Teoría detrás de las herramientas de laboratorio

Miguel Calvo Arnal, Juan Guerrero Marcos y Andrés Laín Sanclemente

Índice

Notación 0.1. Para las medidas no usamos ningún tipo de símbolo en especial x, para la variable aleatoria asociada a una medida x usaremos la notación $\mathring{x}$, para la variable asociada a la variable aleatoria $\mathring{x}$ y que aparece como argumento en su función de densidad de probabilidad y de distribución usaremos $\dot{x}$. Para un valor ideal o «real» de una medida x usaremos $\tilde{x}$ y, por último, para denotar un estimador de una variable aleatoria $\mathring{x}$ usaremos $\hat{x}$.

1. Distribución normal o gaussiana

Proposición 1.1. *Sea $\mathring{x}$ una variable aleatoria gaussiana de centro μ y varianza σ^2. Entonces $\mathring{y} = \mathring{x}+b$, donde $b \in \mathbb{R}$ es una constante, es también una variable aleatoria gaussiana de varianza σ^2, pero de centro $\mu + b$.*

$$\mathring{x} \sim \mathcal{N}\left(\mu, \sigma^2\right) \implies \mathring{y} = \mathring{x} + b \sim \mathcal{N}\left(\mu + b, \sigma^2\right) \qquad \forall b \in \mathbb{R}$$

Demostración. La función de distribución de la variable aleatoria $\mathring{y}$ se relaciona con la de $\mathring{x}$ mediante:

$$F_{\mathring{y}}\left(\dot{y}\right) = \mathrm{P}\left(\mathring{y} < \dot{y}\right) = \mathrm{P}\left(\mathring{x} + b < \dot{y}\right) = \mathrm{P}\left(\mathring{x} < \dot{y} - b\right) = F_{\mathring{x}}\left(\dot{y} - b\right)$$

Derivando la expresión anterior, obtenemos:

$$f_{\mathring{y}}\left(\dot{y}\right) = \frac{\mathrm{d}F_{\mathring{y}}}{\mathrm{d}\dot{y}}\left(\dot{y}\right) = \frac{\mathrm{d}F_{\mathring{x}}}{\mathrm{d}\dot{x}}\left(\dot{y} - b\right) = f_{\mathring{x}}\left(\dot{y} - b\right) =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\dot{y} - b - \mu\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\dot{y} - \left(\mu + b\right)\right)^2}{2\sigma^2}\right)$$

que claramente se corresponde con la la función de densidad de una variable aleatoria gaussiana de centro $\mu + b$ y varianza σ^2. $\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 1.2. *Sea $\mathring{x}$ una variable aleatoria gaussiana de centro μ y varianza σ^2. Entonces $\mathring{y} = a\mathring{x}$, donde $a \in \mathbb{R}$ es una constante, es también una variable aleatoria normal, pero de centro $a\mu$ y de varianza $a^2\sigma^2$.*

$$\mathring{x} \sim \mathcal{N}\left(\mu, \sigma^2\right) \implies \mathring{y} = a\mathring{x} \sim \mathcal{N}\left(a\mu, a^2\sigma^2\right) \qquad \forall a \in \mathbb{R}$$

Demostración. La función de densidad de la variable aleatoria $\mathring{y}$ se relaciona con la de $\mathring{x}$ mediante:

$$F_{\mathring{y}}\left(\dot{y}\right) = \mathrm{P}\left(\mathring{y} < \dot{y}\right) = \mathrm{P}\left(a\mathring{x} < \dot{y}\right) = \mathrm{P}\left(\mathring{x} < \frac{\dot{y}}{a}\right) = F_{\mathring{x}}\left(\frac{\dot{y}}{a}\right)$$

Derivando la expresión anterior, obtenemos:

$$f_{\mathring{y}}\left(\dot{y}\right) = \frac{\mathrm{d}F_{\mathring{y}}}{\mathrm{d}\dot{y}}\left(\dot{y}\right) = \frac{\mathrm{d}F_{\mathring{x}}}{\mathrm{d}\dot{x}}\left(\frac{\dot{y}}{a}\right)\frac{1}{a} = \frac{1}{a}f_{\mathring{x}}\left(\frac{\dot{y}}{a}\right) =$$

$$= \frac{1}{a}\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\frac{\dot{y}}{a} - \mu\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left(-\frac{\left(\frac{\dot{y} - a\mu}{a}\right)^2}{2\sigma^2}\right) =$$

$$= \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left(-\frac{\left(\dot{y} - a\mu\right)^2}{2a^2\sigma^2}\right)$$

Y lo anterior es claramente la función de densidad de probabilidad de una variable aleatoria gaussiana de centro $a\mu$ y varianza $a^2\sigma^2$. $\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 1.3. *Toda variable aleatoria normal puede «normalizarse» a una variable aleatoria gaussiana de centro* 0 *y varianza* 1*.*

$$\mathring{x} \sim \mathcal{N}\left(\mu, \sigma^2\right) \implies \mathring{z} = \frac{\mathring{x} - \mu}{\sigma} \sim \mathcal{N}\left(0, 1\right)$$

Demostración. Sea $\mathring{x} \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Por la proposición 1.1 en la página anterior, la variable aleatoria $\mathring{y} = \mathring{x} - \mu$, cumple $\mathring{y} \sim \mathcal{N}\left(0, \sigma^2\right)$. Por último, por la proposición 1.2 en la página anterior, la variable aleatoria $\mathring{z} = \frac{1}{\sigma}\mathring{y}$ satisface $\mathring{z} \sim \mathcal{N}\left(0, 1\right)$. Por tanto:

$$\mathring{z} = \frac{\mathring{y}}{\sigma} = \frac{\mathring{x} - \mu}{\sigma} \sim \mathcal{N}\left(0, 1\right)$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 1.4. *La combinación lineal de dos variables aleatorias gaussianas normalizadas es otra variable aleatoria normal de centro* 0 *y varianza* $a^2 + b^2$*, siendo* a *y* b *los coeficientes de la combinación lineal.*

$$\mathring{x}_1, \mathring{x}_2 \sim \mathcal{N}\left(0, 1\right) \implies \mathring{y} = a\mathring{x}_1 + b\mathring{x}_2 \sim \mathcal{N}\left(0, a^2 + b^2\right)$$

con $a, b \in \mathbb{R}$*.*

Demostración. Intentemos obtener la función de distribución de la variable $\mathring{y}$:

$$F_{\mathring{y}}\left(\dot{y}\right) = \mathrm{P}\left(\mathring{y} < \dot{y}\right) = \mathrm{P}\left(a\mathring{x}_1 + b\mathring{x}_2 < \dot{y}\right) = \iint_{\{a\dot{x}_1 + b\dot{x}_2 < \dot{y}\}} f_{\mathring{x}_1}\left(\dot{x}_1\right) f_{\mathring{x}_2}\left(\dot{x}_2\right) \mathrm{d}\dot{x}_1 \mathrm{d}\dot{x}_2 =$$

$$= \iint_{\{a\dot{x}_1 + b\dot{x}_2 < \dot{y}\}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\dot{x}_1^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\dot{x}_2^2}{2}\right) \mathrm{d}\dot{x}_1 \mathrm{d}\dot{x}_2 =$$

$$= \iint_{\{a\dot{x}_1 + b\dot{x}_2 < \dot{y}\}} \frac{1}{2\pi} \exp\left(-\frac{\dot{x}_1^2}{2} - \frac{\dot{x}_2^2}{2}\right) \mathrm{d}\dot{x}_1 \mathrm{d}\dot{x}_2 =$$

$$= \iint_{\{a\dot{x}_1 + b\dot{x}_2 < \dot{y}\}} \frac{1}{2\pi} \exp\left(-\frac{\dot{x}_1^2 + \dot{x}_2^2}{2}\right) \mathrm{d}\dot{x}_1 \mathrm{d}\dot{x}_2$$

A continuación, hacemos el cambio de variable:

$$\dot{u}_1 = a\dot{x}_1 + b\dot{x}_2$$

$$\dot{u}_2 = a\dot{x}_1 - b\dot{x}_2$$

Sumando las dos ecuaciones anteriores, podemos despejar $\dot{x}_1$:

$$\dot{u}_1 + \dot{u}_2 = 2a\dot{x}_1 \iff \dot{x}_1 = \frac{\dot{u}_1 + \dot{u}_2}{2a}$$

Y restando las dos ecuaciones anteriores, podemos despejar $\dot{x}_2$:

$$\dot{u}_1 - \dot{u}_2 = 2b\dot{x}_2 \iff \dot{x}_2 = \frac{\dot{u}_1 - \dot{u}_2}{2b}$$

