ESTIMACIÓN DE MÁXIMA LQ-VEROSIMILITUD EN MODELOS CON ERRORES EN LAS VARIABLES FUNCIONALES

Lucas Guarracino y Patricia Giménez

Depto de Matemática. FCEyN. Universidad Nacional de Mar del Plata

Reunión anual de la Unión Matemática Argentina. Bahía Blanca, Septiembre de 2016

Resumen

Resumen

1. Motivación y definición del estimador de máxima Lq-verosimilitud.

Resumen

1. Motivación y definición del estimador de máxima Lq-verosimilitud.
2. Propiedades.

Resumen

1. Motivación y definición del estimador de máxima Lq-verosimilitud.
2. Propiedades.
3. Aplicación.

Resumen

1. Motivación y definición del estimador de máxima Lq-verosimilitud.
2. Propiedades.
3. Aplicación.
4. Estudio de simulación.

Medidas de divergencia

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$\mathcal{D}_{KL}(f, g)$

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$$\mathcal{D}_{KL}(f, g) = E_F \left[\log \left(\frac{f(\boldsymbol{X})}{g(\boldsymbol{X})} \right) \right]$$

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$$\mathcal{D}_{KL}(f, g) = E_F\left[\log\left(\frac{f(\boldsymbol{X})}{g(\boldsymbol{X})}\right)\right] = \int_{\mathcal{X}} f(\boldsymbol{x}) \log\left(\frac{f(\boldsymbol{x})}{g(\boldsymbol{x})}\right) d\mu(\boldsymbol{x}).$$

Medida de divergencia entre densidades de probabilidad

$\mathcal{D} \colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$$\mathcal{D}_{KL}(f, g) = E_F \left[\log \left(\frac{f(\boldsymbol{X})}{g(\boldsymbol{X})} \right) \right] = \int_{\mathcal{X}} f(\boldsymbol{x}) \log \left(\frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} \right) d\mu(\boldsymbol{x}).$$

ϕ-divergencias

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$$\mathcal{D}_{KL}(f, g) = E_F \left[\log \left(\frac{f(\boldsymbol{X})}{g(\boldsymbol{X})} \right) \right] = \int_{\mathcal{X}} f(\boldsymbol{x}) \log \left(\frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} \right) d\mu(\boldsymbol{x}).$$

ϕ-divergencias

$\mathcal{D}_\phi(f, g)$

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f, g) \geq 0$.
2. $\mathcal{D}(f, g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$$\mathcal{D}_{KL}(f, g) = E_F\left[\log\left(\frac{f(\mathbf{X})}{g(\mathbf{X})}\right)\right] = \int_{\mathcal{X}} f(\mathbf{x}) \log\left(\frac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mu(\mathbf{x}).$$

ϕ-divergencias

$$\mathcal{D}_{\phi}(f, g) = E_G\left[\phi\left(\frac{f(\mathbf{X})}{g(\mathbf{X})}\right)\right]$$

Medidas de divergencia

Medida de divergencia entre densidades de probabilidad

$\mathcal{D}\colon \mathcal{G}\times\mathcal{G} \longrightarrow \mathbb{R}_{\geq 0}$ tal que

1. $\mathcal{D}(f,g) \geq 0$.
2. $\mathcal{D}(f,g) = 0$ si y solo si $f = g$.

Divergencia de Kullback-Leibler

$$\mathcal{D}_{KL}(f,g) = E_F\left[\log\left(\frac{f(\mathbf{X})}{g(\mathbf{X})}\right)\right] = \int_{\mathcal{X}} f(\mathbf{x})\log\left(\frac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mu(\mathbf{x}).$$

ϕ-divergencias

$$\mathcal{D}_{\phi}(f,g) = E_G\left[\phi\left(\frac{f(\mathbf{X})}{g(\mathbf{X})}\right)\right] = \int_{\mathcal{X}} g(\mathbf{x})\phi\left(\frac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mu(\mathbf{x}).$$

Medidas de divergencia

Medidas de divergencia

q-divergencia

Medidas de divergencia

q-divergencia

Tomando

$$\phi(x) = \begin{cases} \frac{1}{q(q-1)}\left[x^q - x - (q-1)(x-1)\right], & \text{si } q \neq 0, 1 \\ -x\log x + x - 1, & \text{si } q = 0 \\ x\log x - x + 1, & \text{si } q = 1 \end{cases}$$

q-divergencia

Tomando

$$\phi(x) = \begin{cases} \frac{1}{q(q-1)}\left[x^q - x - (q-1)(x-1)\right], & \text{si } \ q \neq 0, 1 \\ -x\log x + x - 1\,, & \text{si } \ q = 0 \\ x\log x - x + 1\,, & \text{si } \ q = 1 \end{cases}$$

se obtiene
$\mathcal{D}_q(f, g)$

Medidas de divergencia

q-divergencia

Tomando

$$\phi(x) = \begin{cases} \frac{1}{q(q-1)}\left[x^q - x - (q-1)(x-1)\right], & \text{si } q \neq 0, 1 \\ -x\log x + x - 1, & \text{si } q = 0 \\ x\log x - x + 1, & \text{si } q = 1 \end{cases}$$

se obtiene

$\mathcal{D}_q(f,g) = -\frac{1}{q} E_F\left(\log_q\left[\frac{g(\boldsymbol{X})}{f(\boldsymbol{X})}\right]\right)$

q-divergencia

Tomando

$$\phi(x) = \begin{cases} \frac{1}{q(q-1)}\left[x^q - x - (q-1)(x-1)\right], & \text{si } q \neq 0, 1 \\ -x \log x + x - 1, & \text{si } q = 0 \\ x \log x - x + 1, & \text{si } q = 1 \end{cases}$$

se obtiene

$$\mathcal{D}_q(f, g) = -\frac{1}{q} E_F\left(\log_q\left[\frac{g(\boldsymbol{X})}{f(\boldsymbol{X})}\right]\right) = -\frac{1}{q}\int_{\mathcal{X}} f(\boldsymbol{x}) \log_q\left[\frac{g(\boldsymbol{x})}{f(\boldsymbol{x})}\right] d\mu(\boldsymbol{x}),$$

q-divergencia

Tomando

$$\phi(x) = \begin{cases} \frac{1}{q(q-1)}[x^q - x - (q-1)(x-1)]\,, & \text{si } q \neq 0,1 \\ -x\log x + x - 1\,, & \text{si } q = 0 \\ x\log x - x + 1\,, & \text{si } q = 1 \end{cases}$$

se obtiene

$$\mathcal{D}_q(f,g) = -\frac{1}{q}E_F\left(\log_q\left[\frac{g(\boldsymbol{X})}{f(\boldsymbol{X})}\right]\right) = -\frac{1}{q}\int_{\mathcal{X}} f(\boldsymbol{x})\log_q\left[\frac{g(\boldsymbol{x})}{f(\boldsymbol{x})}\right]d\mu(\boldsymbol{x}),$$

donde

$$\log_q x = \begin{cases} \dfrac{x^{1-q}-1}{1-q}\,, & \text{si } q \neq 1 \\ \log x\,, & \text{si } q = 1 \end{cases}$$

Medidas de divergencia

Medidas de divergencia

Si $q = 1$, entonces $\log_q = \log$ y $\mathcal{D}_q(f, g) = \mathcal{D}_{KL}(f, g)$.

