6. Aspectos Computacionales.

La implementación práctica de las técnicas Bayesianas usualmente requiere de un esfuerzo computacional muy alto. La mayor parte de este esfuerzo se requiere para calcular ciertas características de la distribución final del parámetro de interés que permitan resumir de alguna forma la información contenida en ella. En muchos casos estos "resúmenes inferenciales" se reducen a integrales de la forma

$$S_I[\, q(\underset{\sim}{\theta})\,] = \int q(\underset{\sim}{\theta}) \cdot p(\underset{\sim}{\theta}) \cdot p(\underset{\sim}{x} \mid \underset{\sim}{\theta})\, d\underset{\sim}{\theta}_{I^c},$$

donde $\underset{\sim}{x} = (x_1, \ldots, x_n)'$, $\theta = (\theta_1, \ldots, \theta_d)'$, $I \subseteq \{1, \ldots, d\}$, $I^c = \{1, \ldots, d\} \setminus I$ y $\underset{\sim}{\theta}_{I^c} = \{ \theta_i : i \in I^c \}$.

Ejemplos:

- $p(\underset{\sim}{x}) = S_\emptyset[\,1\,]$.

- $E[\underset{\sim}{\theta} / \underset{\sim}{x}] = S_\emptyset[\,\underset{\sim}{\theta}\,] / S_\emptyset[1]$.

- $Var[\underset{\sim}{\theta} / \underset{\sim}{x}] = S_\emptyset[\,\underset{\sim}{\theta}\underset{\sim}{\theta}'\,] / S_\emptyset[1] - \{ S_\emptyset[\underset{\sim}{\theta}]\, S_\emptyset[\underset{\sim}{\theta}]' \} / S_\emptyset[1]^2$

- Sea $\underset{\sim}{\theta} = (\underset{\sim}{\theta}_1', \underset{\sim}{\theta}_2')'$, con $\underset{\sim}{\theta}_1 \in \mathbb{R}^{d_1}$ y $\underset{\sim}{\theta}_2 \in \mathbb{R}^{d-d_1}$, y sea $D_1 = \{1, \ldots, d_1\}$. Entonces

$$p(\underset{\sim}{\theta}_1 / \underset{\sim}{x}) = S_{D_1}[1] / S_\emptyset[1].$$

- $P[\,\underset{\sim}{\theta} \in R\,] = S_\emptyset[\,\mathbb{1}_R(\underset{\sim}{\theta})\,] / S_\emptyset[1]$, con $\mathbb{1}_R(\underset{\sim}{\theta}) = \begin{cases} 1 & \theta \in R \\ 0 & \theta \in R^c \end{cases}$.

Desafortunadamente, en la práctica la dimensión de θ puede ser muy grande, y tanto $p(\theta)$ como $p(x/\tilde{\theta})$ pueden tener formas muy complicadas, por lo que en la mayoría de los casos estas integrales no pueden resolverse de manera analítica. Es por ello que es crucial contar con métodos analíticos y/o numéricos que nos permitan aproximar y/o calcular integrales de manera eficiente.

6.1. Aproximación de Laplace.

En esta sección discutiremos algunas aproximaciones analíticas basadas en argumentos asintóticos, haciendo énfasis en la aproximación de Laplace.

* Órdenes de magnitud.

Sean $f: \mathbb{R} \to \mathbb{R}$ y $g: \mathbb{R} \to \mathbb{R}$ funciones continuas.

- $f = O(g)$ $(x \to \infty)$ si y sólo si existen x_0 y M_0 tales que

$$\frac{f(x)}{g(x)} < M_0 \quad \text{para toda } x > x_0.$$

Ejemplo: $a/(b+n) = O(n^{-1}) \quad (n \to \infty)$

- $f = o(g)$ $(x \to \infty)$ si y sólo si

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0.$$

Ejemplo: $a/(b+n)^2 = o(n^{-1}) \quad (n \to \infty)$

Notación:
- $f = h + O(g) \iff (f-h) = O(g)$
- $f = h\{1 + O(g)\} \iff f/h = 1 + O(g)$
- Similarmente para $o(\cdot)$.

* Aproximación normal.

Aquí discutiremos el caso uniparametral $(d=1)$ con cierto detalle. El argumento es totalmente análogo en el caso multiparametral.