Calculamos ahora la matriz jacobiana del cambio $(\dot{u}_1, \dot{u}_2) \longrightarrow (\dot{x}_1, \dot{x}_2)$:

$$\mathrm{J} = \begin{pmatrix} \frac{\partial \dot{x}_1}{\partial \dot{u}_1} & \frac{\partial \dot{x}_1}{\partial \dot{u}_2} \\ \frac{\partial \dot{x}_2}{\partial \dot{u}_1} & \frac{\partial \dot{x}_2}{\partial \dot{u}_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2a} & \frac{1}{2a} \\ \frac{1}{2b} & -\frac{1}{2b} \end{pmatrix}$$

$$|\det J| = \left| -\frac{1}{4ab} - \frac{1}{4ab} \right| = \frac{1}{2ab}$$

Por otra parte, el conjunto de integración queda:

$$\left\{(\dot{x}_1, \dot{x}_2) \in \mathbb{R}^2 \text{ t.q. } a\dot{x}_1 + b\dot{x}_2 < \dot{y}\right\} = \{(\dot{u}_1, \dot{u}_2) \in (-\infty, \dot{y}) \times \mathbb{R}\}$$

Con este cambio de variable el dominio de integración ha quedado infinitamente más sencillo de manejar. De esta forma, por el teorema del cambio de variable, la integral resulta:

$$F_{\dot{y}}(\dot{y}) = \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{2ab} \frac{1}{2\pi} \exp\left(-\frac{\left(\frac{\dot{u}_1+\dot{u}_2}{2a}\right)^2 + \left(\frac{\dot{u}_1-\dot{u}_2}{2b}\right)^2}{2}\right) \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2 =$$

$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left(-\frac{\frac{\dot{u}_1^2+\dot{u}_2^2+2\dot{u}_1\dot{u}_2}{4a^2} + \frac{\dot{u}_1^2+\dot{u}_2^2-2\dot{u}_1\dot{u}_2}{4b^2}}{2}\right) \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2 =$$

$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left(-\frac{\frac{\dot{u}_1^2+\dot{u}_2^2+2\dot{u}_1\dot{u}_2}{a^2} + \frac{\dot{u}_1^2+\dot{u}_2^2-2\dot{u}_1\dot{u}_2}{b^2}}{8}\right) \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2 =$$

$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left[-\frac{b^2\left(\dot{u}_1^2+\dot{u}_2^2+2\dot{u}_1\dot{u}_2\right) + a^2\left(\dot{u}_1^2+\dot{u}_2^2-2\dot{u}_1\dot{u}_2\right)}{8a^2b^2}\right] \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2 =$$

$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left[-\frac{\left(b^2+a^2\right)\dot{u}_1^2 + \left(b^2+a^2\right)\dot{u}_2^2 + 2\dot{u}_1\dot{u}_2\left(b^2-a^2\right)}{8a^2b^2}\right] \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2 \tag{1.1}$$

A continuación, nuestro propósito es completar cuadrados con el fin de poder hacer la integral en la variable $\dot{u}_2$. Para ello, notemos:

$$\left(\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}\right)^2 =$$

$$= \left(b^2+a^2\right)\dot{u}_2^2 + \dot{u}_1^2\frac{\left(b^2-a^2\right)^2}{b^2+a^2} + 2\dot{u}_1\dot{u}_2\left(b^2-a^2\right) \iff$$

$$\iff \left(\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}\right)^2 - \dot{u}_1^2\frac{\left(b^2-a^2\right)^2}{b^2+a^2} = \left(b^2+a^2\right)\dot{u}_2^2 + 2\dot{u}_1\dot{u}_2\left(b^2-a^2\right)$$

De esta forma, podemos expresar el exponente del integrando de la expresión 1.1 como:

$$-\frac{\left(\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}\right)^2 - \dot{u}_1^2\frac{\left(b^2-a^2\right)^2}{b^2+a^2} + \left(b^2+a^2\right)\dot{u}_1^2}{8a^2b^2} =$$

$$= -\frac{\left(\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}\right)^2 + \frac{\left(b^2+a^2\right)^2-\left(b^2-a^2\right)^2}{b^2+a^2}\dot{u}_1^2}{8a^2b^2} =$$

$$= -\frac{\left(\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}\right)^2 + \frac{b^4+a^4+2a^2b^2-b^4-a^4+2a^2b^2}{b^2+a^2}\dot{u}_1^2}{8a^2b^2} =$$

$$= -\frac{\left(\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}\right)^2 + \frac{4a^2b^2}{b^2+a^2}\dot{u}_1^2}{8a^2b^2} =$$

$$= -\frac{1}{2}\left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{\left(b^2-a^2\right)}{\sqrt{b^2+a^2}}}{2ab}\right)^2 - \frac{\dot{u}_1^2}{2\left(a^2+b^2\right)}$$

Sustituyendo de vuelta en la integral de la expresión 1.1 en la página anterior, se tiene:

$$F_{\mathring{y}}(\dot{y}) = \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left[-\frac{1}{2}\left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab}\right)^2 - \frac{\dot{u}_1^2}{2(a^2+b^2)}\right] \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2 =$$

$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left[-\frac{1}{2}\left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab}\right)^2\right] \exp\left[-\frac{\dot{u}_1^2}{2(a^2+b^2)}\right] \mathrm{d}\dot{u}_1 \mathrm{d}\dot{u}_2$$

Por el teorema de Fubini, se tiene:

$$F_{\mathring{y}}(\dot{y}) = \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{1}{4\pi ab} \exp\left[-\frac{\dot{u}_1^2}{2(a^2+b^2)}\right] \underbrace{\left(\int_{\dot{u}_2=-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab}\right)^2\right] \mathrm{d}\dot{u}_2\right)}_{=:I} \mathrm{d}\dot{u}_1 \tag{1.2}$$

Para proseguir, solucionaremos la integral I de forma separada. Para ello, recuérdese que dentro de esta integral $\dot{u}_1$ no es más que una constante. De esta forma, hacemos el cambio de variable:

$$\dot{t} := \frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab} \iff 2ab\dot{t} = \sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}} \iff$$

$$\iff 2ab\dot{t} - \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}} = \sqrt{b^2+a^2}\dot{u}_2 \iff$$

$$\iff \dot{u}_2 = \frac{2ab\dot{t} - \dot{u}_1\frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{\sqrt{a^2+b^2}}$$

Derivando, se tiene:

$$\mathrm{d}\dot{u}_2 = \frac{2ab}{\sqrt{a^2+b^2}}\mathrm{d}\dot{t}$$

Además:

$$\dot{u}_2 \to \infty \iff \dot{t} \to \infty$$
$$\dot{u}_2 \to -\infty \iff \dot{t} \to -\infty$$

Por consiguiente, la integral I queda:

$$I = \int_{\dot{t}=-\infty}^{\infty} \exp\left(-\frac{1}{2}\dot{t}^2\right) \frac{2ab}{\sqrt{a^2+b^2}}\mathrm{d}\dot{t} =$$

$$= \frac{2ab}{\sqrt{a^2+b^2}}\sqrt{2\pi}\int_{\dot{t}=-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\dot{t}^2\right)}_{*} \mathrm{d}\dot{t}$$

donde el término marcado es claramente la función de densidad de una variable aleatoria gaussiana de centro 0 y varianza unidad: $\mathring{t} \sim \mathcal{N}(0,1)$. Por tanto, la integral que nos resta por evaluar debe ser la unidad por la definición de función de densidad. Así, se llega a:

$$I = \frac{2ab}{\sqrt{a^2+b^2}}\sqrt{2\pi}$$

Sustituyendo de vuelta en la integral 1.2, se tiene:

$$F_{\mathring{y}}(\dot{y}) = \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{1}{4\pi ab} \exp\left(-\frac{\dot{u}_1^2}{2(a^2+b^2)}\right) \frac{2ab}{\sqrt{a^2+b^2}}\sqrt{2\pi}\mathrm{d}\dot{u}_1 =$$

$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{\sqrt{2\pi}}{2\pi} \frac{1}{\sqrt{a^2+b^2}} \exp\left(-\frac{\dot{u}_1^2}{2\left(a^2+b^2\right)}\right) \mathrm{d}\dot{u}_1 =$$
$$= \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{1}{\sqrt{2\pi\left(a^2+b^2\right)}} \exp\left(-\frac{\dot{u}_1^2}{2\left(a^2+b^2\right)}\right) \mathrm{d}\dot{u}_1$$

A continuación, por el Teorema Fundamental del Cálculo Integral, se tiene:

$$f_{\mathring{y}}\left(\dot{y}\right) = \frac{\mathrm{d}F_{\mathring{y}}}{\mathrm{d}\dot{y}}\left(\dot{y}\right) = \frac{1}{\sqrt{2\pi\left(a^2+b^2\right)}} \exp\left(-\frac{\dot{y}^2}{2\left(a^2+b^2\right)}\right)$$

que es claramente la función de densidad de una variable aleatoria normal de centro 0 y de varianza a^2+b^2. Por ende:

$$\mathring{y} \sim \mathcal{N}\left(0, a^2+b^2\right)$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Corolario 1.1. *La combinación lineal de dos variables aleatorias gaussianas es otra variable aleatoria gaussiana. En concreto:*

$$\left.\begin{aligned} \mathring{x}_1 \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right) \\ \mathring{x}_1 \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right) \end{aligned}\right\} \implies \mathring{y} = a\mathring{x}_1 + b\mathring{x}_2 \sim \mathcal{N}\left(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right)$$

con $a, b \in \mathbb{R}$.