Construcción del estimador para una muestra i.i.d.

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.
2. $f_{\boldsymbol{\theta}}$ es una densidad perteneciente a $\mathcal{F}_{\boldsymbol{\theta}}$.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.
2. $f_{\boldsymbol{\theta}}$ es una densidad perteneciente a $\mathcal{F}_{\boldsymbol{\theta}}$.

A partir de una muestra aleatoria $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ proveniente de la distribución desconocida G, queremos minimizar la versión empírica de la q-divergencia

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.
2. $f_{\boldsymbol{\theta}}$ es una densidad perteneciente a $\mathcal{F}_{\boldsymbol{\theta}}$.

A partir de una muestra aleatoria $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ proveniente de la distribución desconocida G, queremos minimizar la versión empírica de la q-divergencia

$$\mathcal{D}_q(\hat{g}, f_{\boldsymbol{\theta}}) = -\frac{1}{q} E_{\hat{G}} \left(\log_q \left[\frac{f_{\boldsymbol{\theta}}(\boldsymbol{X})}{\hat{g}(\boldsymbol{X})} \right] \right),$$

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.
2. $f_{\boldsymbol{\theta}}$ es una densidad perteneciente a $\mathcal{F}_{\boldsymbol{\theta}}$.

A partir de una muestra aleatoria $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ proveniente de la distribución desconocida G, queremos minimizar la versión empírica de la q-divergencia

$$\mathcal{D}_q(\hat{g}, f_{\boldsymbol{\theta}}) = -\frac{1}{q} E_{\hat{G}} \left(\log_q \left[\frac{f_{\boldsymbol{\theta}}(\boldsymbol{X})}{\hat{g}(\boldsymbol{X})} \right] \right),$$

donde

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.
2. $f_{\boldsymbol{\theta}}$ es una densidad perteneciente a $\mathcal{F}_{\boldsymbol{\theta}}$.

A partir de una muestra aleatoria $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ proveniente de la distribución desconocida G, queremos minimizar la versión empírica de la q-divergencia

$$\mathcal{D}_q(\hat{g}, f_{\boldsymbol{\theta}}) = -\frac{1}{q} E_{\hat{G}} \left(\log_q \left[\frac{f_{\boldsymbol{\theta}}(\boldsymbol{X})}{\hat{g}(\boldsymbol{X})} \right] \right),$$

donde

1. $\hat{G}$ es la distribución empírica de la muestra.

Construcción del estimador para una muestra i.i.d.

Modelamos la densidad g con un modelo estadístico paramétrico $\mathcal{F}_{\boldsymbol{\theta}}$, minimizando $\mathcal{D}_q(g, f_{\boldsymbol{\theta}})$ respecto a $\boldsymbol{\theta}$, donde

1. g es la verdadera densidad desconocida que genera los datos.
2. $f_{\boldsymbol{\theta}}$ es una densidad perteneciente a $\mathcal{F}_{\boldsymbol{\theta}}$.

A partir de una muestra aleatoria $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ proveniente de la distribución desconocida G, queremos minimizar la versión empírica de la q-divergencia

$$\mathcal{D}_q(\hat{g}, f_{\boldsymbol{\theta}}) = -\frac{1}{q} E_{\hat{G}} \left(\log_q \left[\frac{f_{\boldsymbol{\theta}}(\boldsymbol{X})}{\hat{g}(\boldsymbol{X})} \right] \right),$$

donde

1. $\hat{G}$ es la distribución empírica de la muestra.
2. $\hat{g}$ es algún estimador no paramétrico de g en base a la muestra.

Construcción del estimador para una muestra i.i.d.

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012):

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012): Si cambiamos g por $g^{\left(\frac{1}{q}\right)}$, obtenemos

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012): Si cambiamos g por $g^{\left(\frac{1}{q}\right)}$, obtenemos

$$\mathcal{D}_q(g^{\left(\frac{1}{q}\right)}, f_{\theta}) = cte\left(\mathcal{H}_q(g, f_{\theta}) - \mathcal{H}_q(g, g^{\left(\frac{1}{q}\right)})\right),$$

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012): Si cambiamos g por $g^{\left(\frac{1}{q}\right)}$, obtenemos

$$\mathcal{D}_q(g^{\left(\frac{1}{q}\right)}, f_{\boldsymbol{\theta}}) = cte\left(\mathcal{H}_q(g, f_{\boldsymbol{\theta}}) - \mathcal{H}_q(g, g^{\left(\frac{1}{q}\right)})\right),$$

donde

1. $\mathcal{H}_q(g, f) = -E_G\left[\log_q f(\boldsymbol{X})\right]$

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012): Si cambiamos g por $g^{\left(\frac{1}{q}\right)}$, obtenemos

$$\mathcal{D}_q(g^{\left(\frac{1}{q}\right)}, f_{\boldsymbol{\theta}}) = cte\left(\mathcal{H}_q(g, f_{\boldsymbol{\theta}}) - \mathcal{H}_q(g, g^{\left(\frac{1}{q}\right)})\right),$$

donde

1. $\mathcal{H}_q(g, f) = -E_G\left[\log_q f(\boldsymbol{X})\right]$ (q-entropía cruzada).