Sea $p_x(\theta) = p(\theta)\, p(\underset{\sim}{x}|\theta)$. Desarrollando

$$\log p_x(\theta) = \log p(\theta) + \log p(\underset{\sim}{x}|\theta)$$

en serie de Taylor alrededor de $\hat{\theta}$, la moda de $p_x(\theta)$, se tiene

$$\log p_x(\theta) = \log p_x(\hat{\theta}) + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{\partial^k \log p_x(\hat{\theta})}{\partial \theta^k} (\theta - \hat{\theta})^k .$$

Más adelante veremos que las aproximaciones se requieren principalmente en regiones tales que $(\theta - \hat{\theta})$ es de orden $O(n^{-1/2})$, por lo que en la discusión siguiente supondremos que

$$(\theta - \hat{\theta}) = O(n^{-1/2}).$$

Por otra parte, notemos que

$$\log p_x(\theta) = \log p(\theta) + \sum_{i=1}^{n} \log p(x_i|\theta),$$

de donde

$$\frac{\partial^k \log p_x(\theta)}{\partial \theta^k} = \frac{\partial^k \log p(\theta)}{\partial \theta^k} + \sum_{i=1}^{n} \frac{\partial^k \log p(x_i|\theta)}{\partial \theta^k}$$

$$= O(n) \qquad (n \to \infty) \qquad \forall\ k = 1, 2, \ldots$$

Por lo tanto

$$\frac{1}{k!} \frac{\partial^k \log p_x(\hat{\theta})}{\partial \theta^k} (\theta - \hat{\theta})^k = O(n^{1-k/2}) \qquad \forall\ k = 1, 2, \ldots$$

Sea
$$V(\theta) = -\left\{ \frac{\partial^2 \log p_x(\theta)}{\partial \theta^2} \right\}^{-1},$$

y notemos que
$$\frac{\partial \log p_x(\hat{\theta})}{\partial \theta} = 0.$$

Entonces
$$\log p_x(\theta) = \log p_x(\hat{\theta}) - \frac{1}{2V(\hat{\theta})}(\theta - \hat{\theta})^2 + O(n^{-1/2}).$$

Ahora
$$p_x(\theta) = p_x(\hat{\theta}) \cdot \exp\left\{ -\frac{1}{2V(\hat{\theta})}(\theta-\hat{\theta})^2 \right\} \cdot \exp\left\{ O(n^{-1/2}) \right\}$$
$$\overset{(1)}{=} p_x(\hat{\theta}) \cdot \exp\left\{ -\frac{1}{2V(\hat{\theta})}(\theta-\hat{\theta})^2 \right\} \left\{ 1 + O(n^{-1/2}) \right\}$$

(1) Nota:
$$e^x = 1 + x + \frac{1}{2}x^2 + \cdots + \frac{1}{k!}x^k + \cdots$$
i.e.
$$e^x = 1 + x + O(x^2) \qquad (x \to 0)$$

- Constante de normalización:

$$p(\theta | \underset{\sim}{x}) = \frac{p_x(\theta)}{p(\underset{\sim}{x})},$$

donde
$$p(\underset{\sim}{x}) = \int p_x(\theta)\, d\theta$$
$$= \int p(\theta) \cdot p(\underset{\sim}{x} | \theta)\, d\theta.$$

Consideremos la expansión

$$\log p_x(\theta) = \log p_x(\hat\theta) - \frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2 + \frac{1}{6}\frac{\partial^3 \log p_x(\hat\theta)}{\partial\theta^3}(\theta-\hat\theta)^3 + O(n^{-1}).$$

Sea $t(\theta) = \frac{\partial^3 \log p_x(\hat\theta)}{\partial\theta^3}(\theta-\hat\theta)^3$. Entonces

$$p_x(\theta) = p_x(\hat\theta)\cdot \exp\left\{-\frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2\right\}\cdot \exp\{t(\theta)/6\}\cdot \exp\{O(n^{-1})\}$$

$$\overset{(1)}{=} p_x(\hat\theta)\cdot \exp\left\{-\frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2\right\}\{1+t(\theta)/6+O(n^{-1})\}\{1+O(n^{-1})\}$$

ya que $t(\theta) = O(n^{-1/2})$.

