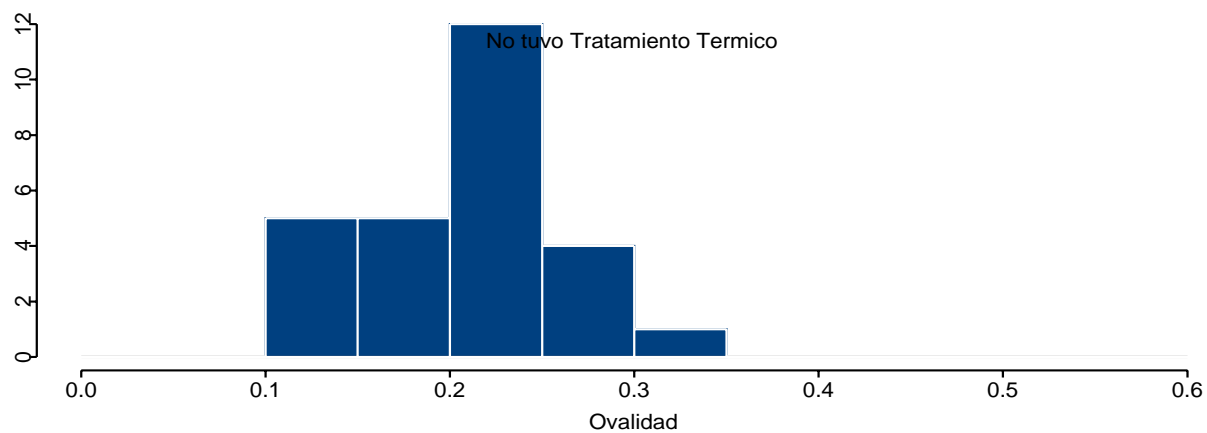
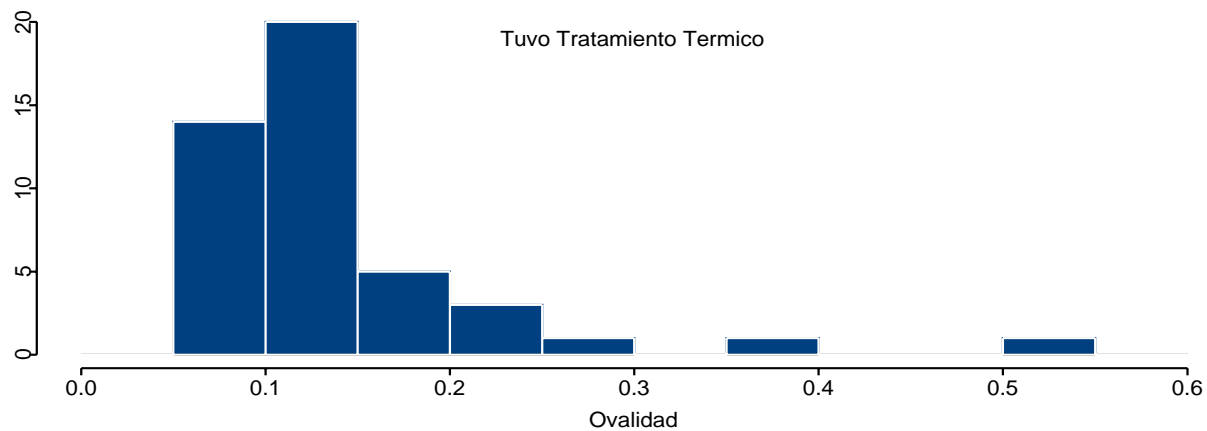