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012): Si cambiamos g por $g^{\left(\frac{1}{q}\right)}$, obtenemos

$$\mathcal{D}_q(g^{\left(\frac{1}{q}\right)}, f_{\boldsymbol{\theta}}) = cte\left(\mathcal{H}_q(g, f_{\boldsymbol{\theta}}) - \mathcal{H}_q(g, g^{\left(\frac{1}{q}\right)})\right),$$

donde

1. $\mathcal{H}_q(g, f) = -E_G\left[\log_q f(\boldsymbol{X})\right]$ (q-entropía cruzada).
2. $g^{\left(\frac{1}{q}\right)}(x) = \dfrac{g(x)^{\frac{1}{q}}}{\int g(x)^{\frac{1}{q}}dx}$

Construcción del estimador para una muestra i.i.d.

Complicaciones al estimar g en problemas multivariados:

1. Difícil selección del ancho de banda.
2. Presición del estimador dependiente de la convergencia del estimador de g.

Ferrari y Laveccia (2012): Si cambiamos g por $g^{\left(\frac{1}{q}\right)}$, obtenemos

$$\mathcal{D}_q(g^{\left(\frac{1}{q}\right)}, f_{\boldsymbol{\theta}}) = cte\left(\mathcal{H}_q(g, f_{\boldsymbol{\theta}}) - \mathcal{H}_q(g, g^{\left(\frac{1}{q}\right)})\right),$$

donde

1. $\mathcal{H}_q(g, f) = -E_G\left[\log_q f(\boldsymbol{X})\right]$ (q-entropía cruzada).
2. $g^{\left(\frac{1}{q}\right)}(x) = \dfrac{g(x)^{\frac{1}{q}}}{\int g(x)^{\frac{1}{q}} dx}$ (transformación potencia de g)

Construcción del estimador para una muestra i.i.d.

Entonces, minimizar $\mathcal{D}_q(g^{(\frac{1}{q})}, f_{\boldsymbol{\theta}})$ es equivalente a minimizar $\mathcal{H}_q(g, f_{\boldsymbol{\theta}})$.

Entonces, minimizar $\mathcal{D}_q(g^{(\frac{1}{q})}, f_{\boldsymbol{\theta}})$ es equivalente a minimizar $\mathcal{H}_q(g, f_{\boldsymbol{\theta}})$. Se evita el problema de estimar g, aunque se cambia de densidad objetivo.

Entonces, minimizar $\mathcal{D}_q(g^{(\frac{1}{q})}, f_{\boldsymbol{\theta}})$ es equivalente a minimizar $\mathcal{H}_q(g, f_{\boldsymbol{\theta}})$. Se evita el problema de estimar g, aunque se cambia de densidad objetivo.
Definimos al estimador de máxima Lq-verosimilitud (mínima q-entropía) como

Construcción del estimador para una muestra i.i.d.

Entonces, minimizar $\mathcal{D}_q(g^{(\frac{1}{q})}, f_{\boldsymbol{\theta}})$ es equivalente a minimizar $\mathcal{H}_q(g, f_{\boldsymbol{\theta}})$. Se evita el problema de estimar g, aunque se cambia de densidad objetivo.
Definimos al estimador de máxima Lq-verosimilitud (mínima q-entropía) como

$$\hat{\boldsymbol{\theta}}^*_q = argmax_{\boldsymbol{\theta}\in\Theta}\left\{\sum_{i=1}^{n}\log_q\left[f(\boldsymbol{z}_i;\boldsymbol{\theta})\right]\right\}.$$

Entonces, minimizar $\mathcal{D}_q(g^{(\frac{1}{q})}, f_{\boldsymbol{\theta}})$ es equivalente a minimizar $\mathcal{H}_q(g, f_{\boldsymbol{\theta}})$. Se evita el problema de estimar g, aunque se cambia de densidad objetivo.

Definimos al estimador de máxima Lq-verosimilitud (mínima q-entropía) como

$$\hat{\boldsymbol{\theta}}_q^* = argmax_{\boldsymbol{\theta}\in\Theta}\left\{\sum_{i=1}^{n}\log_q\left[f(\boldsymbol{z}_i;\boldsymbol{\theta})\right]\right\}.$$

En notación funcional

$$T_q^*(H) = \boldsymbol{\theta}_q^* = argmax_{\boldsymbol{\theta}\in\Theta} E_H\left\{\log_q\left[f(\boldsymbol{Z};\boldsymbol{\theta})\right]\right\}.$$

Observaciones:

Observaciones:

1. T_q^* no es Fisher-consistente, pues $T_q^*(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.

Construcción del estimador para una muestra i.i.d.

Observaciones:

1. T_q^* no es Fisher-consistente, pues $T_q^*(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.
2. $T_q^*(F_{\boldsymbol{\theta}})$ es el parámetro de $f_{\boldsymbol{\theta}}^{(q)}$, siempre que esta transformación sea cerrada en el modelo.

Construcción del estimador para una muestra i.i.d.

Observaciones:

1. T_q^* no es Fisher-consistente, pues $T_q^*(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.
2. $T_q^*(F_{\boldsymbol{\theta}})$ es el parámetro de $f_{\boldsymbol{\theta}}^{(q)}$, siempre que esta transformación sea cerrada en el modelo.
3. $\hat{\boldsymbol{\theta}}_q^*$ es consistente para $T_q^*(G)$.

Observaciones:

1. T_q^* no es Fisher-consistente, pues $T_q^*(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.
2. $T_q^*(F_{\boldsymbol{\theta}})$ es el parámetro de $f_{\boldsymbol{\theta}}^{(q)}$, siempre que esta transformación sea cerrada en el modelo.
3. $\hat{\boldsymbol{\theta}}_q^*$ es consistente para $T_q^*(G)$.
4. Por ser un M-estimador, $\hat{\boldsymbol{\theta}}_q^*$ tiene distribución asintótica conocida.

Construcción del estimador para una muestra i.i.d.

Observaciones:

1. T^*_q no es Fisher-consistente, pues $T^*_q(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.
2. $T^*_q(F_{\boldsymbol{\theta}})$ es el parámetro de $f^{(q)}_{\boldsymbol{\theta}}$, siempre que esta transformación sea cerrada en el modelo.
3. $\hat{\boldsymbol{\theta}}^*_q$ es consistente para $T^*_q(G)$.
4. Por ser un M-estimador, $\hat{\boldsymbol{\theta}}^*_q$ tiene distribución asintótica conocida.

Si está bien definida la biyección τ_q en el espacio paramétrico de manera que $T^*_q(F_{\boldsymbol{\theta}}) = \tau_q^{-1}(\boldsymbol{\theta})$ entonces definimos al estimador de mínima q-divergencia mediante: $\hat{\boldsymbol{\theta}}_q = \tau_q\left(\hat{\boldsymbol{\theta}}^*_q\right)$,

Construcción del estimador para una muestra i.i.d.