Ahora

$$p(\underset{\sim}{x}) = \int p_x(\theta)\, d\theta$$

$$= p_x(\hat\theta)\int\{1+t(\theta)/6+O(n^{-1})\}\cdot \exp\left\{-\frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2\right\}d\theta\cdot\{1+O(n^{-1})\}.$$

Pero

$$\int \exp\left\{-\frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2\right\}d\theta = (2\pi)^{1/2}\, V(\hat\theta)^{1/2}$$

y

$$\int t(\theta)\cdot \exp\left\{-\frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2\right\}d\theta =$$

$$\frac{\partial^3 \log p_x(\hat\theta)}{\partial\theta^3}\underbrace{\int(\theta-\hat\theta)^3 \exp\left\{-\frac{1}{2V(\hat\theta)}(\theta-\hat\theta)^2\right\}d\theta}_{0}$$

$$= 0.$$

Por lo tanto

$$p(\underset{\sim}{x}) = \left\{ p_x(\hat{\theta}) \cdot (2\pi)^{1/2} V(\hat{\theta})^{1/2} + O(n^{-1}) \right\} \cdot \left\{ 1 + O(n^{-1}) \right\}$$

$$\overset{(2)}{=} p_x(\hat{\theta}) \cdot (2\pi)^{1/2} V(\hat{\theta})^{1/2} \left\{ 1 + O(n^{-1}) \right\} \qquad (n \to \infty).$$

(2) Nota: Si a es una función acotada de n entonces

$$\{ a + O(n^{-1}) \} \{ 1 + O(n^{-1}) \} = a \{ 1 + O(n^{-1}) \}.$$

Demostración: notemos que si a es una función acotada de n entonces

$$a \cdot O(n^{-1}) = O(n^{-1}),$$

de donde

$$\{ a + O(n^{-1}) \} \{ 1 + O(n^{-1}) \} =$$

$$a \{ 1 + O(n^{-1}) \} + O(n^{-1}) + O(n^{-2})$$

$$= a \{ 1 + O(n^{-1}) \} + O(n^{-1})$$

$$= a \{ 1 + O(n^{-1}) \} + a \cdot O(n^{-1})$$

$$= a \{ 1 + O(n^{-1}) \}.$$

Finalmente

$$p(\theta \mid \underset{\sim}{x}) = \frac{p_x(\hat{\theta}) \cdot \exp\left\{ -\frac{1}{2V(\hat{\theta})} (\theta - \hat{\theta})^2 \right\} \left\{ 1 + O(n^{-1/2}) \right\}}{p_x(\hat{\theta}) \cdot (2\pi)^{1/2} V(\hat{\theta})^{1/2} \left\{ 1 + O(n^{-1}) \right\}}$$

$$= (2\pi)^{-1/2} \cdot V(\hat{\theta})^{-1/2} \exp\left\{ -\frac{1}{2V(\hat{\theta})} (\theta - \hat{\theta})^2 \right\} \left\{ 1 + O(n^{-1/2}) \right\}$$

$$= N(\theta \mid \hat{\theta}, V(\hat{\theta})) \left\{ 1 + O(n^{-1/2}) \right\} \qquad (n \to \infty).$$

Esta es una versión de la aproximación asintótica normal a la distribución final.

Recordemos que

$$\frac{\partial^2 \log p_x(\theta)}{\partial \theta^2} = O(n),$$

lo cual implica que $V(\theta) = O(n^{-1})$. Se tiene entonces que la desviación estándar asintótica, $\sqrt{V(\hat{\theta})}$, es de orden $O(n^{-1/2})$. Así, las aproximaciones se requieren principalmente en regiones tales que $(\theta - \hat{\theta})$ es de orden $O(n^{-1/2})$, puesto que fuera de estas regiones $p(\theta \mid \underset{\sim}{x})$ es muy pequeña.