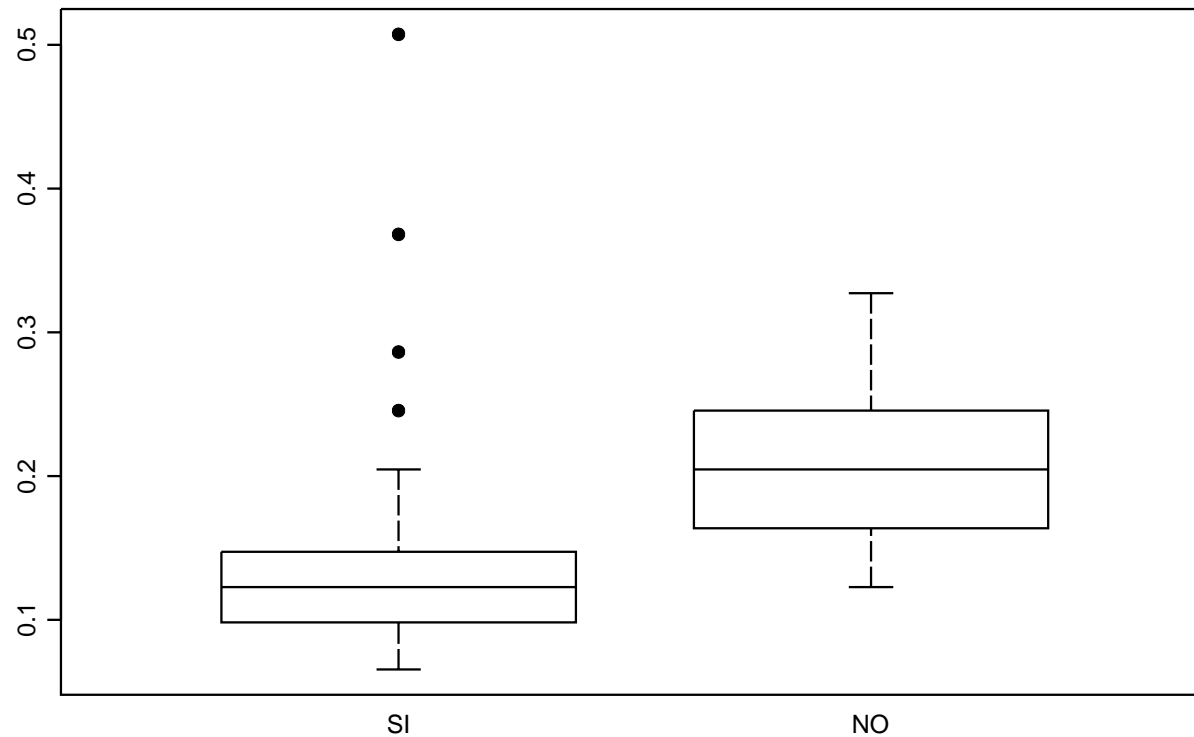