Demostración. Primero, nótese que:

$$\begin{aligned} \mathring{y} = a\mathring{x}_1 + b\mathring{x}_2 = a\left(\mathring{x}_1 - \mu_1 + \mu_1\right) + b\left(\mathring{x}_2 - \mu_2 + \mu_2\right) = \\ = a\left(\mathring{x}_1 - \mu_1\right) + a\mu_1 + b\left(\mathring{x}_2 - \mu_2\right) + b\mu_2 = \\ = a\sigma_1\left(\frac{\mathring{x}_1-\mu_1}{\sigma_1}\right) + a\mu_1 + b\sigma_2\left(\frac{\mathring{x}_2-\mu_2}{\sigma_2}\right) + b\mu_2 = \\ = a\mu_1 + b\mu_2 + a\sigma_1\left(\frac{\mathring{x}_1-\mu_1}{\sigma_1}\right) + b\sigma_2\left(\frac{\mathring{x}_2-\mu_2}{\sigma_2}\right) \end{aligned} \tag{1.3}$$

Ahora bien, por la proposición 1.3 en la página 4:

$$\mathring{z}_1 \coloneqq \frac{\mathring{x}_1-\mu_1}{\sigma_1} \qquad \mathring{z}_2 \coloneqq \frac{\mathring{x}_2-\mu_2}{\sigma_2}$$

son ambas variables aleatorias normales estándar: $\mathring{z}_1, \mathring{z}_2 \sim \mathcal{N}\left(0,1\right)$. Por tanto, el término;

$$a\sigma_1\left(\frac{\mathring{x}_1-\mu_1}{\sigma_1}\right) + b\sigma_2\left(\frac{\mathring{x}_2-\mu_2}{\sigma_2}\right) = a\sigma_1\mathring{z}_1 + b\sigma_2\mathring{z}_2$$

es una combinación lineal de variables aleatorias gaussianas normalizadas. Por la proposición 1.4 en la página 4, se tiene:

$$\mathring{Z} \coloneqq a\sigma_1\left(\frac{\mathring{x}_1-\mu_1}{\sigma_1}\right) + b\sigma_2\left(\frac{\mathring{x}_2-\mu_2}{\sigma_2}\right) \sim \mathcal{N}\left(0, a^2\sigma_1^2 + b^2\sigma_2^2\right)$$

Sustituyendo de nuevo en la ecuación 1.3, obtenemos:

$$\mathring{y} = a\mu_1 + b\mu_2 + \mathring{Z}$$

donde $\mathring{Z} \sim \mathcal{N}\left(0, a^2\sigma_1^2 + b^2\sigma_2^2\right)$. Por la proposición 1.1 en la página 3, se tiene que:

$$\mathring{y} \sim \mathcal{N}\left(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right)$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Corolario 1.2 (¡Importante!). *Cualquier combinación lineal de variables aleatorias gaussianas es otra variable aleatoria gaussiana.*

$$\mathring{x}_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right) \quad \forall i = 1, \dots, n \implies \mathring{y} = \sum_{i=1}^{n} \alpha_i \mathring{x}_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \alpha_i \mu_i, \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2\right)$$

donde $\vec{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

Demostración. Probaremos el resultado por inducción sobre n. Por el corolario 1.1 en la página anterior, el resultado se cumple para $n = 2$. Supongamos que se cumple para n y veamos si se cumple para $n + 1$. Es decir, sean $\{\mathring{x}_i\}_{i=1}^{n+1}$ variables aleatorias tales que:

$$\mathring{x}_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right) \qquad \forall i = 1, \dots, n$$

y sea $\alpha \in \mathbb{R}^{n+1}$. Entonces:

$$\mathring{y} = \sum_{i=1}^{n+1} \alpha_i \mathring{x}_i = \sum_{i=1}^{n} \alpha_i \mathring{x}_i + \alpha_{n+1} \mathring{x}_{n+1}$$

Por hipótesis de inducción:

$$\mathring{z} \coloneqq \sum_{i=1}^{n} \alpha_i \mathring{x}_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \alpha_i \mu_i, \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2\right)$$

Entonces, nos queda por evaluar la suma de dos variables gaussianas:

$$\mathring{y} = \mathring{z} + \alpha_{n+1} \mathring{x}_{n+1}$$

Por el corolario 1.1 en la página anterior, se tiene:

$$\mathring{y} \sim \mathcal{N}\left(\sum_{i=1}^{n} \alpha_i \mu_i + \alpha_{n+1} \mathring{x}_{n+1}, \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 + \alpha_{n+1}^2 \sigma_{n+1}^2\right) =$$

$$= \mathcal{N}\left(\sum_{i=1}^{n+1} \alpha_i \mu_i, \sum_{i=1}^{n+1} \alpha_i^2 \sigma_i^2\right)$$

por lo que el resultado se cumple para $n + 1$, lo que completa la demostración. $\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Corolario 1.3. *Toda combinación lineal normalizada de variables gaussianas normalizadas es otra variable aleatoria gaussiana normalizada.*

$$\left.\begin{array}{r} \mathring{x}_i \sim \mathcal{N}(0,1) \quad \forall i = 1, \dots, n \\ \vec{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \ \ t.q. \ \ ||\vec{\alpha}|| = 1 \end{array}\right\} \implies \alpha^T \vec{\mathring{x}} = \sum_{i=1}^{n} \alpha_i \mathring{x}_i \sim \mathcal{N}(0,1)$$

Demostración. Por el corolario 1.1 en la página anterior, se tiene que:

$$\sum_{i=1}^{n} \alpha_i \mathring{x}_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \alpha_i \cdot 0, \sum_{i=1}^{n} \alpha_i^2 \cdot 1^2\right) = \mathcal{N}\left(0, \sum_{i=1}^{n} \alpha_i^2\right) =$$

$$= \mathcal{N}\left(0, ||\vec{\alpha}||^2\right) = \mathcal{N}\left(0, 1^2\right) = \mathcal{N}(0,1)$$

dado que $\vec{\alpha} \in \mathbb{R}^n$ era de norma unidad. $\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

1.1. Distribución χ^2_ν

Proposición 1.5. *Dado un vector de variables aleatorias gaussianas normalizadas* $\vec{\mathring{x}} = (\mathring{x}_1, \ldots, \mathring{x}_n)$ *que pertenece a un espacio vectorial de dimensión* $n-k$*, entonces, el cuadrado de la norma euclídea del vector sigue una distribución* χ^2 *con* $n-k$ *grados de libertad. Es decir:*

$$\left.\begin{array}{r} \mathring{x}_i \sim \mathcal{N}(0,1) \quad \forall i = 1, \ldots, n \\ \dim\left(\left\langle \vec{\mathring{x}} \right\rangle\right) = n-k \end{array}\right\} \implies \left|\left|\vec{\mathring{x}}\right|\right|^2 = \mathring{x}_1^2 + \cdots + \mathring{x}_n^2 = \sum_{i=1}^{n} \mathring{x}_i^2 \sim \chi^2_{n-k}$$

Demostración. Llamemos W al espacio vectorial generado por el vector aleatorio $\vec{\mathring{x}}$, que tiene, por hipótesis, dimensión $n-k$. Entonces, sabemos que podemos descomponer $\mathbb{R}^n$ como suma directa de W y de su complementario ortogonal. Es decir:

$$\mathbb{R}^n = W \oplus W^\perp \tag{1.4}$$

A continuación, sea $\{\vec{a}_1, \ldots, \vec{a}_{n-k}\}$ una base ortonormal de W (que sabemos que siempre existe al menos una) y sea $\left\{\vec{b}_1, \ldots, \vec{b}_k\right\}$ una base ortonormal de $W^\perp$. Entonces, por la ecuación 1.4 $U := \{\vec{a}_1, \ldots, \vec{a}_{n-k}\} \cup \left\{\vec{b}_1, \ldots, \vec{b}_k\right\} = \left\{\vec{a}_1, \ldots, \vec{a}_{n-k}, \vec{b}_1, \ldots, \vec{b}_k\right\}$ es una base ortonormal de $\mathbb{R}^n$.