En el caso multiparametral, un argumento análogo permite demostrar que

$$p(\underset{\sim}{\theta} \mid \underset{\sim}{x}) = N_d(\underset{\sim}{\theta} \mid \hat{\underset{\sim}{\theta}}, V(\hat{\theta})) \cdot \{1 + O(n^{-1/2})\},$$

donde $\hat{\underset{\sim}{\theta}} \in \mathbb{R}^d$ es la moda de $p(\underset{\sim}{\theta} \mid \underset{\sim}{x})$ y

$$V(\theta) = -\left\{ \frac{\partial^2 \log p(\underset{\sim}{\theta} \mid \underset{\sim}{x})}{\partial \underset{\sim}{\theta}' \partial \underset{\sim}{\theta}} \right\}^{-1}.$$

Si esta aproximación es adecuada, prácticamente cualquier resumen inferencial de interés puede aproximarse fácilmente. Por ejemplo,

- $E[\underset{\sim}{\theta} \mid \underset{\sim}{x}] \approx \hat{\theta}$
- $Var[\underset{\sim}{\theta} \mid \underset{\sim}{x}] \approx V(\hat{\underset{\sim}{\theta}})$
- Sea $\underset{\sim}{\theta} = (\underset{\sim}{\theta}_1', \underset{\sim}{\theta}_2')'$, con $\underset{\sim}{\theta}_1 \in \mathbb{R}^{d_1}$ y $\underset{\sim}{\theta}_2 \in \mathbb{R}^{d-d_1}$.

  Si $\hat{\underset{\sim}{\theta}} = (\hat{\underset{\sim}{\theta}}_1', \hat{\underset{\sim}{\theta}}_2')'$ y

$$V(\hat{\theta}) = \begin{bmatrix} V_{11}(\hat{\theta}) & V_{12}(\hat{\theta}) \\ V_{21}(\hat{\theta}) & V_{22}(\hat{\theta}) \end{bmatrix}$$

denotan la partición correspondiente de la media y varianza asintóticas, respectivamente, entonces

$$p(\underset{\sim}{\theta}_1 | \underset{\sim}{x}) \approx N(\underset{\sim}{\theta}_1 | \underset{\sim}{\hat{\theta}}_1, V_{11}(\hat{\theta})).$$

Desafortunadamente, en aplicaciones específicas no siempre es fácil determinar si la aproximación normal es adecuada para el tamaño de muestra dado.

* Aproximación de Laplace (forma estándar).

Supongamos que se desea calcular una integral de la forma

$$S_\phi[q(\underset{\sim}{\theta})] = \int q(\underset{\sim}{\theta}) \cdot \exp\{-n\,h(\underset{\sim}{\theta})\}\,d\underset{\sim}{\theta},$$

donde $h: \mathbb{R}^d \to \mathbb{R}$ es una función suave, con un mínimo en $\underset{\sim}{\hat{\theta}}$, y $q: \mathbb{R}^d \to \mathbb{R}$ es una función suave.

El método de Laplace aproxima $S_\phi[q(\underset{\sim}{\theta})]$ a través de

$$\hat{S}_\phi[q(\underset{\sim}{\theta})] = q(\underset{\sim}{\hat{\theta}}) \cdot (2\pi/n)^{d/2} |\Sigma(\underset{\sim}{\hat{\theta}})|^{1/2} \cdot \exp\{-n\,h(\underset{\sim}{\hat{\theta}})\},$$

donde

$$\Sigma(\underset{\sim}{\theta}) = \left\{ \frac{\partial^2 h(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}' \partial \underset{\sim}{\theta}} \right\}^{-1}.$$

De hecho

$$\hat{S}_\phi[q(\underset{\sim}{\theta})] = S_\phi[q(\underset{\sim}{\theta})]\,\{1 + O(n^{-1})\} \qquad (n \to \infty).$$

La derivación de esta aproximación es similar a la de la aproximación normal, y se basa en la expansión en serie de Taylor, alrededor de $\underset{\sim}{\hat{\theta}}$, tanto de $h(\underset{\sim}{\theta})$ como de $q(\underset{\sim}{\theta})$. Aquí discutiremos con cierto detalle sólo el caso uniparametral $(d=1)$. El caso multiparametral es totalmente análogo. Como en el caso de la aproximación normal, aquí supondremos que

$$(\theta - \hat{\theta}) = O(n^{-1/2}).$$

Desarrollando $n\,h(\theta)$ alrededor de $\hat{\theta}$, tenemos

$$n\,h(\theta) = n\,h(\hat{\theta}) + \sum_{k=1}^{\infty} \frac{n}{k!} \frac{\partial^k h(\hat{\theta})}{\partial \theta^k} (\theta - \hat{\theta})^k .$$