Ejemplo 1: Ovalidad de Tubos

Histogramas de Ovalidades



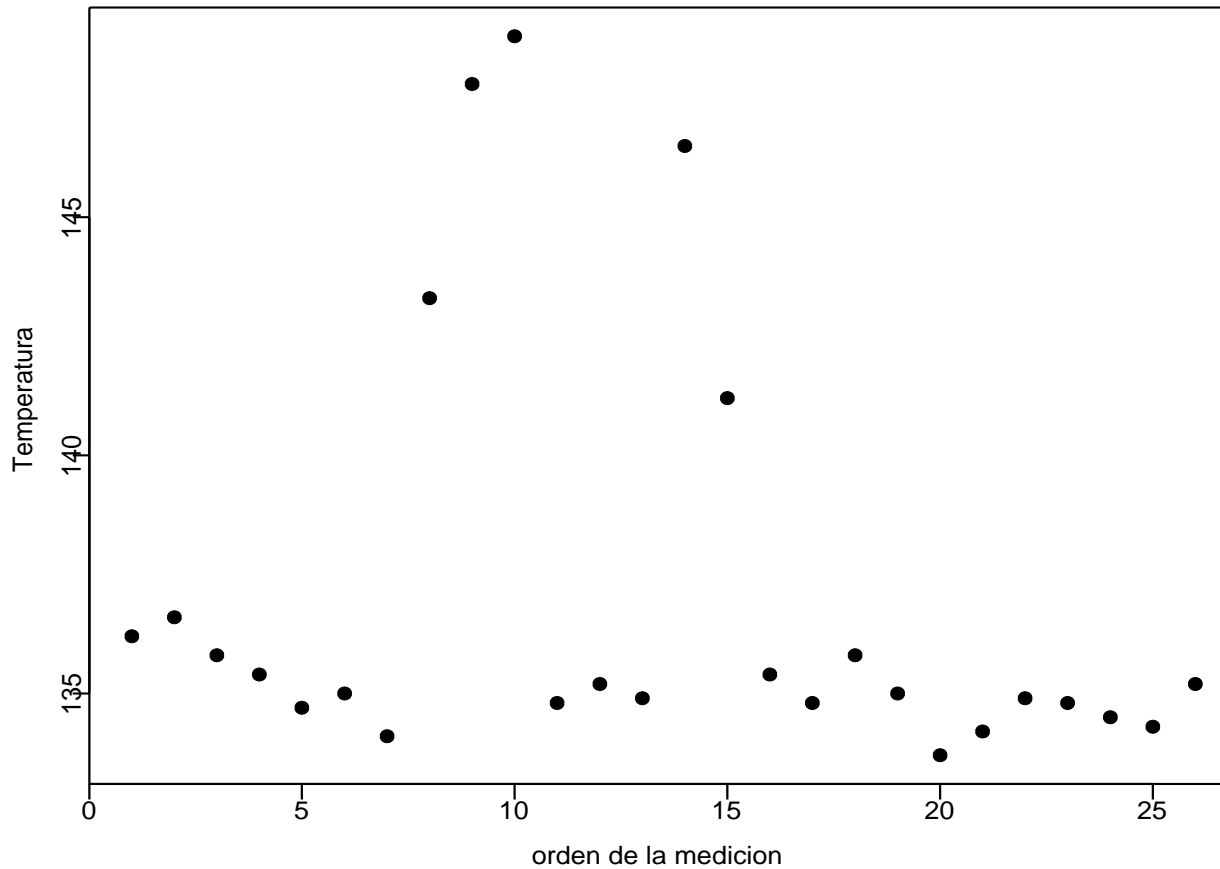


- $\bar{x} = 0.1425$ para los 45 tubos sometidos a tratamiento térmico
- Mediana= 0.1227
- Si eliminamos los 4 datos atípicos $\bar{x} = 0.1161$.

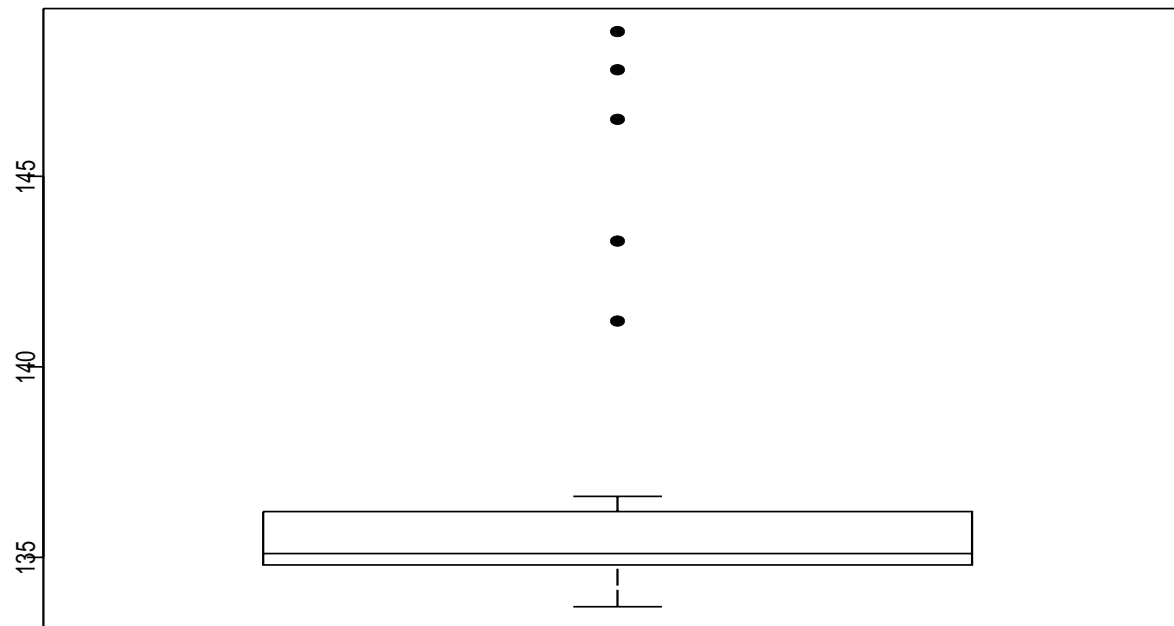
Ejemplo 2: Temperatura de Sublimación del Platino

