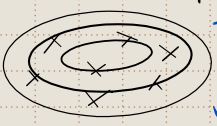


## Metodos profundidad

- \*  $X_1, X_2, \dots, X_n \sim F_X \in \mathbb{R}^d$  vs  $H_0: F_X = F_0$  vs  $H_1: F_X \neq F_0$ 
  - \* Motivación Kolmogorov-Smirnov (util dim 2)
  - \* función característica empírica (cara computacional)
  - \* pruebas basadas en profundidad.
- \* profundidad  $\Rightarrow$   medida centralidad respecto a nube de datos
  - \* Medidas profundidad caracterizan (F)
    - profundidad medio espacio Tukey.
    - profundidad simplicial.
- \* profundidad semiespacio Tukey: determina forma única la función distribución continua soporte compacto.
- \* profundidad Tukey y zonoides convergen forma casi segura y uniforme a sus versiones poblacionales

## Problema una muestra (profundidad Tukey y zonoides)

\* semiespacio Tukey (D1):  $HD_F(X) = \inf \{ D_F(H) : H \text{ semiespacio cerrado } \mathbb{R}^d \text{ tal que } H \ni X \}$

\* profundidad zonoid (D2):  $D_\alpha(F) = \left\{ \int_{\mathbb{R}^d} x g(x) dF(x) : g(x) : \mathbb{R}^d \rightarrow [0, \frac{1}{\alpha}] \text{ y } \int_{\mathbb{R}^d} g(x) dF(x) = 1 \right\}$

$$z_{D_F}(X) = \begin{cases} 0 & \text{otc} \\ \sup \{ \alpha : y \in D_\alpha(F) \} & \text{si no vacío} \end{cases}$$

\* (la versión muestral de estas profundidades se tiene reemplazando  $F$  por  $F_n(X)$  empírica.

## Carta control Ewma (Ejemplos)

- \* carta Ewma general:  $z_T = \lambda X_T + (1-\lambda)z_{T-1}$   $\lambda \in [0,1]$  constante suavizamiento
  - $\lambda = 1$  solo info reciente
  - $\lambda \approx 0$  info reciente poco peso
- \*  $\lambda$  caracteriza profundidad de la memoria del ewma (rata decaimiento pesos)

$$z_T = \lambda \sum_{j=0}^{T-1} (1-\lambda)^j X_{T-j} + (1-\lambda)^T z_0$$

## Ejemplo Ewma:

$$T_j = f(x_i, y_j) \quad \begin{matrix} i=1 \dots n \\ j=1 \dots m \end{matrix}$$

$$= \sum_{i=1}^m \left( \frac{|r_{ij} - (m+n+1)/2|}{\sqrt{mn(m+n+1)/2}} \right)$$

$$T = \sum T_j^2$$

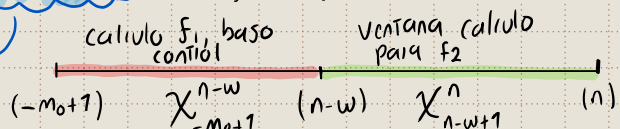
\* calcular distancias profundidad que pueda caracterizar  $F(\cdot)$

\* usamos prueba banded ajuste con el estadístico solucionado

\* construcción esquema Ewma

Estadístico

$$X_{a,b}^b = \{X_{j,a} \dots X_{j,b}\}$$



$$T_j = \sum_{i=n-w+1}^n \left( \frac{|r_{ij} - (m_0+n+1)/2|}{\sqrt{w(m_0+n-w)(m_0+n+1)/2}} \right)$$

Ewma

$$T_j = \sum_{i=n-w+1}^n (1-\lambda)^{n-i} \left( \frac{|r_{ij} - (m_0+n+1)/2|}{\sqrt{w(m_0+n-w)(m_0+n+1)/2}} \right)$$