Observaciones:

1. T_q^* no es Fisher-consistente, pues $T_q^*(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.
2. $T_q^*(F_{\boldsymbol{\theta}})$ es el parámetro de $f_{\boldsymbol{\theta}}^{(q)}$, siempre que esta transformación sea cerrada en el modelo.
3. $\hat{\boldsymbol{\theta}}_q^*$ es consistente para $T_q^*(G)$.
4. Por ser un M-estimador, $\hat{\boldsymbol{\theta}}_q^*$ tiene distribución asintótica conocida.

Si está bien definida la biyección τ_q en el espacio paramétrico de manera que $T_q^*(F_{\boldsymbol{\theta}}) = \tau_q^{-1}(\boldsymbol{\theta})$ entonces definimos al estimador de mínima q-divergencia mediante: $\hat{\boldsymbol{\theta}}_q = \tau_q\left(\hat{\boldsymbol{\theta}}_q^*\right)$, o en notación funcional: $T_q(H) = \tau_q\left(T_q^*(H)\right)$.

Construcción del estimador para una muestra i.i.d.

Observaciones:

1. T_q^* no es Fisher-consistente, pues $T_q^*(F_{\boldsymbol{\theta}}) \neq \boldsymbol{\theta}$.
2. $T_q^*(F_{\boldsymbol{\theta}})$ es el parámetro de $f_{\boldsymbol{\theta}}^{(q)}$, siempre que esta transformación sea cerrada en el modelo.
3. $\hat{\boldsymbol{\theta}}_q^*$ es consistente para $T_q^*(G)$.
4. Por ser un M-estimador, $\hat{\boldsymbol{\theta}}_q^*$ tiene distribución asintótica conocida.

Si está bien definida la biyección τ_q en el espacio paramétrico de manera que $T_q^*(F_{\boldsymbol{\theta}}) = \tau_q^{-1}(\boldsymbol{\theta})$ entonces definimos al estimador de mínima q-divergencia mediante: $\hat{\boldsymbol{\theta}}_q = \tau_q\left(\hat{\boldsymbol{\theta}}_q^*\right)$, o en notación funcional: $T_q(H) = \tau_q\left(T_q^*(H)\right)$.
Las propiedades de T_q y $\hat{\boldsymbol{\theta}}_q$ se deducen inmediatamente de las de T_q^* y $\hat{\boldsymbol{\theta}}_q^*$ a través de τ_q.

Muestra no idénticamente distribuída

Muestra no idénticamente distribuída

Si cada elemento $\boldsymbol{Z}_i$ de la muestra tiene distribución G_i, modelamos cada g_i con un modelo estadístico paramétrico $\mathcal{F}_{i,\boldsymbol{\theta}}$, todos con el mismo espacio parametral, y definimos el estimador de máxima Lq-verosimilitud mediante

Muestra no idénticamente distribuída

Si cada elemento $\boldsymbol{Z}_i$ de la muestra tiene distribución G_i, modelamos cada g_i con un modelo estadístico paramétrico $\mathcal{F}_{i,\boldsymbol{\theta}}$, todos con el mismo espacio parametral, y definimos el estimador de máxima Lq-verosimilitud mediante

$$\hat{\boldsymbol{\theta}}_q^* = argmax_{\boldsymbol{\theta}\in\Theta}\left\{\sum_{i=1}^{n}\log_q\left[f_i(\boldsymbol{z}_i;\boldsymbol{\theta})\right]\right\},$$

Muestra no idénticamente distribuída

Si cada elemento $\boldsymbol{Z}_i$ de la muestra tiene distribución G_i, modelamos cada g_i con un modelo estadístico paramétrico $\mathcal{F}_{i,\boldsymbol{\theta}}$, todos con el mismo espacio parametral, y definimos el estimador de máxima Lq-verosimilitud mediante

$$\hat{\boldsymbol{\theta}}_q^* = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[f_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\},$$

o en notación funcional

Muestra no idénticamente distribuída

Si cada elemento $\boldsymbol{Z}_i$ de la muestra tiene distribución G_i, modelamos cada g_i con un modelo estadístico paramétrico $\mathcal{F}_{i,\boldsymbol{\theta}}$, todos con el mismo espacio parametral, y definimos el estimador de máxima Lq-verosimilitud mediante

$$\hat{\boldsymbol{\theta}}_q^* = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[f_i(\boldsymbol{z}_i;\boldsymbol{\theta}) \right] \right\},$$

o en notación funcional

$$T_q^*(H_1,\dots,H_n) = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q f_i(\boldsymbol{Z}_i;\boldsymbol{\theta}) \right] \right\}.$$

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\left\{ \begin{array}{l} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{array} \right. ,$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes.

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes. Entonces

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes.
Entonces

1. $X_i \sim N(\xi_i, \phi)$.

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes. Entonces

1. $X_i \sim N(\xi_i, \phi)$.
2. $\boldsymbol{Y}_i \sim \boldsymbol{N}_p(\boldsymbol{\alpha} + \xi_i \boldsymbol{\beta}, \phi \boldsymbol{Id}_p)$.

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes. Entonces

1. $X_i \sim N(\xi_i, \phi)$.
2. $\boldsymbol{Y}_i \sim \boldsymbol{N}_p(\boldsymbol{\alpha} + \xi_i \boldsymbol{\beta}, \phi \boldsymbol{Id}_p)$.
3. $\boldsymbol{Z}_i \sim \boldsymbol{N}_{p+1}(\boldsymbol{\mu}_i, \phi \boldsymbol{Id}_{p+1})$, con $\boldsymbol{\mu}_i = (\xi_i, \boldsymbol{\alpha}^t + \xi_i \boldsymbol{\beta}^t)^t$.

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes. Entonces

1. $X_i \sim N(\xi_i, \phi)$.
2. $\boldsymbol{Y}_i \sim \boldsymbol{N}_p(\boldsymbol{\alpha} + \xi_i \boldsymbol{\beta}, \phi \boldsymbol{Id}_p)$.
3. $\boldsymbol{Z}_i \sim \boldsymbol{N}_{p+1}(\boldsymbol{\mu}_i, \phi \boldsymbol{Id}_{p+1})$, con $\boldsymbol{\mu}_i = (\xi_i, \boldsymbol{\alpha}^t + \xi_i \boldsymbol{\beta}^t)^t$.
4. $\boldsymbol{\theta} = (\boldsymbol{\alpha}^t, \boldsymbol{\beta}^t, \phi)^t$: parámetro estructural.