136.2	136.6	135.8	135.4	134.7	135.0	134.1	143.3
147.8	148.8	134.8	135.2	134.9	146.5	141.2	135.4
134.8	135.8	135.0	133.7	134.2	134.9	134.8	134.5
134.3	135.2						

Ejemplo 2: Temperatura de Sublimación del Platino



Ejemplo 2: Temperatura de Sublimación del Platino



Con todos los datos: $\bar{x} = 137.05$

Sacando los datos anómalos: $\bar{x} = 134.9$.

Mediana $M = 135.1$, $d_I = 1.7$.

Media α -podada

- Ordenamos los datos,
- Descartamos el $100\alpha\%$ de los datos más pequeños y el $100\alpha\%$ de los más grandes.
- Tomamos la media de los restantes.

$$\bar{x}_\alpha = \frac{x_{([n\alpha]+1)} + \dots + x_{(n-[n\alpha])}}{n - 2[n\alpha]}$$

La Mediana puede ser vista como una 50% media podada.

Ejemplo 2. La 20% media podada correspondiente a los datos de *temperatura de sublimación de platino* $\bar{x}_{0.2} = 135.29$

El Problema de Robustez

- Modelo de posición y escala: $x_i = \mu + \sigma u_i$, $u_i \sim F$ i.i.d.
- $x_i \sim F_{\mu\sigma}(x) = F\left(\frac{x-\mu}{\sigma}\right)$
- $F = \Phi \implies x_i \sim N(\mu, \sigma^2)$.



- \bar{x} es IMVU y minimax.
- Un *entorno de contaminación* de tamaño ϵ de la distribución Φ se define por

$$\mathcal{V}_\epsilon = \{F : F = (1 - \epsilon)\Phi + \epsilon H \text{ con } H \text{ arbitraria}\}.$$

El Problema de Robustez

- Si x_1, \dots, x_n i.i.d. $x_i \sim F \in V_\epsilon \implies$
 - a) Una proporción $(1 - \epsilon)$ de las observaciones estarán dadas por $x_i = \mu + \sigma u_i$ con $u_i \sim \Phi$,
 - b) Una proporción ϵ estarán dadas por $x_i = \mu + \sigma u_i$ con $u_i \sim H$. Estas últimas observaciones serán denominadas **puntos atípicos o outliers**, .
- Si $E_H(u) = 0$ entonces

$$\text{var}_F(\bar{x}) = \frac{\sigma^2}{n} \quad \mathbf{var}_F(u) = \frac{\sigma^2}{n} ((1 - \epsilon) + \epsilon \mathbf{var}_H(u))$$

El Problema de Robustez: $\text{var}_F(\bar{x})$

$$H \sim N(0, \tau^2)$$

ϵ	0	0.1	0.2	0.3	0.4
$\tau = 2$	1.0	1.3	1.6	1.9	2.2
$\tau = 5$	1.0	3.4	5.8	8.2	10.6
$\tau = 10$	1.0	10.9	20.8	30.7	40.6
$\tau = 20$	1.0	40.9	80.8	120.7	160.6

El Problema de Robustez: $\text{var}_F(\tilde{\mu}_n)$

$$\tilde{\mu}_n = \text{median}(x_i)$$

$$\text{var}_F(\tilde{\mu}_n) = \frac{\sigma^2}{4f^2(0)}$$

$$H \sim N(0, \tau^2)$$

ϵ	0	0.1	0.2	0.3	0.4
$\tau = 2$	1.570796	1.762224	1.990901	2.267161	2.605179
$\tau = 5$	1.570796	1.857455	2.230389	2.728045	3.412934
$\tau = 10$	1.570796	1.897076	2.336672	2.948847	3.837344
$\tau = 20$	1.570796	1.917912	2.394215	3.072774	4.086688