# Bondad Ajuste

\*  $S_n = \{X_1 \dots X_n\} \stackrel{iid}{\sim} F(\tau)$   $\begin{cases} H_0: F(\tau) = F_0(\tau) \quad \forall \tau \in D_F \rightarrow D_F: \text{dominio de } F(\tau) \\ H_a: F(\tau) \neq F_0(\tau) \text{ para algun } \tau \in D_F \end{cases}$

$$\hat{F}_n(\tau) = \frac{1}{n} \sum I(X_i \leq \tau)$$

\* Zhang 2002 Gof basados en la razon de verosimilitud no parametrica

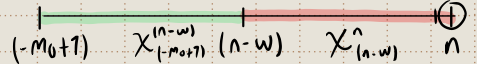
$$H_0: p \equiv P(X \leq \tau) = F_0(\tau) \quad L(\tau) = n \left\{ \hat{F}_n(\tau) \ln \left( \frac{\hat{F}_n(\tau)}{F_0(\tau)} \right) + (1 - \hat{F}_n(\tau)) \ln \left( \frac{1 - \hat{F}_n(\tau)}{1 - F_0(\tau)} \right) \right\}$$

(Cuando  $H_0$  falsa  $L(\tau)$  grande  $Z = \int_{-\infty}^{\infty} L(\tau) dW(\tau) = \int_{-\infty}^{\infty} L(\tau) [\hat{F}_n(\tau)(1 - \hat{F}_n(\tau))]^{-1} d\hat{F}_n(\tau)$   
 No depende de  $F_0(\cdot)$  estadística no parametrica y robusta.

## Carta control propuesta

\* Control datos funcionales híbridos, observaciones linea individuales, cambios distribución, carta suma, límite de distribución.

\* ESTRUCTURA SISTEMA:



\* llega una observacion n y creamos una ventana tamaño w hacia atras

\* El proceso está en control  $X_{(-m_0+1)}^{(n-w)}$  y comparamos con  $X_{(n-w)}^{(n-w)}$  ¿sigue en control?

\*  $H_0: F(\tau) = F_0(\tau) \quad \forall \tau$  vs  $H_a: F(\tau) \neq F_0(\tau)$  algun  $\tau$

\* Encontramos componentes principales y los puntuajes FMACP híbrido  $(\hat{\mu}_{ij})_{i=1, \dots, n}^{j=1, \dots, K}$   
 K componentes, cada observación es  $\hat{\mu}_i = [\hat{\mu}_{i1} \dots \hat{\mu}_{iK}]^T$ .

\* pasamos observaciones híbridas a vectoriales, luego usamos medida profundidad que guarda la distribución de probabilidad (profundidad Tukey) con ella pasamos una dimensión.

\* Tenemos observaciones  $D_N(\hat{\mu}_1), D_N(\hat{\mu}_2) \dots, D_N(\hat{\mu}_n)$  una dimensión.

$$\begin{aligned} \hat{F}_0(\tau) &= (m_0 + n - w)^{-1} \sum_{i=-m_0+1}^{n-w} I(D_N(\hat{\mu}_i) \leq \tau) \\ \hat{F}(\tau) &= (w \alpha_\lambda)^{-1} \sum_{i=n-w}^n (1-\lambda)^{n-i} I(D_N(\hat{\mu}_i) \leq \tau) \\ \alpha_\lambda &= \sum_{i=n-w}^n (1-\lambda)^{n-i} = \frac{1}{\lambda} [1 - (1-\lambda)^w] \end{aligned}$$

\* usamos algun Test de Bondad de ajuste para comparar ambas distribuciones, en nuestro caso serán prueba razon de verosimilitud no parametrica. (implimentada R)

\* mediante metodos permutación para un  $\alpha$  y  $\tau_0$  dados, encontramos límites carta cada paso

\* Evaluamos cada paso si está en control, y mueve la ventana (w) añadiendo calculos para  $\hat{F}_0(\cdot)$

## Contenido

Fecha:    /    /

## Contenido

Fecha:    /    /