Ahora bien,

$$\frac{n}{k!} \frac{\partial^k h(\theta)}{\partial \theta^k} = O(n) \quad (n \to \infty) \qquad \forall\, k = 1, 2, \ldots,$$

de donde

$$\frac{n}{k!} \frac{\partial^k h(\hat{\theta})}{\partial \theta^k} (\theta - \hat{\theta})^k = O(n^{1-k/2}) \qquad \forall\, k = 1, 2, \ldots$$

Por lo tanto

$$n\,h(\theta) = n\,h(\hat{\theta}) + \frac{n}{2\Sigma(\hat{\theta})} (\theta - \hat{\theta})^2 + \frac{n\,\tau(\theta)}{6} + O(n^{-1}),$$

donde $\Sigma(\theta) = \left\{ \frac{\partial^2 h(\theta)}{\partial \theta^2} \right\}^{-1}$ y $\tau(\theta) = \frac{\partial^3 h(\hat{\theta})}{\partial \theta^3} (\theta - \hat{\theta})^3$.

En otras palabras

$$\exp\{-n\,h(\theta)\} = \exp\{-n\,h(\hat{\theta})\} \cdot \exp\left\{ - \frac{n}{2\Sigma(\hat{\theta})} (\theta - \hat{\theta})^2 \right\}$$

$$\times \exp\left\{ - \frac{n\,\tau(\theta)}{6} \right\} \cdot \exp\{ O(n^{-1}) \}$$

$$\overset{(1)}{=} \exp\{-n\,h(\hat{\theta})\} \cdot \exp\left\{ - \frac{n}{2\Sigma(\hat{\theta})} (\theta - \hat{\theta})^2 \right\}$$

$$\times \left\{ 1 - \frac{n\,\tau(\theta)}{6} + O(n^{-1}) \right\} \cdot \left\{ 1 + O(n^{-1}) \right\}$$

De manera similar,

$$q(\theta) = q(\hat{\theta}) + q'(\hat{\theta})\,(\theta-\hat{\theta}) + O(n^{-1}),$$

donde $\quad q'(\theta) = \dfrac{\partial q(\theta)}{\partial \theta}$.

Entonces el integrando de $S_\theta[q(\theta)]$ puede escribirse como

$$q(\theta)\exp\{-n\,h(\theta)\} =$$

$$\left\{q(\hat{\theta}) + q'(\hat{\theta})(\theta-\hat{\theta}) + O(n^{-1})\right\}\left\{1 - \frac{n\,\tau(\theta)}{6} + O(n^{-1})\right\}$$

$$\times\ \exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}\cdot \exp\{-n\,h(\hat{\theta})\}\cdot\left\{1+O(n^{-1})\right\}$$

$$= \left\{q(\hat{\theta}) + q'(\hat{\theta})(\theta-\hat{\theta}) - \frac{n\,q(\hat{\theta})\,\tau(\theta)}{6} - \frac{n\,q'(\hat{\theta})\,\tau(\theta)}{6}(\theta-\hat{\theta}) + O(n^{-1})\right\}$$

$$\times\ \exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}\cdot \exp\{-n\,h(\hat{\theta})\}\cdot\left\{1+O(n^{-1})\right\}.$$

Pero

$$\int \exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}d\theta = (2\pi/n)^{1/2}\,\Sigma(\hat{\theta})^{1/2},$$

$$\int (\theta-\hat{\theta})\cdot\exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}d\theta = 0,$$

$$\int \tau(\theta)\cdot\exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}d\theta =$$

$$\frac{\partial^3 h(\hat{\theta})}{\partial\theta}\ \underbrace{\int (\theta-\hat{\theta})^3\cdot\exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}d\theta}_{0}$$

$$= 0$$

y

$$\int n\,\tau(\theta)(\theta-\hat{\theta})\cdot \exp\left\{-\frac{n}{2\Sigma(\hat{\theta})}(\theta-\hat{\theta})^2\right\}d\theta = O(n^{-1}).$$