M-estimadores de posición

- $x_i = \mu + \sigma u_i$, $u_i \sim F$ i.i.d. conocida σ conocido.
- f densidad de F , $f_{\mu\sigma}(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right)$,
- Función de verosimilitud correspondiente a x_1, \dots, x_n

$$L(\mu) = \frac{1}{\sigma^n} \prod_{i=1}^n f\left(\frac{x_i - \mu}{\sigma}\right)$$

- El EMV $\hat{\mu}_f$ minimiza

$$S(\mu) = \sum_{i=1}^n \rho_f\left(\frac{x_i - \mu}{\sigma}\right), \quad (1)$$

$$\rho_f(u) = -\log f(u) + \log f(0).$$

M-estimadores de posición

- Huber (1964) definió los **M-estimadores** para el **modelo de posición** como el valor $\hat{\mu}$ que minimiza

$$S(\mu) = \sum_{i=1}^n \rho \left(\frac{x_i - \mu}{\sigma} \right), \quad (2)$$

- ρ se elige independientemente de f y de tal manera que tenga las propiedades deseadas:
 1. El estimador es **altamente eficiente** cuando f corresponde a la distribución $N(0,1)$.
 2. El estimador es **poco sensible a contaminación por outliers**, en particular es altamente eficiente para toda f correspondiente a una distribución de \mathcal{V}_ϵ .

M-estimadores de posición

- Propiedades de la función ρ que define al M-estimador

A1. ρ es derivable con derivada $\psi = \rho'$.

A2. ρ es par.

A3. $\rho(u)$ es monótona no decreciente en $|u|$.

A4. $\rho(0) = 0$.

- Huber propuso una familia de funciones ρ_k^H

$$\rho_k^H(u) = \begin{cases} -ku - \frac{k^2}{2} & \text{si } u < -k \\ \frac{u^2}{2} & \text{si } |u| \leq k \\ ku - \frac{k^2}{2} & \text{si } u > k \end{cases} .$$