Ahora bien, en este momento, estamos expresando el vector $\vec{\mathring{x}}$ en base canónica (E), que es una base ortonormal. Por tanto, el cambio de base entre E y U vendrá dado por una matriz ortogonal. Es decir, existe $Q \in \mathbb{R}^{n\times n}$ tal que $Q^T Q = \mathbb{I}$ y $\vec{\mathring{z}} \in \mathbb{R}^n$ tal que:

$$\vec{\mathring{z}} = Q\vec{\mathring{x}} \tag{1.5}$$

es la representación de $\vec{\mathring{x}}$ en base U. Además, como Q es una matriz ortogonal, sus columnas son vectores de norma unidad de $\mathbb{R}^n$. De esta forma, cada $\mathring{z}_i$ es una combinación lineal normalizada de las $\{\mathring{x}_i\}_{i=1}^n$ y, por el lema 1.3 en la página anterior, cada $\mathring{z}_i$ es también una variable aleatoria gaussiana normalizada. Como $\vec{\mathring{x}} \in W$, sabemos que las últimas k componentes de $\vec{\mathring{z}}$ serán, necesariamente, nulas. Es decir:

$$\mathring{z}_{n-k+1} = \mathring{z}_{n-k+2} = \cdots = \mathring{z}_n = 0 \tag{1.6}$$

Por otra parte:

$$\left|\left|\vec{\mathring{x}}\right|\right|^2 = \sum_{i=1}^{n} \mathring{x}_i^2 = \vec{\mathring{x}}^T\vec{\mathring{x}} = \vec{\mathring{x}}^T\mathbb{I}\vec{\mathring{x}} = \vec{\mathring{x}}^T Q^T Q\vec{\mathring{x}}$$

ya que $Q^T Q = \mathbb{I}$. Si seguimos operando, se tiene:

$$\left|\left|\vec{\mathring{x}}\right|\right|^2 = \left(Q\vec{\mathring{x}}\right)^T Q\vec{\mathring{x}} = \vec{\mathring{z}}^T\vec{\mathring{z}} = \sum_{i=1}^{n} \mathring{z}_i^2$$

ya que era $\vec{\mathring{z}} = Q\vec{\mathring{x}}$ por la ecuación 1.5. Ahora bien, aplicando la ecuación 1.6, se obtiene:

$$\left|\left|\vec{\mathring{x}}\right|\right|_2^2 = \sum_{i=1}^{n-k} \mathring{z}_i^2 + \sum_{i=n-k+1}^{n} \mathring{z}_i^2 = \sum_{i=1}^{n-k} \mathring{z}_i^2$$

Como $\{\mathring{z}_i\}_{i=1}^{n-k}$ es un conjunto de $n-k$ variables aleatorias gaussianas normalizadas independientes entre sí, entonces:

$$\left|\left|\vec{\mathring{x}}\right|\right|_2^2 = \sum_{i=1}^{n} \mathring{x}_i^2 = \sum_{i=1}^{n-k} \mathring{z}_i \sim \chi^2_{n-k}$$

por definición de la distribución χ^2_{n-k}. $\mathbb{Q.E.D.}$

2. Propagación de errores

Teorema 2.1. *Sea Ω un conjunto abierto de $\mathbb{R}^n$. A continuación, sea $\mathring{\vec{x}} = (\mathring{x}_1, \ldots, \mathring{x}_n)$ un vector de variables aleatorias y sea $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ una función escalar de clase $C^{(1)}$. Al primer orden no nulo, se tiene:*

$$\mathrm{E}\left(f\left(\mathring{\vec{x}}\right)\right) = f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(2\right)$$

$$\mathrm{Var}\left(f\left(\mathring{\vec{x}}\right)\right) = \left[\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right]^T \mathrm{Cov}\left(\mathring{\vec{x}}\right) \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(3\right)$$

Demostración. Como la función f es de clase $C^{(1)}$, admite un desarrollo de Taylor de orden 1 en torno a $\mathrm{E}\left(\mathring{\vec{x}}\right)$.

$$f\left(\dot{\vec{x}}\right) = f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left[\dot{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right] + o\left(2\right) \tag{2.1}$$

donde es importante distinguir $\dot{\vec{x}}$ de $\mathring{\vec{x}}$. A continuación, obtengamos el valor esperado de $f\left(\mathring{\vec{x}}\right)$:

$$\mathrm{E}\left[f\left(\mathring{\vec{x}}\right)\right] = \mathrm{E}\left[f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left[\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right] + o\left(2\right)\right] =$$

$$= f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \mathrm{E}\left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(2\right)$$

Ahora bien, recordemos que $\mathrm{E}\left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right) = 0$ para cualquier variable aleatoria. De esta forma, tenemos:

$$\mathrm{E}\left[f\left(\mathring{\vec{x}}\right)\right] = f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(2\right)$$

con lo que llegamos a la primera parte del enunciado.

Vamos con la varianza. Aplicaremos la expresión:

$$\mathrm{Var}\left(\mathring{y}\right) = \mathrm{E}\left(\mathring{y}^2\right) - \left(\mathrm{E}\left(\mathring{y}\right)\right)^2$$

a la variable aleatoria $\mathring{y} = f\left(\mathring{\vec{x}}\right)$, que sabemos que debe cumplirse para cualquier variable aleatoria. Elevando al cuadrado la expresión 2.1, se tiene:

$$f\left(\dot{\vec{x}}\right)^2 = f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)^2 + \left[\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left(\dot{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right]^2 + 2f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left(\dot{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(3\right)$$

agrupando en $o\left(3\right)$ todos los términos de orden 3 y superior. A continuación, podemos calcular el valor esperado de $\left[f\left(\mathring{\vec{x}}\right)\right]^2$:

$$\mathrm{E}\left[f\left(\mathring{\vec{x}}\right)^2\right] = \mathrm{E}\left[f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)^2 + \left[\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right]^2 + \right.$$

$$\left. + 2f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(3\right)\right] =$$

$$= f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)^2 + \mathrm{E}\left[\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)^2\right] + 2f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \mathrm{E}\left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(3\right)$$

De nuevo, como es $\mathrm{E}\left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right) = 0$, se tiene:

$$\mathrm{E}\left[f\left(\mathring{\vec{x}}\right)^2\right] = f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)^2 + \mathrm{E}\left[\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) \cdot \left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)^2\right] + o\left(3\right)$$

De esta forma, ya estamos en disposición de obtener la varianza de $f\left(\mathring{\vec{x}}\right)$:

$$\mathrm{Var}\left(f\left(\mathring{\vec{x}}\right)\right) = \mathrm{E}\left(f\left(\mathring{\vec{x}}\right)^2\right) - \left[\mathrm{E}\left(f\left(\mathring{\vec{x}}\right)\right)\right]^2 =$$

$$= f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)^2 + \mathrm{E}\left[\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\cdot\left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)^2\right] - f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)^2 + o\left(3\right) =$$

$$= \mathrm{E}\left[\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\cdot\left(\mathring{\vec{x}} - \mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)^2\right] + o\left(3\right)$$

Para proseguir, desarrollamos el producto escalar dentro del cuadrado:

$$\mathrm{Var}\left(f\left(\mathring{\vec{x}}\right)\right) = \mathrm{E}\left[\left(\sum_{i=1}^{n}\frac{\partial f}{\partial \dot{x}_i}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\left(\mathring{x}_i - \mathrm{E}\left(\mathring{x}_i\right)\right)\right)^2\right] + o\left(3\right)$$

Ahora, expandimos el cuadrado anterior como el doble sumatorio de cualquier combinación de términos:

$$\mathrm{Var}\left(f\left(\mathring{\vec{x}}\right)\right) = \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial f}{\partial \dot{x}_i}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\frac{\partial f}{\partial \dot{x}_j}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\left(\mathring{x}_i - \mathrm{E}\left(\mathring{x}_i\right)\right)\left(\mathring{x}_j - \mathrm{E}\left(\mathring{x}_j\right)\right)\right] + o\left(3\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial f}{\partial \dot{x}_i}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\frac{\partial f}{\partial \dot{x}_j}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\mathrm{E}\left[\left(\mathring{x}_i - \mathrm{E}\left(\mathring{x}_i\right)\right)\left(\mathring{x}_j - \mathrm{E}\left(\mathring{x}_j\right)\right)\right] + o\left(3\right)$$

Justo el término $\mathrm{E}\left[\left(\mathring{x}_i - \mathrm{E}\left(\mathring{x}_i\right)\right)\left(\mathring{x}_j - \mathrm{E}\left(\mathring{x}_j\right)\right)\right]$ es la definición de covarianza entre las variables aleatorias $\mathring{x}_i$ y $\mathring{x}_j$. De esta forma, obtenemos:

$$\mathrm{Var}\left(f\left(\mathring{\vec{x}}\right)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial f}{\partial \dot{x}_i}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\frac{\partial f}{\partial \dot{x}_j}\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\mathrm{Cov}\left(\mathring{x}_i, \mathring{x}_j\right) + o\left(3\right)$$

A continuación, es nuestra intención convertir lo anterior en un producto vector*matriz*vector. Para ello, recordamos:

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial \dot{x}_1}, \ldots, \frac{\partial f}{\partial \dot{x}_n}\right)$$

$$\mathrm{Cov}\left(\mathring{\vec{x}}\right) = \begin{pmatrix} \mathrm{Var}\left(\mathring{x}_1\right) & \mathrm{Cov}\left(\mathring{x}_1, \mathring{x}_2\right) & \cdots & \mathrm{Cov}\left(\mathring{x}_1, \mathring{x}_{n-1}\right) & \mathrm{Cov}\left(\mathring{x}_1, \mathring{x}_n\right) \\ \mathrm{Cov}\left(\mathring{x}_2, \mathring{x}_1\right) & \ddots & & & \mathrm{Cov}\left(\mathring{x}_2, \mathring{x}_n\right) \\ \vdots & & \ddots & & \vdots \\ \mathrm{Cov}\left(\mathring{x}_{n-1}, \mathring{x}_1\right) & & & \ddots & \mathrm{Cov}\left(\mathring{x}_{n-1}, \mathring{x}_n\right) \\ \mathrm{Cov}\left(\mathring{x}_n, \mathring{x}_1\right) & \mathrm{Cov}\left(\mathring{x}_n, \mathring{x}_2\right) & \cdots & \mathrm{Cov}\left(\mathring{x}_n, \mathring{x}_{n-1}\right) & \mathrm{Var}\left(\mathring{x}_n\right) \end{pmatrix}$$

Con esto en mente, es fácil ver que:

$$\mathrm{Var}\left(f\left(\mathring{\vec{x}}\right)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)_i\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)_j\left(\mathrm{Cov}\left(\mathring{\vec{x}}\right)\right)_{i,j} + o\left(3\right) =$$

$$= \sum_{i=1}^{n}\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)_i\left(\sum_{j=1}^{n}\left(\mathrm{Cov}\left(\mathring{\vec{x}}\right)\right)_{i,j}\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)_j\right) + o\left(3\right) =$$

$$= \sum_{i=1}^{n}\left(\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)_i\left(\mathrm{Cov}\left(\mathring{\vec{x}}\right)\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right)_i = \left[\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right)\right]^T \mathrm{Cov}\left(\mathring{\vec{x}}\right)\vec{\nabla} f\left(\mathrm{E}\left(\mathring{\vec{x}}\right)\right) + o\left(3\right)$$

donde hemos usado los subíndices i y j para indicar los elementos de las matrices o vectores correspondientes y T indica «transpuesta». $\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Corolario 2.1. *Sea Ω un abierto de $\mathbb{R}^n$. A continuación, sea $\vec{\mathring{x}} = (\mathring{x}_1, \dots, \mathring{x}_n)$ un vector de variables aleatorias independientes dos a dos y sea $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ una función escalar de clase $C^{(1)}$. Al primer orden no nulo, se tiene:*

$$\mathrm{E}\left[f\left(\vec{\mathring{x}}\right)\right] = f\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right) + o\,(2)$$

$$\mathrm{Var}\left[f\left(\vec{\mathring{x}}\right)\right] = \sum_{i=1}^{n}\left[\frac{\partial f}{\partial \dot{x}_i}\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right)\right]^2 \mathrm{Var}\,(\mathring{x}_i) + o\,(3)$$

Demostración. El primer resultado se obtiene directamente del teorema 2.1 en la página 10. Por ese mismo teorema se tiene:

$$\mathrm{Var}\left(f\left(\vec{\mathring{x}}\right)\right) = \left[\vec{\nabla} f\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right)\right]^T \mathrm{Cov}\left(\vec{\mathring{x}}\right)\vec{\nabla} f\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right) + o\,(3)$$

En nuestro caso, como las variables $\vec{\mathring{x}} = (\mathring{x}_1, \dots, \mathring{x}_n)$ son independientes dos a dos, se tiene:

$$\mathrm{Cov}\,(\mathring{x}_i, \mathring{x}_j) = 0 \qquad \forall i \neq j \ \wedge \ i, j = 1, \dots, n$$

Por tanto, la matriz $\mathrm{Cov}\left(\vec{\mathring{x}}\right)$ es diagonal. De esta forma, el producto vector*matriz*vector queda, sencillamente:

$$\mathrm{Var}\left(f\left(\vec{\mathring{x}}\right)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\vec{\nabla} f\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right)\right)_i \left(\mathrm{Cov}\left(\vec{\mathring{x}}\right)\right)_{i,j}\left(\vec{\nabla} f\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right)\right)_j =$$

$$= \sum_{i=1}^{n}\left[\left(\vec{\nabla} f\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right)\right)_i\right]^2 \left(\mathrm{Cov}\left(\vec{\mathring{x}}\right)\right)_{i,i}$$

ya que $\mathrm{Cov}\,(\mathring{x}_i, \mathring{x}_j) = 0 \ \forall i \neq j$. De esta forma, recordando que $\left(\mathrm{Cov}\left(\vec{\mathring{x}}\right)\right)_{i,i} = \mathrm{Var}\,(\mathring{x}_i)$ y que $\left(\vec{\nabla} f\right)_i = \frac{\partial f}{\partial \dot{x}_i}$, se llega a:

$$\mathrm{Var}\left(f\left(\vec{\mathring{x}}\right)\right) = \sum_{i=1}^{n}\left[\frac{\partial f}{\partial \dot{x}_i}\left(\mathrm{E}\left(\vec{\mathring{x}}\right)\right)\right]^2 \mathrm{Var}\,(\mathring{x}_i)$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

3. Ajustes lineales con errores únicamente en las ordenadas

Suposición 3.1. *Los datos medidos $\{x_i\}_{i=1}^n$ son exactos.*
Suposición 3.2. *Cada dato medido y_i es una muestra de una distribución gaussiana de centro «el valor real» $\tilde{y}_i$ desconocido y de varianza σ_i^2 conocida, que denotaremos $\mathring{y}_i$. Es decir:*

$$\mathring{y}_i \sim \mathcal{N}\left(\tilde{y}_i, \sigma_i^2\right) \qquad \forall i = 1, \dots, n$$

Además, las variables aleatorias asociadas a los datos medidos $\{\mathring{y}_i\}_{i=1}^n$ son independientes dos a dos.
Suposición 3.3. *Los valores reales $\{\tilde{y}_i\}_{i=1}^n$ satisfacen la ecuación:*

$$\tilde{y}_i = a x_i + b \qquad \forall i = 1, \dots, n$$

donde $a, b \in \mathbb{R}$ son dos constantes (los parámetros del modelo).

Proposición 3.1. *Los estimadores de máxima verosimilitud para los parámetros a y b del modelo vienen dados por:*

$$\hat{a} = \frac{\left(\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n} \frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n} \frac{x_i y_i}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2}\right)}$$

$$\hat{b} = \frac{\left(\sum_{i=1}^{n} \frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}\right)\hat{a}}{\left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)}$$

Demostración. Por la suposición 3.2 en la página anterior, sabemos que los datos medidos $\{y_i\}_{i=1}^n$ son muestras de variables aleatorias gaussianas de centro $\tilde{y}_i$ y desviación típica σ_i. Es decir:

$$\mathring{y}_i \sim \mathcal{N}\left(\tilde{y}_i, \sigma_i^2\right) \qquad \forall i = 1, \ldots, n$$

Ahora bien, por la suposición 3.3 en la página anterior, se tiene $\tilde{y}_i = ax_i + b$. Entonces, las variables aleatorias $\{\mathring{y}_i\}_{i=1}^n$ siguen distribuciones normales de centro $ax_i + b$ y de desviación típica σ_i. Dicho de otra forma:

$$\mathring{y}_i \sim \mathcal{N}\left(ax_i + b, \sigma_i^2\right) \qquad \forall i = 1, \ldots, n$$

Sabemos que lo anterior está bien definido ya que los $\{x_i\}_{i=1}^n$ son exactos por la suposición 3.1 en la página anterior.

De esta manera, la función de densidad de probabilidad de cada variable aleatoria $\mathring{y}_i$ viene dada por:

$$f_{\mathring{y}_i}\left(\dot{y}_i\right) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\left(\dot{y}_i - ax_i - b\right)^2}{2\sigma_i^2}\right) \qquad \forall i = 1, \ldots, n$$

donde no debe confundirse la variable aleatoria $\mathring{y}_i$ con el argumento de la función de densidad de probabilidad $f_{\mathring{y}_i}$: $\dot{y}_i$.

