta_1, \ldots, \beta_p)$

Modelo Bayesiano Logístico

Ve existir una secuencia (Y_i, X_i, β) de vectores aleatorios.

De forma analítica o de forma computacional.

$Y_1 | x_1, \beta, \ldots Y_n | x_n, \beta$

El modelo supone que

$$Y_i \mid \mu_i, \beta \sim \text{Bernoulli}$$

π

$$\beta \sim N(a, b)$$

Hiperparametros a y b.

$$x_i^T \beta.$$

$$\mu_i = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}} \text{ ; } P(\beta) \text{ un modelo } N(a, b)$$

$\beta = (1, \beta_1, \dots, \beta_p)$

Por estimativa se toma

$$\mathbb{E}(\beta_i \mid Y_i, X_i) = \beta_{0i}$$

Modelo para $\beta_i \mid Y_i, X_i$
no tiene forma cerrada

El modelo logístico formal

$\mathcal{X} = \{0,1\}^n$,

$(\mathcal{X}, \mathcal{F}, \{P_\beta : \beta \in \mathbb{R}^{p+1}\})$

Moddos conjugado

Modelo no parámetrico (funcional)

En general, cuando describo un modelo estadistico en base a una caracteristica desconocida, esta media es llamada de "parametro" o de "target".

Funcionales de un modelo $(\mathcal{X}, \mathcal{G}, P)$

$P \to F(x)$

$\to E(X) = \int x \, dF(x)$

$\to Var(X) = \iint \frac{1}{2}(x_1 - x_2)^2 \, dF(x_1) \, dF(x_2)$

$\to F(x).$

Modelo de ruido blanco Gaussián con drift

$\{Y(t) : t > 0\}$

$$\boxed{Y(t) = m(t)dt + \varepsilon dW(t)}$$

$m(t) \in L^2(\mathbb{R})$: $\int m^2(t)\,dt < \infty$

$W(t)$: P.e. de Weiner.

$E(Y(t)) = m(t)\,dt.$

Paramétrico	Funcional
θ	$m(t)$
$\theta \mid x$	
Θ	$L^2(\mathbb{R})$
$\hat{\theta}$	$\hat{m}(t)$ //

$$Y(t) = c + m(t)\,dt + \varepsilon\, dW(t),$$

c: conocido

ε: conocido

$W(t)$: es conocido

$m(t)$: es desconocida.

Presupuesto

$m(t) \in L^2(\mathbb{R})$: $\int m^2(t)\,dt < \infty$

$m^{(i)}(t)$: existe $i = 1, \ldots, r, \ldots$

$\exists \{\psi_k(t)\}$ para $L^2(\mathbb{R})$

que es una base lineal:

$\{a_k\}$

$$\boxed{m(t) = \sum_k a_k \psi_k(t).}$$

$\{\psi_k\}_k$: Spline

Fourier

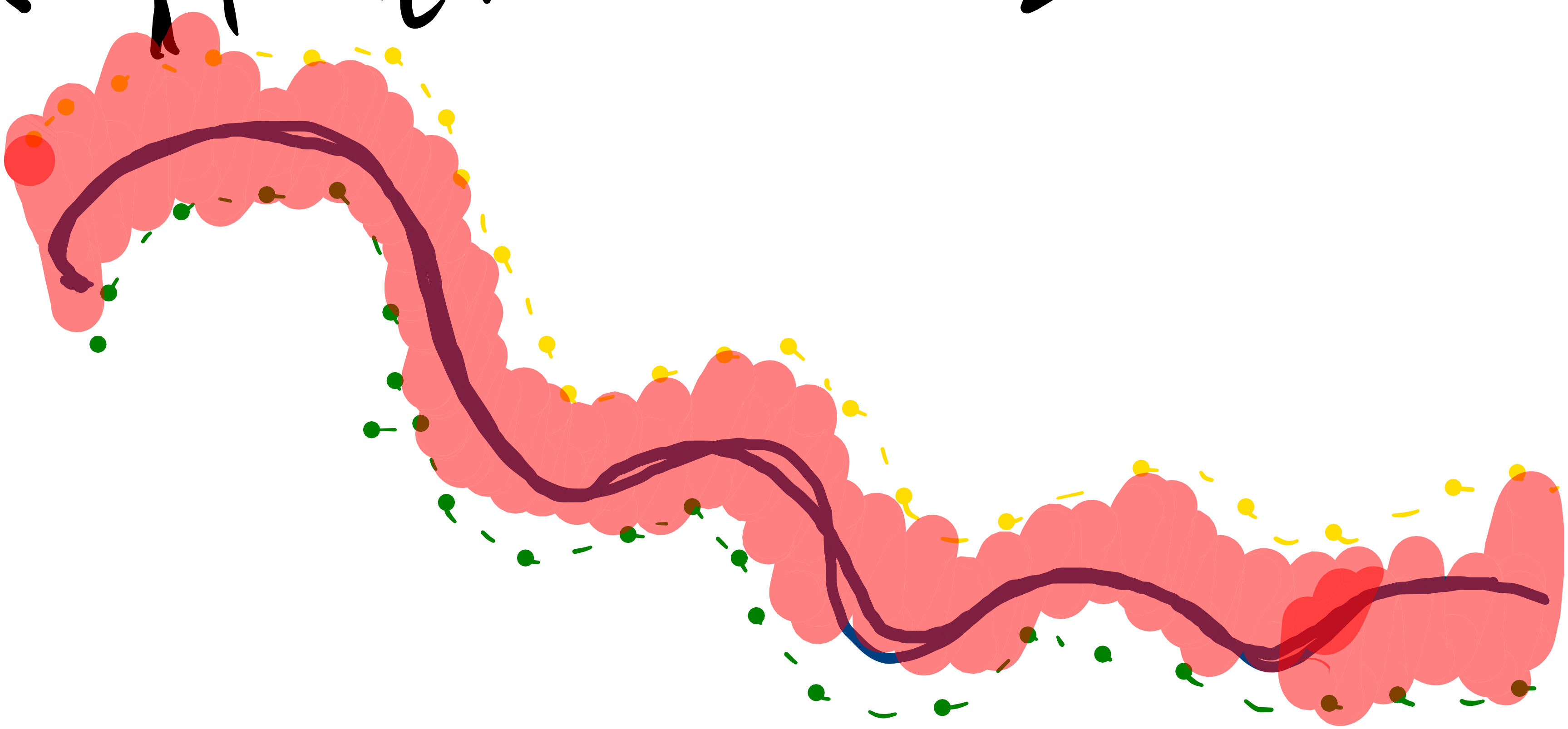
Polinomial

Wavelets

$$\hat{m}(t) = \sum_{k=1}^{50} \hat{a}_k \psi_k(t).$$

$$\hat{m}(t) = \sum_{k=1}^{s} \hat{a}_k \psi_k(t)$$

$$(\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_s)$$



$$(a_1, a_2, \ldots, a_n)$$

$$a_i \sim N(0, \sigma_i^2)$$

$$\text{"}x\text{"} = (Y(t_1), Y(t_2), \ldots, Y(t_{1020}))$$

$$a_i | x \sim \text{Posterior}$$

$$I(x,y) = m(x,y)\,dt + \varepsilon\, dW(x,y)$$

Denoising image.

Familia de modelos estadísticos

- Familia exponencial paramétrica: Poisson, Normal
- Familia Generalizada de valores extremos: Pareto, Gumbel
- Familia elíptica.
- Familia Conjugada.

Def: Un modelo $(\mathcal{X}, \mathcal{G}, P_{\Theta})$ pertenece a la familia exponencial si $f_\theta = \frac{P_\theta}{d\mu}$ se expresa como:

$$f_\theta(x) = h(x)\, e^{\langle \eta(\theta), T(x) \rangle - A(\theta)}$$

$\eta(\theta) = (\eta_1(\theta), \dots, \eta_d(\theta))$ y $T(x) = (T_1(x), \dots, T_d(x))^T$

$$\langle x, y \rangle = \sum_{i=1}^{d} x_i y_i .$$

Modelo de Poisson

$$P_\theta(x) = \frac{e^{-\theta}\cdot\theta^{x}}{x!}\,, \quad \theta \in \Theta = \mathbb{R}_+$$

$$= \frac{1}{x!}\exp\{-\theta + \log\theta^{x}\}$$

$$= \frac{1}{x!}\exp\{x\log\theta - \theta\}\,, \quad \phi = 1$$

$h(x) = \frac{1}{x!}$; $\eta(\theta) = \log\theta$; $A(\theta) = \theta$ //

Modelo Normal $\mu \in \mathbb{R}_+$ y $\sigma^2 = 1$

$$f_\mu(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(x-\mu)^2}{2} \right\}$$

"$\theta = \mu$"

$$= \dots$$

$$= \underbrace{\frac{e^{-x^2/2}}{\sqrt{2\pi}}}_{h(x)} \exp\left\{ x\underbrace{\mu}_{\eta(\mu)=\mu} - \underbrace{\frac{\mu^2}{2}}_{A(\mu)} \right\}, \quad d = 1$$

Estadístico $(\mathcal{X}, \mathcal{G}, \mathcal{P})$

$$T : \mathcal{X} \longrightarrow A$$

$A \subseteq \mathbb{R}, \mathbb{R}^n$

- Información sobre θ (Suficiente)
- sin Inf. sobre θ, que mejora la precisión. (Ancilares)

Informalmente, un estadístico T es suficiente para P_θ (para θ) si T contiene toda la información de x acerca de θ.

Un estadístico $A(x)$ es ancilar si su medida inducida P_A no depende de θ.

Sir Ronald Fisher.

Def. Un estadístico $T: \mathcal{X} \to A$ es suficiente para P_θ si

(i) $P_{\theta}(x \mid T(x) = t) = P(x \mid T(x) = t)$

(ii) $P_\theta(X = x \mid T(x) = t)$ no depende de θ

(iii) $f_\theta(x \mid T(x) = t)$

Def: T' es suficiente minimal sii es función de toda estadística suficiente.

$\forall T$ suf $\exists h: T' = h(T)$.

Minimal   suf.

$$\sigma(T') \subseteq \sigma(T)$$

Def.: $A: \mathcal{X} \to B$ es ancilar para cada P_θ en $(\mathcal{X}, \mathcal{G}, \mathcal{P}_\Theta)$ si

$$P_{\theta A} = P_A \text{ no depende de } \theta.$$

Caso Bernoulli

$$P(X=x \mid T(x)=t) = \frac{1}{\binom{n}{t}} \quad \text{no depende de } \theta$$

$\downarrow$

modelo uniforme

$(x_1, \ldots, x_n)$

t: nº de 1's.

$$T(x) = \sum_{i=1}^{n} X_i$$