M -estimadores de posición

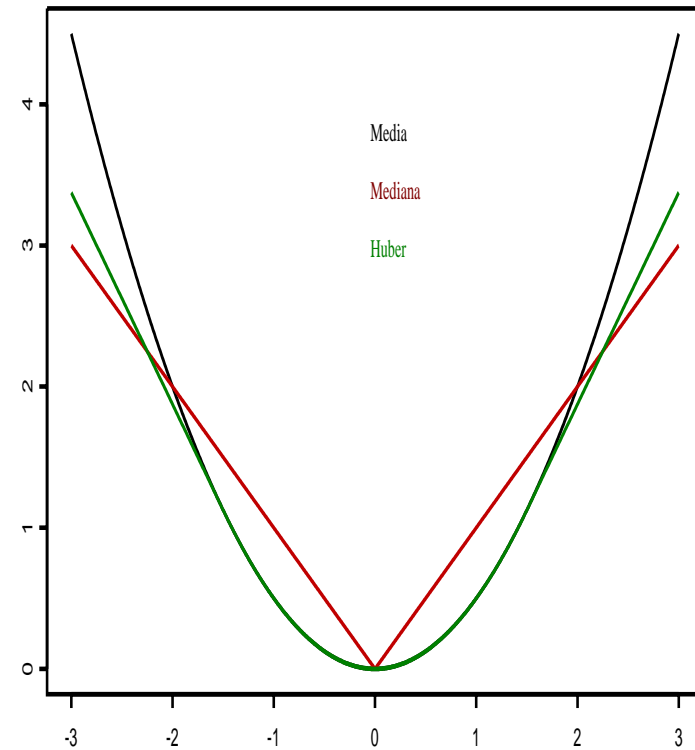
Ejemplo 2

M -estimador con $k = 1.5$ 135.38

Mediana 135.1

\bar{x} 137.05

\bar{x} sin atípicos 134.9



M-estimadores de posición

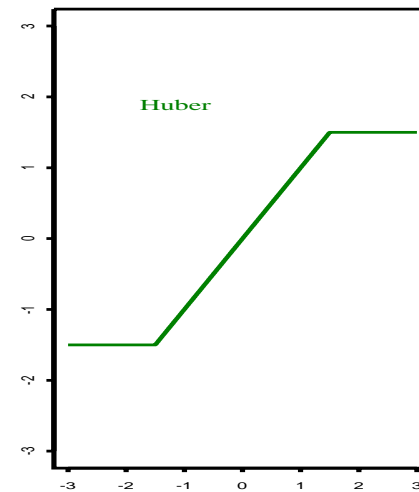
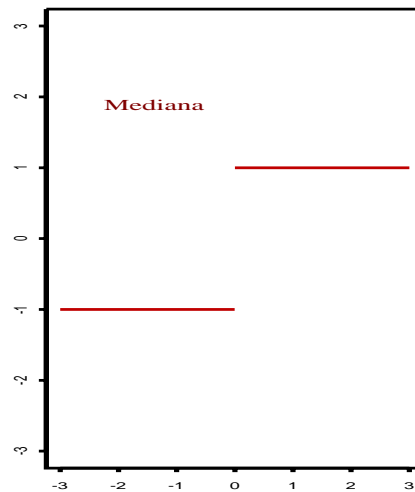
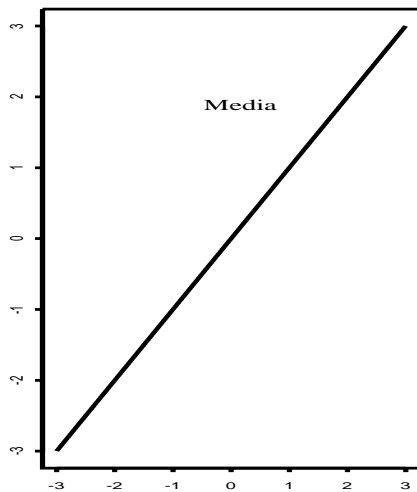
$$A(\mu) = \sum_{i=1}^n \psi \left(\frac{x_i - \mu}{\sigma} \right) = 0. \quad (3)$$

Teorema. *Supongamos que ψ es continua impar, no decreciente y para algún a se tiene $\psi(a) > 0$. Entonces*

- (i) La ecuación (3) tiene al menos una raíz.*
- (ii) Toda raíz de (3) corresponde a un mínimo de $S(\mu)$.*
- (iii) Las raíces de (3) forman un intervalo.*
- (iv) Si ψ es estrictamente creciente hay una única raíz de (3).*

M-estimadores de posición

$$\psi_k^H(u) = \begin{cases} -k & \text{si } u < -k \\ u & \text{si } |u| \leq k \\ k & \text{si } u > k \end{cases} .$$



Propiedades asintóticas de los M -estimadores

- Fisher-consistencia:

$$E_{F_{\mu\sigma}} \left(\psi \left(\frac{x - \mu}{\sigma} \right) \right) = 0, \quad (4)$$

- Es equivalente a $E_F(\psi(u)) = 0$.
- Si F tiene una densidad simétrica respecto de 0 y ψ es impar, se cumple (4).

Teorema Sean x_1, \dots, x_n i.i.d. $\mathbf{x}_i = \mu + \sigma \mathbf{u}_i$, $u_i \sim F$. Sea $\hat{\mu}_n$ solución de (3), donde ψ y F satisfacen $E_F(\psi(u)) = 0$. Luego $\hat{\mu}_n \xrightarrow{a.s.} \mu$ en cualquiera de los siguientes casos

1. La función ψ es estrictamente creciente.
2. La función ψ es no decreciente, $\psi(u) > \psi(0)$ y $F(u) > F(0)$ para todo $u > 0$.

Propiedades asintóticas de los M -estimadores

Teorema. *i.i.d.* $\mathbf{x}_i = \boldsymbol{\mu} + \boldsymbol{\sigma}\mathbf{u}_i$, $u_i \sim F$. Sea $\hat{\boldsymbol{\mu}}_n$ solución de (3), donde ψ y F satisfacen $E_F(\psi(u)) = 0$.