A continuación como, por la suposición 3.2 en la página anterior, las variables aleatorias $\{\mathring{y}_i\}_{i=1}^n$ son independientes dos a dos, la función de verosimilitud será simplemente el producto de las funciones de densidad $f_{\mathring{y}_i}$ evaluadas en los datos medidos:

$$L\left(a, b\right) = \prod_{i=1}^{n} f_{\mathring{y}_i}\left(\dot{y}_i = y_i\right) = \frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^{n} \sigma_i^2}} \exp\left(-\sum_{i=1}^{n}\left(\frac{y_i - ax_i - b}{\sqrt{2}\sigma_i}\right)^2\right)$$

donde hemos usado la propiedad $\mathrm{e}^a \mathrm{e}^b = \mathrm{e}^{a+b}$ $\forall a, b \in \mathbb{R}$. Para proseguir, consideramos el logaritmo de la función de verosimilitud:

$$l\left(a, b\right) = \ln L = \ln\left(\frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^{n} \sigma_i^2}}\right) - \sum_{i=1}^{n}\left(\frac{y_i - ax_i - b}{\sqrt{2}\sigma_i}\right)^2$$

donde hemos hecho uso de la propiedad del logaritmo de transformar productos en sumas y hemos empleado, asimismo, que la exponencial y el logaritmo son operaciones inversas.

Con el fin de obtener los operadores de máxima verosimilitud, debemos buscar aquellos puntos en los que l es máxima (que ya sabemos que coincidirán con los de L por ser el logaritmo neperiano una función estrictamente creciente). Para ello, derivamos la función l con respecto a los parámetros a y b:

$$\frac{\partial l}{\partial a} = -\sum_{i=1}^{n} 2\frac{y_i - ax_i - b}{\sqrt{2}\sigma_i}\left(-\frac{x_i}{\sqrt{2}\sigma_i}\right) = \sum_{i=1}^{n} x_i\frac{y_i - ax_i - b}{\sigma_i^2} = \sum_{i=1}^{n} x_i\frac{y_i - ax_i - b}{\sigma_i^2} \tag{3.1}$$

$$\frac{\partial l}{\partial b} = -\sum_{i=1}^{n} 2\frac{y_i - ax_i - b}{\sqrt{2}\sigma_i}\left(-\frac{1}{\sqrt{2}\sigma_i}\right) = \sum_{i=1}^{n} \frac{y_i - ax_i - b}{\sigma_i^2} = \sum_{i=1}^{n} \frac{y_i - ax_i - b}{\sigma_i^2} \tag{3.2}$$

donde hemos hecho uso de que la derivada parcial es una aplicación lineal y de la regla de la cadena. Ahora, imponemos $\vec{\nabla} l = \left(\frac{\partial l}{\partial a}, \frac{\partial l}{\partial b}\right) = \vec{0}$ para obtener el punto crítico de l:

$$\frac{\partial l}{\partial b} = 0 \iff \sum_{i=1}^{n} x_i\frac{y_i - ax_i - b}{\sigma_i^2} = 0$$

$$\frac{\partial l}{\partial a} = 0 \iff \sum_{i=1}^{n} \frac{y_i - ax_i - b}{\sigma_i^2} = 0$$

Expandiendo las sumas anteriores y sacando los parámetros a y b como factor común de los sumatorios, se tiene:

$$\frac{\partial l}{\partial b} = 0 \iff \left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)a - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)b = 0 \tag{3.3}$$

$$\frac{\partial l}{\partial a} = 0 \iff \left(\sum_{i=1}^{n}\frac{x_i y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)a - \left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)b = 0 \tag{3.4}$$

Nuestro objetivo es obtener los valores de a y b que resuelven el sistema de ecuaciones anterior. Para tal fin, aplicaremos el método de la reducción, buscando que el coeficiente de b sea el mismo en ambas ecuaciones. Por consiguiente, multiplicamos la ecuación 3.3 por $\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)$ y la ecuación 3.4 por $\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)$. De esta forma, se obtiene:

$$\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2 a - \left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)b = 0 \tag{3.5}$$

$$\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)a - \left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)b = 0 \tag{3.6}$$

Nótese que, como estábamos buscando, ahora el coeficiente que acompaña a b en ambas ecuaciones es el mismo. De esta forma, restándole a la ecuación 3.5 la ecuación 3.6, se tiene:

$$\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i y_i}{\sigma_i^2}\right) - \left[\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)\right]a = 0$$

ya que el término en b desaparece. De la ecuación anterior resulta razonablemente sencillo despejar a:

$$\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i y_i}{\sigma_i^2}\right) = \left[\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)\right]a \iff$$

$$\iff a = \frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i y_i}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}$$

Conocido el valor de a, podemos simplemente despejar b de la ecuación 3.3 en la página anterior:

$$\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)a-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)b=0 \iff \left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)a=\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)b \iff$$

$$\iff b=\frac{\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)a}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}$$

Por tanto, los estimadores de máxima verosimilitud para los operadores a y b serán:

$$\hat{a}=\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_iy_i}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}$$

$$\hat{b}=\frac{\left(\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\hat{a}}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}$$

Q.E.D.

Proposición 3.2. *Los estimadores $\hat{a}$ y $\hat{b}$ son no sesgados.*

Demostración. Recordemos que un operador $\hat{a}$ es no sesgado si $\mathrm{E}(\hat{a})=a$. Para este cálculo, consideraremos que las muestras y_i que aparecen en las expresiones de la proposición 3.1 en la página 13 representan sus variables aleatorias asociadas $\mathring{y}_i$.

$$\hat{a}=\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{\mathring{y}_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i\mathring{y}_i}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}$$

Como, por la suposición 3.1 en la página 12, los $\{x_i\}_{i=1}^{n}$ son datos exactos, se tiene:

$$\mathrm{E}(\hat{a})=\frac{\mathrm{E}\left[\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{\mathring{y}_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i\mathring{y}_i}{\sigma_i^2}\right)\right]}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}=$$

$$=\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\mathrm{E}\left(\sum_{i=1}^{n}\frac{\mathring{y}_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\mathrm{E}\left(\sum_{i=1}^{n}\frac{x_i\mathring{y}_i}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}=$$

$$=\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{\mathrm{E}(\mathring{y}_i)}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i\mathrm{E}(\mathring{y}_i)}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}$$

A continuación, combinando las suposiciones 3.2 en la página 12 y 3.3 en la página 12, se tiene que:

$$\mathrm{E}(\mathring{y}_i)=\tilde{y}_i=ax_i+b \qquad \forall i=1,\ldots,n \tag{3.7}$$

donde los parámetros a y b que aparecen en la ecuación anterior no son sus estimadores, sino sus valores verdaderos (por eso no llevan circunflejo). De esta forma, se tiene:

$$\mathrm{E}(\hat{a})=\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{ax_i+b}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i(ax_i+b)}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}$$

Operando, llegamos a:

$$\mathrm{E}\left(\hat{a}\right)=\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(a\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)+b\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(a\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)+b\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}=$$

$$=\frac{a\left[\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)\right]+b\left[\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\right]}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}=$$

$$=a\frac{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)}=a$$

De esta forma, efectivamente $\mathrm{E}\left(\hat{a}\right)=a$ y, por tanto, $\hat{a}$ es un operador no sesgado. Vayamos con $\hat{b}$:

$$\hat{b}=\frac{\left(\sum_{i=1}^{n}\frac{\mathring{y}_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\hat{a}}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}$$

De nuevo, para calcular el valor esperado de $\hat{b}$, tenemos que considerar que todo es constante, salvo las variables aleatorias $\{\mathring{y}_i\}_{i=1}^n$ y el estimador $\hat{a}$ (que también es una variable aleatoria). De esta manera:

$$\mathrm{E}\left(\hat{b}\right)=\frac{\mathrm{E}\left[\left(\sum_{i=1}^{n}\frac{\mathring{y}_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\hat{a}\right]}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}=\frac{\mathrm{E}\left(\sum_{i=1}^{n}\frac{\mathring{y}_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\mathrm{E}\left(\hat{a}\right)}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}=$$

$$=\frac{\left(\sum_{i=1}^{n}\frac{\mathrm{E}(\mathring{y}_i)}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\mathrm{E}\left(\hat{a}\right)}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}$$

Al igual que en el caso anterior, combinando las suposiciones 3.2 en la página 12 y 3.3 en la página 12, se tiene la ecuación 3.7 en la página anterior. Además, como hemos probado antes $\mathrm{E}\left(\hat{a}\right)=a$. Así, se obtiene:

$$\mathrm{E}\left(\hat{b}\right)=\frac{\left(\sum_{i=1}^{n}\frac{ax_i+b}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)a}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}=$$

$$=\frac{a\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)+b\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)a}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}=$$

$$=\frac{a\left[\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)\right]+b\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}=b\frac{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}{\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)}=b$$

Y, por consiguiente, $\mathrm{E}\left(\hat{b}\right)=b$ y $\hat{b}$ es un operador no sesgado. $\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 3.3. *La matriz de covarianza de los estimadores $\hat{a},\hat{b}$ satisface:*

$$\mathrm{Cov}\left(\hat{a},\hat{b}\right)\geq\frac{1}{\left(\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\right)\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right)-\left(\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\right)^2}\begin{pmatrix}\sum_{i=1}^{n}\frac{1}{\sigma_i^2} & -\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}\\ -\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2} & \sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2}\end{pmatrix}$$

donde la desigualdad $\geq$ se entiende como $A\geq B\iff A-B$ es semidefinida positiva.