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes. Entonces

1. $X_i \sim N(\xi_i, \phi)$.
2. $\boldsymbol{Y}_i \sim \boldsymbol{N}_p(\boldsymbol{\alpha} + \xi_i \boldsymbol{\beta}, \phi \boldsymbol{Id}_p)$.
3. $\boldsymbol{Z}_i \sim \boldsymbol{N}_{p+1}(\boldsymbol{\mu}_i, \phi \boldsymbol{Id}_{p+1})$, con $\boldsymbol{\mu}_i = (\xi_i, \boldsymbol{\alpha}^t + \xi_i \boldsymbol{\beta}^t)^t$.
4. $\boldsymbol{\theta} = (\boldsymbol{\alpha}^t, \boldsymbol{\beta}^t, \phi)^t$: parámetro estructural.
5. ξ_i's : parámetros incidentales.

Modelo de calibración comparativa

Sea $\boldsymbol{Z}_1 = (X_1, \boldsymbol{Y}_1^t)^t, \ldots, \boldsymbol{Z}_n = (X_n, \boldsymbol{Y}_n^t)^t$ una muestra aleatoria, con

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{\alpha} + \xi_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \\ X_i = \xi_i + U_i \end{cases},$$

donde $\boldsymbol{\epsilon}_i \sim \boldsymbol{N}_p(\boldsymbol{0}, \phi \boldsymbol{Id}_p)$ y $U_i \sim N(0, \phi)$ son independientes. Entonces

1. $X_i \sim N(\xi_i, \phi)$.
2. $\boldsymbol{Y}_i \sim \boldsymbol{N}_p(\boldsymbol{\alpha} + \xi_i \boldsymbol{\beta}, \phi \boldsymbol{Id}_p)$.
3. $\boldsymbol{Z}_i \sim \boldsymbol{N}_{p+1}(\boldsymbol{\mu}_i, \phi \boldsymbol{Id}_{p+1})$, con $\boldsymbol{\mu}_i = (\xi_i, \boldsymbol{\alpha}^t + \xi_i \boldsymbol{\beta}^t)^t$.
4. $\boldsymbol{\theta} = (\boldsymbol{\alpha}^t, \boldsymbol{\beta}^t, \phi)^t$: parámetro estructural.
5. ξ_i's : parámetros incidentales.

La muestra no es idénticamente distribuída.

Modelo de calibración comparativa

Modelo de calibración comparativa

$$f_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\|\boldsymbol{z}-\boldsymbol{\mu}_i\|^2\right\}.$$

$$f_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\|\boldsymbol{z}-\boldsymbol{\mu}_i\|^2\right\}.$$

EMLqV de los parámetros incidentales

$$\hat{\xi}_i = \frac{1}{c}\left((\boldsymbol{y}_i - \boldsymbol{\alpha})^t\boldsymbol{\beta} + x_i\right),$$

Modelo de calibración comparativa

$$f_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\|\boldsymbol{z}-\boldsymbol{\mu}_i\|^2\right\}.$$

EMLqV de los parámetros incidentales

$$\hat{\xi}_i = \frac{1}{c}\left((\boldsymbol{y}_i - \boldsymbol{\alpha})^t\boldsymbol{\beta} + x_i\right),$$

con $c = 1 + \boldsymbol{\beta}^t\boldsymbol{\beta}$.

Modelo de calibración comparativa

$$f_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\|\boldsymbol{z}-\boldsymbol{\mu}_i\|^2\right\}.$$

EMLqV de los parámetros incidentales

$$\hat{\xi}_i = \frac{1}{c}\left((\boldsymbol{y}_i - \boldsymbol{\alpha})^t\boldsymbol{\beta} + x_i\right),$$

con $c = 1 + \boldsymbol{\beta}^t\boldsymbol{\beta}$.

Densidad perfilada

$$\tilde{f}_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\left[\boldsymbol{y} - (\boldsymbol{\alpha} + x\boldsymbol{\beta})\right]^t \boldsymbol{B}\left[\boldsymbol{y} - (\boldsymbol{\alpha} + x\boldsymbol{\beta})\right]\right\},$$

Modelo de calibración comparativa

$$f_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\|\boldsymbol{z}-\boldsymbol{\mu}_i\|^2\right\}.$$

EMLqV de los parámetros incidentales

$$\hat{\xi}_i = \frac{1}{c}\left((\boldsymbol{y}_i-\boldsymbol{\alpha})^t\boldsymbol{\beta}+x_i\right),$$

con $c = 1+\boldsymbol{\beta}^t\boldsymbol{\beta}$.

Densidad perfilada

$$\tilde{f}_i(\boldsymbol{z};\boldsymbol{\theta}) = (2\pi\phi)^{-\frac{p+1}{2}} \exp\left\{-\frac{1}{2\phi}\left[\boldsymbol{y}-(\boldsymbol{\alpha}+x\boldsymbol{\beta})\right]^t \boldsymbol{B}\left[\boldsymbol{y}-(\boldsymbol{\alpha}+x\boldsymbol{\beta})\right]\right\},$$

donde $\boldsymbol{B} = \frac{1}{c}\left(\boldsymbol{Id}_p-\boldsymbol{\beta}\boldsymbol{\beta}^t\right)$.

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\}.$$

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

$$T_q^{\dagger}(H_1, \ldots, H_n) = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q \tilde{f}_i(\boldsymbol{z}; \boldsymbol{\theta}) \right] \right\}.$$

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta} \in \Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

$$T_q^{\dagger}(H_1, \ldots, H_n) = argmax_{\boldsymbol{\theta} \in \Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q \tilde{f}_i(\boldsymbol{z}; \boldsymbol{\theta}) \right] \right\}.$$

Ecuación de estimación

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

$$T_q^{\dagger}(H_1, \ldots, H_n) = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q \tilde{f}_i(\boldsymbol{z}; \boldsymbol{\theta}) \right] \right\}.$$

Ecuación de estimación

$\sum_{i=1}^{n} \mathbf{u}_i^{\dagger}(\mathbf{Z}_i; \boldsymbol{\theta})$

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i; \boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

$$T_q^{\dagger}(H_1, \ldots, H_n) = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q \tilde{f}_i(\boldsymbol{z}; \boldsymbol{\theta}) \right] \right\}.$$