Por lo tanto

$$S_\phi[q(\theta)] = \left\{q(\hat{\theta})\cdot(2\pi/n)^{1/2}\,\Sigma(\hat{\theta})^{1/2}\exp\{-nh(\hat{\theta})\}+O(n^{-1})\right\}\left\{1+O(n^{-1})\right\}$$

$$\overset{(2)}{\equiv} q(\hat{\theta})\cdot(2\pi/n)^{1/2}\,\Sigma(\hat{\theta})^{1/2}\cdot\exp\{-nh(\hat{\theta})\}\cdot\left\{1+O(n^{-1})\right\}$$

$$= \hat{S}_\phi[q(\theta)]\cdot\left\{1+O(n^{-1})\right\} \qquad (n\to\infty).$$

— o —

Notemos que, en general, una integral dada

$$\int f(\underset{\sim}{\theta})\,d\underset{\sim}{\theta}$$

puede escribirse como

$$\int q(\underset{\sim}{\theta})\cdot\exp\{-nh(\theta)\}\,d\underset{\sim}{\theta}$$

para distintas funciones $q(\underset{\sim}{\theta})$ y $h(\underset{\sim}{\theta})$. Para n fijo, la precisión de la aproximación de Laplace depende de la elección particular de $q(\underset{\sim}{\theta})$ y $h(\underset{\sim}{\theta})$, así como de la parametrización que se utilice.

Es posible lograr una mayor precisión si se mantienen algunos términos de orden mayor en la expansión en serie de Taylor, pero esto requiere a su vez del cálculo de derivadas de orden mayor.

// En ocasiones $q(\underset{\sim}{\theta})$ y $h(\underset{\sim}{\theta})$ se eligen de manera que tanto las derivadas de $h(\cdot)$ como el valor de $\hat{\theta}$ puedan encontrarse fácilmente. De esta manera se obtiene una aproximación rápida y simple, aunque no siempre suficientemente precisa. //

Ejemplo: Sea $h(\underset{\sim}{\theta}) = -\frac{1}{n} \log p_x(\underset{\sim}{\theta})$, de manera que

$$p_x(\underset{\sim}{\theta}) = \exp\{-n\, h(\underset{\sim}{\theta})\}.$$

Se desea calcular

$$E[q(\underset{\sim}{\theta}) \mid \underset{\sim}{x}] = S_\phi[q(\underset{\sim}{\theta})] \Big/ S_\phi[1].$$

$$\hat{S}_\phi[q(\underset{\sim}{\theta})] = q(\hat{\underset{\sim}{\theta}})\ (2\pi/n)^{d/2} \, |\Sigma(\hat{\underset{\sim}{\theta}})|^{1/2} \cdot \exp\{-n\, h(\hat{\underset{\sim}{\theta}})\}$$

y $$\hat{S}_\phi[1] = (2\pi/n)^{d/2} \, |\Sigma(\hat{\underset{\sim}{\theta}})|^{1/2} \cdot \exp\{-n\, h(\hat{\underset{\sim}{\theta}})\}.$$

Por lo tanto

$$\hat{E}[q(\underset{\sim}{\theta}) \mid \underset{\sim}{x}] = q(\hat{\underset{\sim}{\theta}}).$$

De hecho,

$$\hat{E}[q(\underset{\sim}{\theta}) \mid \underset{\sim}{x}] = E[q(\underset{\sim}{\theta}) \mid \underset{\sim}{x}] \cdot \{1 + O(n^{-1})\}$$

* Aproximación de Laplace (forma exponencial).

Notemos que, si $q(\underset{\sim}{\theta})$ es una función positiva, la integral

$$S_\phi[q(\underset{\sim}{\theta})] = \int q(\underset{\sim}{\theta}) \exp\{-n\, h(\theta)\}\, d\underset{\sim}{\theta}$$

también puede escribirse como

$$S_\phi[q(\underset{\sim}{\theta})] = \int q^*(\underset{\sim}{\theta}) \exp\{-n\, h^*(\underset{\sim}{\theta})\}\, d\underset{\sim}{\theta},$$

con $q^*(\underset{\sim}{\theta}) \equiv 1$ y

$$h^*(\underset{\sim}{\theta}) = -\frac{1}{n} \log q(\underset{\sim}{\theta}) + h(\theta)$$