Demostración. En la demostración de la proposición 3.1 en la página 13 llegamos a las siguientes expresiones para las derivadas parciales del logaritmo de la función de verosimilitud (ecuaciones 3.1 en la página 14 y 3.2 en la página 14):

$$\frac{\partial l}{\partial a} = \sum_{i=1}^{n} x_i \frac{y_i - ax_i - b}{\sigma_i^2}$$

$$\frac{\partial l}{\partial b} = \sum_{i=1}^{n} \frac{y_i - ax_i - b}{\sigma_i^2}$$

A continuación, debemos hallar las derivadas segundas:

$$\frac{\partial^2 l}{\partial a^2} = -\sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} \qquad \frac{\partial^2 l}{\partial b^2} = -\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \qquad \frac{\partial^2 l}{\partial a \partial b} = -\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}$$

De esta forma, el hessiano del logaritmo de la función de verosimilitud queda:

$$\mathrm{H}l = \begin{pmatrix} -\sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} & -\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \\ -\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} & -\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \end{pmatrix}$$

Y, por ende, la matriz de información de Fisher resulta:

$$I = \mathrm{E}\left[-\mathrm{H}l\right] = \begin{pmatrix} \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \end{pmatrix}$$

Sabemos que la matriz de información de Fisher satisface la desigualdad:

$$\mathrm{Cov}\left(\hat{a}, \hat{b}\right) \geq I^{-1} = \frac{1}{\left(\sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2}\right)\left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right) - \left(\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}\right)^2} \begin{pmatrix} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} & -\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \\ -\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} \end{pmatrix}$$

donde hemos aplicado que la inversa de una matriz 2x2 viene dada por la fórmula:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 3.4. *La variable aleatoria:*

$$\mathring{Z} = \sum_{i=1}^{n} \frac{\left(\mathring{y}_i - ax_i - b\right)^2}{\sigma_i^2}$$

sigue una distribución χ^2 con $n-2$ grados de libertad.

Demostración. Por la suposición 3.2 en la página 12, $\mathring{y}_i \sim \mathcal{N}\left(\tilde{y}_i, \sigma_i\right)$ $\forall i = 1, \ldots, n$. Además, por la suposición 3.3 en la página 12, se tiene:

$$\mathring{y}_i \sim \mathcal{N}\left(ax_i + b, \sigma_i\right) \qquad \forall i = 1, \ldots, n$$

Por tanto:

$$\mathring{\varepsilon}_i := \frac{\mathring{y}_i - (ax_i + b)}{\sigma_i} = \frac{\mathring{y}_i - ax_i - b}{\sigma_i} \sim \mathcal{N}\left(0, 1\right) \qquad \forall i = 1, \ldots, n$$

por la proposición 1.3 en la página 4. Por otra parte, las expresiones de $\hat{a}$ y $\hat{b}$ obtenidas en la proposición 3.1 en la página 13 imponen dos ligaduras en los posibles valores de las variables aleatorias $\{\mathring{\varepsilon}_i\}_{i=1}^n$. Por tanto, el vector aleatorio $\vec{\mathring{\varepsilon}} \coloneqq (\mathring{\varepsilon}_1, \ldots, \mathring{\varepsilon}_n)$ pertenece a un espacio vectorial de dimensión $n-2$. Por la proposición 1.5 en la página 9:

$$\mathring{Z} = \sum_{i=1}^n \frac{(\mathring{y}_i - a x_i - b)^2}{\sigma_i^2} = \sum_{i=1}^n \mathring{\varepsilon}_i^2 = \left|\left|\vec{\mathring{\varepsilon}}\right|\right|^2 \sim \chi^2_{n-2}$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Implementación en el ordenador 3.1. *Para implementar todo esto en el ordenador resulta más sencillo definirse las cantidades:*

$$S_\sigma \coloneqq \sum_{i=1}^n \frac{1}{\sigma_i^2} \qquad S_x \coloneqq \sum_{i=1}^n \frac{x_i}{\sigma_i^2} \qquad S_{xx} \coloneqq \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \qquad S_y \coloneqq \sum_{i=1}^n \frac{y_i}{\sigma_i^2} \qquad S_{xy} \coloneqq \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2}$$

De esta forma, se obtienen las siguientes expresiones:

$$\hat{a} = \frac{S_x S_y - S_\sigma S_{xy}}{S_x^2 - S_\sigma S_{xx}}$$

$$\hat{b} = \frac{S_y - S_x \hat{a}}{S_\sigma}$$

$$\operatorname{Cov}\left(\hat{a}, \hat{b}\right) \geq \frac{1}{S_{xx} S_\sigma - S_x^2} \begin{pmatrix} S_\sigma & -S_x \\ -S_x & S_{xx} \end{pmatrix}$$

En la práctica, la desigualdad anterior se interpreta como una igualdad por comodidad.

4. Ajustes a funciones arbitrarias con errores únicamente en las ordenadas

Suposición 4.1. *Los datos medidos $\{x_i\}_{i=1}^n$ son exactos.*
Suposición 4.2. *Cada dato medido y_i es una muestra de una distribución gaussiana de centro «el valor real» $\tilde{y}_i$ desconocido y de varianza σ_i^2 conocida, que denotaremos $\mathring{y}_i$. Es decir:*

$$\mathring{y}_i \sim \mathcal{N}\left(\tilde{y}_i, \sigma_i^2\right) \qquad \forall i = 1, \ldots, n$$

Además, las variables aleatorias asociadas a los datos medidos $\{\mathring{y}_i\}_{i=1}^n$ son independientes dos a dos.
Suposición 4.3. *Los valores reales $\{\tilde{y}_i\}_{i=1}^n$ satisfacen la ecuación:*

$$\tilde{y}_i = g\left(x_i, a_1, \ldots, a_m\right) \qquad \forall i = 1, \ldots, n$$

donde $a_1, \ldots, a_m \in \mathbb{R}$ son constantes (los parámetros del modelo) y g es una función $g : \Omega \subseteq \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}$, siendo Ω un conjunto abierto de $\mathbb{R}^{m+1}$.
Proposición 4.1. *Los estimadores de máxima verosimilitud para los parámetros del modelo $(\hat{a}_1, \ldots, \hat{a}_m)$ son los valores de $a_1, \ldots, a_m$ que resuelven el sistema de ecuaciones:*

$$\frac{\partial l}{\partial a_j} = \sum_{i=1}^n \frac{y_i - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i^2} \frac{\partial g}{\partial a_j}\left(x_i, a_1, \ldots, a_m\right) = 0 \qquad \forall j = 1, \ldots, m$$

Demostración. Por la suposición 4.2, sabemos que los datos medidos $\{y_i\}_{i=1}^n$ son muestras de variables aleatorias gaussianas de centro $\tilde{y}_i$ y desviación típica σ_i. Es decir:

$$\mathring{y}_i \sim \mathcal{N}\left(\tilde{y}_i, \sigma_i^2\right) \qquad \forall i = 1, \ldots, n$$

Ahora bien, por la suposición 4.3 en la página anterior, se tiene $\tilde{y}_i = g(x_i, a_1, \ldots, a_m)$. Entonces, las variables aleatorias $\{\mathring{y}_i\}_{i=1}^n$ siguen distribuciones normales de centro $g(x_i, a_1, \ldots, a_m)$ y de desviación típica σ_i. Dicho de otra forma:

$$\mathring{y}_i \sim \mathcal{N}\left(g(x_i, a_1, \ldots, a_m), \sigma_i^2\right) \qquad \forall i = 1, \ldots, n$$

Sabemos que lo anterior está bien definido ya que los $\{x_i\}_{i=1}^n$ son exactos por la suposición 4.1 en la página anterior.

De esta manera, la función de densidad de probabilidad de cada variable aleatoria $\mathring{y}_i$ viene dada por:

$$f_{\mathring{y}_i}(\dot{y}_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(\dot{y}_i - g(x_i, a_1, \ldots, a_m))^2}{2\sigma_i^2}\right) \qquad \forall i = 1, \ldots, n$$

donde no debe confundirse la variable aleatoria $\mathring{y}_i$ con el argumento de la función de densidad de probabilidad $f_{\mathring{y}_i}$: $\dot{y}_i$.