Supongamos que $\hat{\boldsymbol{\mu}}_n \xrightarrow{p} \boldsymbol{\mu}$, y que además ψ tiene dos derivadas continuas y ψ'' es acotada. Entonces,

$$n^{1/2}(\hat{\boldsymbol{\mu}}_n - \boldsymbol{\mu}) \xrightarrow{D} N(0, \sigma^2 V(\psi, F)),$$

donde

$$V(\psi, F) = \frac{E_F \psi^2(u)}{(E_F \psi'(u))^2}. \quad (5)$$

M -estimador minimax

$$\mathcal{V}_{1,\epsilon} = \{F : F = (1 - \epsilon)\Phi + \epsilon H \text{ con } H \text{ simétrica}\}. \quad (6)$$

- La mayor varianza posible en este entorno del M -estimador basado en ψ está dada por $V(\psi) = \sup_{F \in \mathcal{V}_{1,\epsilon}} V(\psi, F)$.
- Huber eligió el M -estimador con función $\tilde{\psi}$, $V(\tilde{\psi}) = \min_{\psi} V(\psi)$.
- Estos estimadores se denominarán **minimax** (minimizan la máxima varianza asintótica en el entorno de contaminación $\mathcal{V}_{1,\epsilon}$).
- Huber (1964) mostró que $\tilde{\psi}$ está en la familia ψ_k^H , donde k depende de la cantidad de contaminación ϵ .

M-estimador con escala desconocida

M-estimador se definirá por el valor $\hat{\mu}$ que minimiza

$$S(\mu) = \sum_{i=1}^n \rho \left(\frac{x_i - \mu}{\hat{\sigma}_n} \right). \quad (7)$$

- $\hat{\sigma}_n$ un estimador robusto y consistente de σ
- Por ejemplo, la MAD, $\hat{\sigma}_n^2 = \kappa \operatorname{median}_{1 \leq i \leq n} \{|x_i - \tilde{\mu}_n|\}$,
- $\tilde{\mu}_n = \operatorname{median}_{1 \leq i \leq n} \{x_i\}$, y
- κ una constante tal que $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$ si $x_i \sim N(\mu, \sigma^2)$.
- $\kappa = \frac{1}{\Phi^{-1}(0.75)} = \frac{1}{0.675}$

M -estimador con escala desconocida

$$\sum_{i=1}^n \psi \left(\frac{x_i - \mu}{\hat{\sigma}} \right) = 0. \quad (8)$$

Teorema. Sean x_1, \dots, x_n i.i.d. $\mathbf{x}_i = \mu + \sigma \mathbf{u}_i$, $u_i \sim F$.

Consideremos un estimador $\hat{\mu}_n$ solución de (8), donde ψ es impar y F es simétrica respecto de 0. Supongamos que $\hat{\mu}_n \xrightarrow{p} \mu$, $\hat{\sigma}_n \xrightarrow{p} \sigma$, y que $\psi \in \mathcal{C}^2$, ψ'' es acotada. Luego se tiene que

$$n^{1/2}(\hat{\mu}_n - \mu) \xrightarrow{D} N(0, \sigma^2 V(\psi, F)),$$

donde $V(\psi, F)$ está dada por

$$V(\psi, F) = \frac{E_F \psi^2(u)}{(E_F \psi'(u))^2}.$$

Otros ejemplos: diagrama de Hertzsprung–Russell

X = logaritmo de la temperatura en la superficie de la estrella

Y = logaritmo de intensidad lumínica.

obs	X	Y	obs	X	Y	obs	X	Y	obs	X	Y
1	4.37	5.23	25	4.38	5.02	13	4.48	5.42	37	4.53	5.10
2	4.56	5.74	26	4.42	4.66	14	4.01	4.05	38	4.45	5.22
3	4.26	4.93	27	4.29	4.66	15	4.29	4.26	39	4.53	5.18
4	4.56	5.74	28	4.38	4.90	16	4.42	4.58	40	4.43	5.57
5	4.30	5.19	29	4.22	4.39	17	4.23	3.94	41	4.38	4.62
6	4.46	5.46	30	3.48	6.05	18	4.42	4.18	42	4.45	5.06
7	3.84	4.65	31	4.38	4.42	19	4.23	4.18	43	4.50	5.34
8	4.57	5.27	32	4.56	5.10	20	3.49	5.89	44	4.45	5.34
9	4.26	5.57	33	4.45	5.22	21	4.29	4.38	45	4.55	5.54
10	4.37	5.12	34	3.49	6.29	22	4.29	4.22	46	4.45	4.98
11	3.49	5.73	35	4.23	4.34	23	4.42	4.42	47	4.42	4.50
12	4.43	5.45	36	4.62	5.62	24	4.49	4.85			