En otras palabras, podemos escribir la integral como

$$S_\phi[q(\underset{\sim}{\theta})] = \int \exp\{-n\, h^*(\underset{\sim}{\theta})\}\, d\underset{\sim}{\theta}.$$

Aplicando ahora la aproximación de Laplace se tiene

$$\tilde{S}_\phi[q(\underset{\sim}{\theta})] = (2\pi/n)^{d/2}\, |\Sigma^*(\hat{\underset{\sim}{\theta}}^*)|^{1/2} \cdot \exp\{-n\, h^*(\hat{\underset{\sim}{\theta}}^*)\}$$

$$= q(\hat{\underset{\sim}{\theta}}^*) \cdot (2\pi/n)^{d/2} |\Sigma^*(\hat{\underset{\sim}{\theta}}^*)|^{1/2} \cdot \exp\{-n\, h(\hat{\underset{\sim}{\theta}}^*)\},$$

donde $\hat{\underset{\sim}{\theta}}^*$ es el valor de $\underset{\sim}{\theta}$ que minimiza $h^*(\underset{\sim}{\theta})$ y

$$\Sigma^*(\underset{\sim}{\theta}) = \left\{ \frac{\partial^2 h^*(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}'\, \partial \underset{\sim}{\theta}} \right\}^{-1}.$$

Este es un caso particular importante y se conoce como la <u>forma exponencial</u> de la aproximación de Laplace.

<u>Ejemplo</u>: Sea $h(\underset{\sim}{\theta}) = -\frac{1}{n} \log p_x(\underset{\sim}{\theta})$, de manera que

$$p_x(\underset{\sim}{\theta}) = \exp\{-n\, h(\underset{\sim}{\theta})\}$$

y

$$h^*(\underset{\sim}{\theta}) = -\frac{1}{n} \{\log q(\underset{\sim}{\theta}) + \log p_x(\underset{\sim}{\theta})\}.$$

Se desea calcular

$$E[\, q(\underset{\sim}{\theta}) \mid \underset{\sim}{x}\,] = S_\phi[q(\underset{\sim}{\theta})] \,/\, S_\phi[1].$$

$$\tilde{S}_\phi[q(\underset{\sim}{\theta})] = q(\hat{\theta}^*)\cdot(2\pi/n)^{d/2}\cdot|\Sigma^*(\hat{\underset{\sim}{\theta}}^*)|^{1/2}\cdot\exp\{-nh(\hat{\underset{\sim}{\theta}}^*)\}$$

y

$$\tilde{S}_\phi[1] = (2\pi/n)^{d/2}\cdot|\Sigma(\hat{\underset{\sim}{\theta}})|^{1/2}\cdot\exp\{-nh(\hat{\underset{\sim}{\theta}})\}$$

Por lo tanto

$$\tilde{E}[q(\underset{\sim}{\theta})|\underset{\sim}{x}] = q(\hat{\underset{\sim}{\theta}}^*)\cdot\frac{|\Sigma^*(\hat{\underset{\sim}{\theta}}^*)|^{1/2}\cdot\exp\{-nh(\hat{\underset{\sim}{\theta}}^*)\}}{|\Sigma(\hat{\underset{\sim}{\theta}})|^{1/2}\cdot\exp\{-nh(\hat{\underset{\sim}{\theta}})\}}$$

$$= q(\hat{\underset{\sim}{\theta}}^*)\cdot\frac{p_x(\hat{\underset{\sim}{\theta}}^*)\ |\Sigma^*(\hat{\underset{\sim}{\theta}}^*)|^{1/2}}{p_x(\hat{\underset{\sim}{\theta}})\ |\Sigma(\hat{\underset{\sim}{\theta}})|^{1/2}}$$

De hecho, puede demostrarse que

$$\tilde{E}[q(\underset{\sim}{\theta})|\underset{\sim}{x}] = E[q(\underset{\sim}{\theta})|\underset{\sim}{x}]\{1+O(n^{-2})\}$$

(ver Tierney & Kadane 1986; JASA 81, 82-86).