A continuación, como por la suposición 4.2 en la página anterior, las variables aleatorias $\{\mathring{y}_i\}_{i=1}^n$ son independientes dos a dos, la función de verosimilitud será simplemente el producto de las funciones de densidad $f_{\mathring{y}_i}$ evaluadas en los datos medidos:

$$L(a,b) = \prod_{i=1}^n f_{\mathring{y}_i}(\dot{y}_i = y_i) = \frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^n \sigma_i^2}} \exp\left(-\sum_{i=1}^n \left(\frac{y_i - g(x_i, a_1, \ldots, a_m)}{\sqrt{2}\sigma_i}\right)^2\right)$$

donde hemos usado la propiedad $\mathrm{e}^a\mathrm{e}^b = \mathrm{e}^{a+b}\ \forall a, b \in \mathbb{R}$. Para proseguir, consideramos el logaritmo de la función de verosimilitud:

$$l(a,b) = \ln L = \ln\left(\frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^n \sigma_i^2}}\right) - \sum_{i=1}^n \left(\frac{y_i - g(x_i, a_1, \ldots, a_m)}{\sqrt{2}\sigma_i}\right)^2$$

donde hemos hecho uso de la propiedad del logaritmo de transformar productos en sumas y hemos empleado, asimismo, que la exponencial y el logaritmo son operaciones inversas.

Con el fin de obtener los operadores de máxima verosimilitud, debemos buscar aquellos puntos en los que l es máxima (que ya sabemos que coincidirán con los de L por ser el logaritmo neperiano una función estrictamente creciente). Para ello, derivamos la función l con respecto a cada parámetro a_j:

$$\frac{\partial l}{\partial a_j} = -\sum_{i=1}^n 2\frac{y_i - g(x_i, a_1, \ldots, a_m)}{\sqrt{2}\,\sigma_i}\left(-\frac{1}{\sqrt{2}\,\sigma_i}\right)\frac{\partial g}{\partial a_j}(x_i, a_1, \ldots, a_m) =$$

$$= \sum_{i=1}^n \frac{y_i - g(x_i, a_1, \ldots, a_m)}{\sigma_i^2}\frac{\partial g}{\partial a_j}(x_i, a_1, \ldots, a_m)$$

Los estimadores de máxima verosimilitud $\hat{a}_1, \ldots, \hat{a}_m$ son los valores de $a_1, \ldots, a_m$ tales que:

$$\frac{\partial l}{\partial a_j} = 0 \qquad \forall j = 1, \ldots, m \iff$$

$$\sum_{i=1}^n \frac{y_i - g(x_i, a_1, \ldots, a_m)}{\sigma_i^2}\frac{\partial g}{\partial a_j}(x_i, a_1, \ldots, a_m) = 0 \qquad \forall j = 1, \ldots, m$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 4.2. *La matriz de información de Fisher viene dada por:*

$$I_{j,k} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k} \left(x_i, a_1, \ldots, a_m\right) \frac{\partial g}{\partial a_j} \left(x_i, a_1, \ldots, a_m\right) \qquad \forall j, k = 1, \ldots, m$$

Demostración. Recuperamos la expresión para la derivada del logaritmo de la función de verosimilitud respecto a un parámetro a_j dada en la proposición 4.1 en la página 18:

$$\frac{\partial l}{\partial a_j} = \sum_{i=1}^{n} \frac{y_i - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i^2} \frac{\partial g}{\partial a_j} \left(x_i, a_1, \ldots, a_m\right)$$

Volviendo a derivar con respecto a a_k, se tiene:

$$\frac{\partial^2 l}{\partial a_k \partial a_j} = \sum_{i=1}^{n} \left[-\frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k} \left(x_i, a_1, \ldots, a_m\right) \frac{\partial g}{\partial a_j} \left(x_i, a_1, \ldots, a_m\right) + \right.$$

$$\left. + \frac{y_i - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i^2} \frac{\partial^2 g}{\partial a_k \partial a_j} \left(x_i, a_1, \ldots, a_m\right) \right]$$

Ahora bien, sabemos que el elemento j, k de la matriz de información de Fisher viene dado por:

$$I_{j,k} = \mathrm{E}\left(- \left(\mathrm{H}l\right)_{j,k} \right) = \mathrm{E}\left(- \frac{\partial^2 l}{\partial a_k \partial a_j} \right)$$

donde en la parcial segunda correspondiente hay que sustituir y_i por $\mathring{y}_i$, su variable aleatoria correspondiente.

$$I_{j,k} = \mathrm{E}\left[\sum_{i=1}^{n} \left[\frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k} \left(x_i, a_1, \ldots, a_m\right) \frac{\partial g}{\partial a_j} \left(x_i, a_1, \ldots, a_m\right) + \right. \right.$$

$$\left. \left. - \frac{\mathring{y}_i - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i^2} \frac{\partial^2 g}{\partial a_k \partial a_j} \left(x_i, a_1, \ldots, a_m\right) \right] \right] =$$

$$= \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k} \left(x_i, a_1, \ldots, a_m\right) \frac{\partial g}{\partial a_j} \left(x_i, a_1, \ldots, a_m\right) +$$

$$- \sum_{i=1}^{n} \frac{\mathrm{E}\left(\mathring{y}_i\right) - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i^2} \frac{\partial^2 g}{\partial a_k \partial a_j} \left(x_i, a_1, \ldots, a_m\right)$$

Combinando las suposiciones 4.2 en la página 18 y 4.3 en la página 18, se tiene que:

$$\mathrm{E}\left(\mathring{y}_i\right) = g\left(x_i, a_1, \ldots, a_m\right) \qquad \forall i = 1, \ldots, n$$

Y, por tanto, el elemento j, k de la matriz de información de Fisher queda, simplemente:

$$I_{j,k} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k} \left(x_i, a_1, \ldots, a_m\right) \frac{\partial g}{\partial a_j} \left(x_i, a_1, \ldots, a_m\right)$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Proposición 4.3. *La variable aleatoria:*

$$\mathring{Z} = \sum_{i=1}^{n} \frac{\left(\mathring{y}_i - g\left(x_i, a_1, \ldots, a_m\right)\right)^2}{\sigma_i^2}$$

sigue una distribución χ^2 *con* $n - m$ *grados de libertad.*

Demostración. Por la suposición 4.2 en la página 18, $\mathring{y}_i \sim \mathcal{N}\left(\tilde{y}_i, \sigma_i^2\right)$ $\forall i = 1, \ldots, n$. Además, por la suposición 4.3 en la página 18, se tiene:

$$\mathring{y}_i \sim \mathcal{N}\left(g\left(x_i, a_1, \ldots, a_m\right), \sigma_i^2\right) \qquad \forall i = 1, \ldots, n$$

Por tanto:

$$\mathring{\varepsilon}_i := \frac{\mathring{y}_i - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i} \sim \mathcal{N}\left(0, 1\right) \qquad \forall i = 1, \ldots, n$$

por la proposición 1.3 en la página 4. Por otra parte, las ecuaciones obtenidas en la proposición 4.1 en la página 18 imponen m ligaduras en los posibles valores de las variables aleatorias $\{\mathring{\varepsilon}_i\}_{i=1}^n$. Por tanto, el vector aleatorio $\mathring{\vec{\varepsilon}} := \left(\mathring{\varepsilon}_1, \ldots, \mathring{\varepsilon}_n\right)$ pertenece a un espacio vectorial de dimensión $n - m$. Por la proposición 1.5 en la página 9:

$$\mathring{Z} = \sum_{i=1}^{n} \frac{\left(\mathring{y}_i - g\left(x_i, a_1, \ldots, a_m\right)\right)^2}{\sigma_i^2} = \sum_{i=1}^{n} \mathring{\varepsilon}_i^2 = \left\|\mathring{\vec{\varepsilon}}\right\|^2 \sim \chi^2_{n-m}$$

$\mathbb{Q}.\mathbb{E}.\mathbb{D}.$

Implementación en el ordenador 4.1. *Este caso resulta más complicado que los ajustes lineales puesto que no hay solución analítica. El método a seguir es el siguiente:*

1. *Se calculan analíticamente (si se puede) las derivadas de la función g:* $\frac{\partial g}{\partial x_j}$ $\forall j = 1, \ldots, m$.
2. *Se resuelven numéricamente las ecuaciones 4.1 en la página 18.*

$$\frac{\partial l}{\partial a_j} = \sum_{i=1}^{n} \frac{y_i - g\left(x_i, a_1, \ldots, a_m\right)}{\sigma_i^2} \frac{\partial g}{\partial a_j}\left(x_i, a_1, \ldots, a_m\right) = 0 \qquad \forall j = 1, \ldots, m$$

3. *Se calcula la matriz de información de Fisher mediante la expresión dada en la proposición 4.2 en la página anterior:*

$$I_{j,k} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}\left(x_i, a_1, \ldots, a_m\right) \frac{\partial g}{\partial a_j}\left(x_i, a_1, \ldots, a_m\right) \qquad \forall j, k = 1, \ldots, m$$

4. *Se invierte la matriz de información de Fisher numéricamente para obtener la matriz de covarianza:*

$$\mathrm{Cov}\left(\hat{a}_1, \ldots, \hat{a}_m\right) = I^{-1}$$