Ecuación de estimación

$\sum_{i=1}^{n} \mathbf{u}_i^{\dagger}(\mathbf{Z}_i; \boldsymbol{\theta}) = \sum_{i=1}^{n} \tilde{f}_i(\mathbf{Z}_i; \boldsymbol{\theta})^{1-q} \tilde{\mathbf{u}}_i(\mathbf{Z}_i; \boldsymbol{\theta})$

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\theta}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i;\boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

$$T_q^{\dagger}(H_1,\ldots,H_n) = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q \tilde{f}_i(\boldsymbol{z};\boldsymbol{\theta}) \right] \right\}.$$

Ecuación de estimación

$\sum_{i=1}^n \mathbf{u}_i^{\dagger}(\mathbf{Z}_i;\theta) = \sum_{i=1}^n \tilde{f}_i(\mathbf{Z}_i;\theta)^{1-q}\tilde{\mathbf{u}}_i(\mathbf{Z}_i;\theta) = \mathbf{0}$,

Modelo de calibración comparativa

EMLqV en el modelo de calibración comparativa

$$\hat{\boldsymbol{\theta}}_q^{\dagger} = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} \log_q \left[\tilde{f}_i(\boldsymbol{z}_i;\boldsymbol{\theta}) \right] \right\}.$$

En notación funcional

$$T_q^{\dagger}(H_1,\ldots,H_n) = argmax_{\boldsymbol{\theta}\in\Theta} \left\{ \sum_{i=1}^{n} E_{H_i} \left[\log_q \tilde{f}_i(\boldsymbol{z};\boldsymbol{\theta}) \right] \right\}.$$

Ecuación de estimación

$\sum_{i=1}^{n} \mathbf{u}_i^{\dagger}(\mathbf{Z}_i;\boldsymbol{\theta}) = \sum_{i=1}^{n} \tilde{f}_i(\mathbf{Z}_i;\boldsymbol{\theta})^{1-q}\tilde{\mathbf{u}}_i(\mathbf{Z}_i;\boldsymbol{\theta}) = \mathbf{0}$,
con $\tilde{\mathbf{u}}_i = \frac{\partial}{\partial\boldsymbol{\theta}} \log \tilde{f}_i$.

EMLqV de los parámetros incidentales

EMLqV de los parámetros incidentales

Asumimos la existencia de un único $\boldsymbol{\theta}_q^{\dagger}$, independiente del índice i, tal que para todo $i \in \mathbb{N}$

Asumimos la existencia de un único $\boldsymbol{\theta}_q^\dagger$, independiente del índice i, tal que para todo $i \in \mathbb{N}$

$$E_{G_i}\left[\mathbf{u}_i^\dagger(\boldsymbol{Z}_i; \boldsymbol{\theta}_q^\dagger)\right] = \mathbf{0}.$$

EMLqV de los parámetros incidentales

Asumimos la existencia de un único $\boldsymbol{\theta}_q^\dagger$, independiente del índice i, tal que para todo $i \in \mathbb{N}$

$$E_{G_i}\left[\mathbf{u}_i^\dagger(\boldsymbol{Z}_i; \boldsymbol{\theta}_q^\dagger)\right] = \mathbf{0}.$$

Bajo condiciones de regularidad

EMLqV de los parámetros incidentales

EMLqV de los parámetros incidentales

1. $\hat{\boldsymbol{\theta}}_q^\dagger$ converge en probabilidad a $\boldsymbol{\theta}_q^\dagger$.

EMLqV de los parámetros incidentales

1. $\hat{\boldsymbol{\theta}}_q^\dagger$ converge en probabilidad a $\boldsymbol{\theta}_q^\dagger$.
2. $\hat{\boldsymbol{\theta}}_q^\dagger$ es asintóticamente normal con media $\boldsymbol{\theta}_q^\dagger$ y matriz de covarianza asintótica $\frac{1}{n}\boldsymbol{V}_n^\dagger(\boldsymbol{\theta}_q^\dagger)$,

1. $\hat{\boldsymbol{\theta}}_q^{\dagger}$ converge en probabilidad a $\boldsymbol{\theta}_q^{\dagger}$.
2. $\hat{\boldsymbol{\theta}}_q^{\dagger}$ es asintóticamente normal con media $\boldsymbol{\theta}_q^{\dagger}$ y matriz de covarianza asintótica $\frac{1}{n}\boldsymbol{V}_n^{\dagger}(\boldsymbol{\theta}_q^{\dagger})$, siendo

$$\boldsymbol{V}_n^{\dagger}(\boldsymbol{\theta}) = \left[\boldsymbol{\Psi}_n^{\dagger}(\boldsymbol{\theta})\right]^{-1} \boldsymbol{\Omega}_n^{\dagger}(\boldsymbol{\theta}) \left[\boldsymbol{\Psi}_n^{\dagger}(\boldsymbol{\theta})^t\right]^{-1},$$

EMLqV de los parámetros incidentales

1. $\hat{\boldsymbol{\theta}}_q^\dagger$ converge en probabilidad a $\boldsymbol{\theta}_q^\dagger$.
2. $\hat{\boldsymbol{\theta}}_q^\dagger$ es asintóticamente normal con media $\boldsymbol{\theta}_q^\dagger$ y matriz de covarianza asintótica $\frac{1}{n}\boldsymbol{V}_n^\dagger(\boldsymbol{\theta}_q^\dagger)$, siendo

$$\boldsymbol{V}_n^\dagger(\boldsymbol{\theta}) = \left[\boldsymbol{\Psi}_n^\dagger(\boldsymbol{\theta})\right]^{-1} \boldsymbol{\Omega}_n^\dagger(\boldsymbol{\theta}) \left[\boldsymbol{\Psi}_n^\dagger(\boldsymbol{\theta})^t\right]^{-1},$$

donde

$$\boldsymbol{\Omega}_n^\dagger(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} E_{G_i}\left[\mathbf{u}_i^\dagger(\boldsymbol{Z}_i;\boldsymbol{\theta})\mathbf{u}_i^\dagger(\boldsymbol{Z}_i;\boldsymbol{\theta})^t\right]$$