Diagrama de Hertzsprung–Russell

$$g(x) = \theta_1 x + \theta_2. \quad \hat{\theta}_1 = -0.4133039 \quad \text{y} \quad \hat{\theta}_2 = 6.793467$$

Diagrama de Hertzsprung--Russell del cluster de estrellas CYG OB1.
Ajuste por Minimos Cuadrados

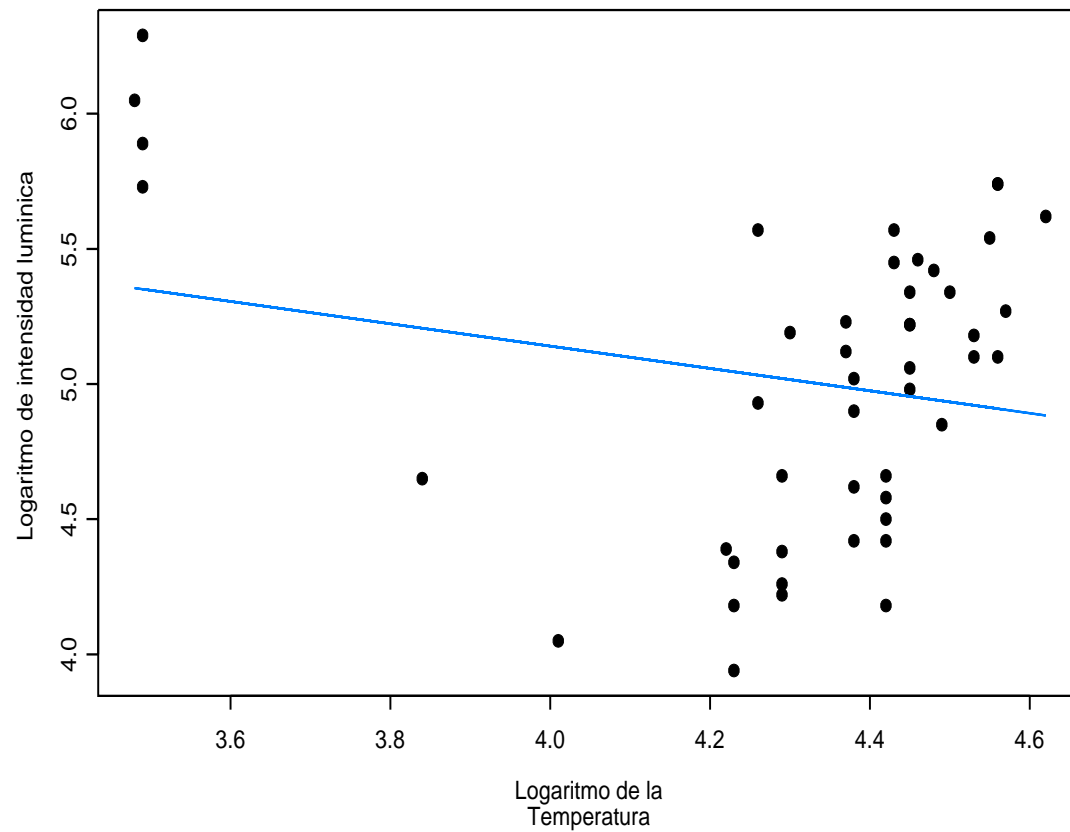


Diagrama de Hertzsprung–Russell: *Least Median of Squares*

$$\operatorname{median}_i \left(Y_i - \left[\tilde{\theta}_2 + \tilde{\theta}_1 X_i \right] \right)^2 = \min_{\theta \in \Theta} \operatorname{median}_i \left(Y_i - [\theta_2 + \theta_1 X_i] \right)^2,$$

$$\tilde{\theta}_1 = 3.970588 \quad \tilde{\theta}_2 = -12.62794$$

Diagrama de Hertzsprung–Russell del cluster de estrellas CYG OB1.

En Azul: Ajuste por Mínimos Cuadrados

En Rojo: Ajuste por LMS