EMLqV de los parámetros incidentales

1. $\hat{\boldsymbol{\theta}}_q^{\dagger}$ converge en probabilidad a $\boldsymbol{\theta}_q^{\dagger}$.
2. $\hat{\boldsymbol{\theta}}_q^{\dagger}$ es asintóticamente normal con media $\boldsymbol{\theta}_q^{\dagger}$ y matriz de covarianza asintótica $\frac{1}{n}\boldsymbol{V}_n^{\dagger}(\boldsymbol{\theta}_q^{\dagger})$, siendo

$$\boldsymbol{V}_n^{\dagger}(\boldsymbol{\theta}) = \left[\boldsymbol{\Psi}_n^{\dagger}(\boldsymbol{\theta})\right]^{-1} \boldsymbol{\Omega}_n^{\dagger}(\boldsymbol{\theta}) \left[\boldsymbol{\Psi}_n^{\dagger}(\boldsymbol{\theta})^t\right]^{-1},$$

donde

$$\boldsymbol{\Omega}_n^{\dagger}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} E_{G_i}\left[\mathbf{u}_i^{\dagger}(\boldsymbol{Z}_i;\boldsymbol{\theta})\mathbf{u}_i^{\dagger}(\boldsymbol{Z}_i;\boldsymbol{\theta})^t\right]$$

y

$$\boldsymbol{\Psi}_n^{\dagger}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} E_{G_i}\left[\mathbf{I}_i^{\dagger}(\boldsymbol{Z}_i;\boldsymbol{\theta})\right].$$

EMLqV de los parámetros incidentales

EMLqV de los parámetros incidentales

Supongamos que estamos en la situación ideal en que cada $g_i \in \mathcal{F}_{i,\boldsymbol{\theta}}$, es decir, existe $\boldsymbol{\theta}_0$ y $\xi_1^0, \dots, \xi_n^0$ tales que $g_i = f_{i,\boldsymbol{\theta}_0}$, $\forall\ i = 1, \dots, n$.

EMLqV de los parámetros incidentales

Supongamos que estamos en la situación ideal en que cada $g_i \in \mathcal{F}_{i,\boldsymbol{\theta}}$, es decir, existe $\boldsymbol{\theta}_0$ y $\xi_1^0, \ldots, \xi_n^0$ tales que $g_i = f_{i,\boldsymbol{\theta}_0}$, $\forall\ i = 1, \ldots, n$. En este contexto se puede calcular $\boldsymbol{\theta}_q^\dagger$ y la matriz de covarianza asintótica, a saber

EMLqV de los parámetros incidentales

Supongamos que estamos en la situación ideal en que cada $g_i \in \mathcal{F}_{i,\boldsymbol{\theta}}$, es decir, existe $\boldsymbol{\theta}_0$ y $\xi_1^0, \ldots, \xi_n^0$ tales que $g_i = f_{i,\boldsymbol{\theta}_0}$, $\forall\ i = 1, \ldots, n$. En este contexto se puede calcular $\boldsymbol{\theta}_q^\dagger$ y la matriz de covarianza asintótica, a saber

$$\begin{cases} \boldsymbol{\alpha}^\dagger = \boldsymbol{\alpha}_0 \\ \boldsymbol{\beta}^\dagger = \boldsymbol{\beta}_0 \\ \phi^\dagger = k\phi_0 \end{cases}$$

EMLqV de los parámetros incidentales

Supongamos que estamos en la situación ideal en que cada $g_i \in \mathcal{F}_{i,\boldsymbol{\theta}}$, es decir, existe $\boldsymbol{\theta}_0$ y $\xi_1^0, \ldots, \xi_n^0$ tales que $g_i = f_{i,\boldsymbol{\theta}_0}$, $\forall\ i = 1, \ldots, n$. En este contexto se puede calcular $\boldsymbol{\theta}_q^\dagger$ y la matriz de covarianza asintótica, a saber

$$\begin{cases} \boldsymbol{\alpha}^\dagger = \boldsymbol{\alpha}_0 \\ \boldsymbol{\beta}^\dagger = \boldsymbol{\beta}_0 \\ \phi^\dagger = k\phi_0 \end{cases}$$

$$\frac{1}{n}\boldsymbol{V}_n^\dagger(\boldsymbol{\theta}^\dagger) = \frac{{\phi_0}^2}{n}\left(\frac{a^{(1)}}{c^{(1)}}\right)^2 \frac{c^{(2)}}{a^{(2)}} \begin{pmatrix} \mathbf{K} \otimes \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{4{\phi^\dagger}^2 t}{(p+1)^2} \end{pmatrix}$$

Función de influencia

Función de influencia

Como $T_q^\dagger$ es un M-estimador, la función de influencia al contaminar el dato i_0 con una distribución degenerada con masa uno en el punto de contaminación $\boldsymbol{t}_{i_0}$ toma la forma

Función de influencia

Como $T_q^\dagger$ es un M-estimador, la función de influencia al contaminar el dato i_0 con una distribución degenerada con masa uno en el punto de contaminación $\boldsymbol{t}_{i_0}$ toma la forma

$$\boldsymbol{IF}(T_q^\dagger; \boldsymbol{t}_{i_0}; G_1, \ldots, G_n) = -\frac{1}{n}\left[\boldsymbol{\Psi}_n(\boldsymbol{\theta}_q^\dagger)\right]^{-1} \boldsymbol{u}_i^\dagger(\boldsymbol{t}_{i_0}, \boldsymbol{\theta}_q^\dagger)$$

Función de influencia

Como $T_q^\dagger$ es un M-estimador, la función de influencia al contaminar el dato i_0 con una distribución degenerada con masa uno en el punto de contaminación $\boldsymbol{t}_{i_0}$ toma la forma

$$\boldsymbol{IF}(T_q^\dagger; \boldsymbol{t}_{i_0}; G_1, \ldots, G_n) = -\frac{1}{n}\left[\boldsymbol{\Psi}_n(\boldsymbol{\theta}_q^\dagger)\right]^{-1} \boldsymbol{u}_i^\dagger(\boldsymbol{t}_{i_0}, \boldsymbol{\theta}_q^\dagger)$$

En general no es acotada, pero para el modelo de calibración comparativa se observa que los puntos con residuos grandes (alejados del modelo) tienen muy poca influencia.

EFICIENCIA

EFICIENCIA

q	1	0.99	0.98	0.95	0.90	0.85	0.80	0.75
ARE($\hat{\beta}^\dagger$)	100	99.96	99.82	98.88	95.56	90.13	82.81	73.85
ARE($\hat{\phi}^\dagger$)	100	99.91	99.64	97.78	91.35	81.45	69.13	55.64

Table: Eficiencia relativa asintótica de los estimadores de máxima L_q-verosimilitud $\hat{\alpha}^\dagger$, $\hat{\beta}^\dagger$ y $\hat{\phi}^\dagger$ para varios valores de q ($p = 2$)

Algunos estudios de simulación

ROBUSTEZ

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

1. $U_i \sim N(0,1)$.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

1. $U_i \sim N(0,1)$. Este esquema representa el caso de un modelo puro.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

1. $U_i \sim N(0,1)$. Este esquema representa el caso de un modelo puro.
2. $U_i \sim (1-\epsilon)N(0,\phi_0) + \epsilon N(2,\phi_0)$

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

1. $U_i \sim N(0,1)$. Este esquema representa el caso de un modelo puro.
2. $U_i \sim (1-\epsilon)N(0,\phi_0) + \epsilon N(2,\phi_0)$, donde ϵ es el porcentaje de contaminación.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

1. $U_i \sim N(0,1)$. Este esquema representa el caso de un modelo puro.
2. $U_i \sim (1-\epsilon)N(0,\phi_0) + \epsilon N(2,\phi_0)$, donde ϵ es el porcentaje de contaminación.

Usamos 1000 replicaciones para estimar sesgo y MSE (error cuadrático medio) de los estimadores.

ROBUSTEZ

Para explorar la estabilidad de los estimadores bajo contaminación consideramos el modelo de calibración comparativa con $p = 2$. Fijamos $n = 50, \boldsymbol{\alpha}_0 = (0,0)^t, \boldsymbol{\beta}_0 = (1,1)^t$ y consideramos dos valores de ϕ_0 : 0.1 y 0.5. Vamos a considerar solo contaminaciones en la variable x_i (o U_i), de manera que $\boldsymbol{e}_i \sim N_2(\mathbf{0}, \phi_0 \boldsymbol{Id}_2)$.

1. $U_i \sim N(0,1)$. Este esquema representa el caso de un modelo puro.
2. $U_i \sim (1-\epsilon)N(0,\phi_0) + \epsilon N(2,\phi_0)$, donde ϵ es el porcentaje de contaminación.

Usamos 1000 replicaciones para estimar sesgo y MSE (error cuadrático medio) de los estimadores. Consideramos $\epsilon = 0, 0.05, 0.10, 0.15$ y 0.20 y valores de $q = 0.75, 0.8, 0.85, 0.9, 0.95$ y 1.

		q					
ϕ_0	ϵ	1	0.95	0.90	0.85	0.80	0.75
0.1	0	**0.0168**	0.0171	0.0179	0.0192	0.0214	0.0249
	0.05	0.0581	0.0351	0.0234	**0.0215**	0.0227	0.0263
	0.10	0.1363	0.0877	0.0459	0.0271	**0.0252**	0.0290
	0.15	0.2436	0.1812	0.1084	0.0519	**0.0316**	0.0339
	0.20	0.3620	0.2967	0.2128	0.1163	0.0643	**0.0538**
0.5	0	**0.1007**	0.1025	0.1072	0.1155	0.1299	0.1595
	0.05	0.1339	0.1294	**0.1286**	0.1328	0.1442	0.1680
	0.10	0.2081	0.1936	0.1833	**0.1797**	0.1854	0.2057
	0.15	0.3113	0.2899	0.2716	0.2590	**0.2573**	0.2843
	0.20	0.4316	0.4072	0.3846	0.3663	**0.3560**	0.3643

Table: MSE simulado para el estimador de máxima L_q-verosimilitud $\hat{\theta}^{\dagger}$.

Algunos estudios de simulación

Del análisis de la tabla anterior surge que:

Del análisis de la tabla anterior surge que:

1. Para cualquier nivel de contaminación ϵ fijo y valor de q fijo, el MSE crece cuando la varianza del error de medición ϕ_0 crece.

Del análisis de la tabla anterior surge que:

1. Para cualquier nivel de contaminación ϵ fijo y valor de q fijo, el MSE crece cuando la varianza del error de medición ϕ_0 crece.
2. Bajo el modelo puro ($\epsilon = 0$), el EMV es el más eficiente (el mínimo MSE se da para $q = 1$), como se esperaba.

Del análisis de la tabla anterior surge que:

1. Para cualquier nivel de contaminación ϵ fijo y valor de q fijo, el MSE crece cuando la varianza del error de medición ϕ_0 crece.
2. Bajo el modelo puro ($\epsilon = 0$), el EMV es el más eficiente (el mínimo MSE se da para $q = 1$), como se esperaba.
3. Para ϕ_0 fijo, a medida que aumenta el nivel de contaminación decrece el valor de q para el cual MSE es mínimo.

Del análisis de la tabla anterior surge que:

1. Para cualquier nivel de contaminación ϵ fijo y valor de q fijo, el MSE crece cuando la varianza del error de medición ϕ_0 crece.
2. Bajo el modelo puro ($\epsilon = 0$), el EMV es el más eficiente (el mínimo MSE se da para $q = 1$), como se esperaba.
3. Para ϕ_0 fijo, a medida que aumenta el nivel de contaminación decrece el valor de q para el cual MSE es mínimo.
4. Para un nivel de contaminación ϵ fijo, el valor de q para el cual MSE es mínimo aumenta o permanece igual cuando ϕ_0 aumenta. Es decir, el valor de q que minimiza el MSE es una función no decreciente de ϕ_0.

La siguiente tabla muestra los valores de q que minimizan el error cuadrático medio (MSE) para distintos valores de ϕ_0 y niveles de contaminación ϵ.

La siguiente tabla muestra los valores de q que minimizan el error cuadrático medio (MSE) para distintos valores de ϕ_0 y niveles de contaminación ϵ.

	ϵ			
ϕ_0	0.05	0.10	0.15	0.20
0.1	0.85	0.82	0.80	0.76
0.3	0.86	0.80	0.77	0.75
0.5	0.92	0.85	0.82	0.79
0.7	0.95	0.90	0.87	0.86

Table: Valores de q óptimo

La siguiente tabla muestra los valores de q que minimizan el error cuadrático medio (MSE) para distintos valores de ϕ_0 y niveles de contaminación ϵ.

	ϵ			
ϕ_0	0.05	0.10	0.15	0.20
0.1	0.85	0.82	0.80	0.76
0.3	0.86	0.80	0.77	0.75
0.5	0.92	0.85	0.82	0.79
0.7	0.95	0.90	0.87	0.86

Table: Valores de q óptimo