Si $q(\underset{\sim}{\theta})$ no es positiva, es posible utilizar la aproximación de Laplace sobre

$$E[q(\underset{\sim}{\theta})+a\,|\,\underset{\sim}{x}]$$

donde a es una constante suficientemente grande. Entonces

$$\tilde{E}[q(\underset{\sim}{\theta})|\underset{\sim}{x}] = \tilde{E}[q(\underset{\sim}{\theta})+a\,|\,\underset{\sim}{x}] - a.$$

Comentario: En términos generales, tanto la aproximación de Laplace como los métodos discutidos más adelante en este capítulo serán más precisos y/o eficientes a medida que la aproximación normal a la distribución final sea más adecuada.

En la mayoría de los casos resulta conveniente trabajar en términos de una parametrización $\underset{\sim}{\varphi} = \varphi(\underset{\sim}{\theta})$ tal que los componentes φ_i de $\underset{\sim}{\varphi}$ tomen valores en $\mathbb{R}$ y la distribución final de $\underset{\sim}{\varphi}$ sea aproximadamente normal. También es importante que la correlación entre los componentes de $\underset{\sim}{\varphi}$ bajo la distribución final no sea muy alta.

* Aproximación de Laplace para densidades marginales.

• Sea $\underset{\sim}{\theta} = (\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$, con $\underset{\sim}{\theta}_1 \in \mathbb{R}^{d_1}$ y $\underset{\sim}{\theta}_2 \in \mathbb{R}^{d-d_1}$. Supongamos que la distribución final de $\underset{\sim}{\theta}$ se escribe como

$$p(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) \propto q(\theta_1, \theta_2) \cdot \exp\{-n\, h(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)\},$$

donde $h: \mathbb{R}^d \to \mathbb{R}$ y $q: \mathbb{R}^d \to \mathbb{R}$ son funciones suaves.

Interesa calcular la densidad marginal de $\underset{\sim}{\theta}_1$, dada por

$$p(\underset{\sim}{\theta}_1) \propto \int q(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) \cdot \exp\{-n\, h(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)\}\, d\underset{\sim}{\theta}_2 .$$

Para cada valor de $\underset{\sim}{\theta}_1$, sean

$$q_{\theta_1}(\underset{\sim}{\theta}_2) = q(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$$

y

$$h_{\theta_1}(\underset{\sim}{\theta}_2) = h(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$$

(es decir, $q(\cdot)$ y $h(\cdot)$ vistas como función de $\underset{\sim}{\theta}_2$ solamente), y supongamos que $h_{\theta_1}(\underset{\sim}{\theta}_2)$ tiene un mínimo en $\underset{\sim}{\hat{\theta}}_2 = \underset{\sim}{\hat{\theta}}_2(\underset{\sim}{\theta}_1)$.

- La forma estándar de la aproximación de Laplace produce entonces la siguiente aproximación para $p(\underset{\sim}{\theta}_1)$:

$$\hat{p}(\underset{\sim}{\theta}_1) \propto q_{\underset{\sim}{\theta}_1}(\hat{\underset{\sim}{\theta}}_2(\underset{\sim}{\theta}_1)) \cdot |\Sigma(\underset{\sim}{\theta}_1)|^{1/2} \cdot \exp\{-n\, h_{\theta_1}(\hat{\underset{\sim}{\theta}}_2(\underset{\sim}{\theta}_1))\}$$

$$\propto |\Sigma(\underset{\sim}{\theta}_1)|^{1/2}\, p(\underset{\sim}{\theta}_1, \hat{\underset{\sim}{\theta}}_2(\underset{\sim}{\theta}_1)),$$

donde

$$\Sigma(\underset{\sim}{\theta}_1) = \Sigma_{\theta_1}(\hat{\underset{\sim}{\theta}}_2(\underset{\sim}{\theta}_1)),$$

con

$$\Sigma_{\theta_1}(\theta_2) = \left\{ \frac{\partial^2 h_{\theta_1}(\underset{\sim}{\theta}_2)}{\partial \underset{\sim}{\theta}_2' \partial \underset{\sim}{\theta}_2} \right\}^{-1}.$$

- Notemos que $p(\underset{\sim}{\theta}_1)$ también puede escribirse como

$$p(\underset{\sim}{\theta}_1) \propto \int q^*(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) \cdot \exp\{-n\, h^*(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)\}\, d\underset{\sim}{\theta}_2$$

con $q^*(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) \equiv 1$ y

$$h^*(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) = -\frac{1}{n} \log q(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) + h(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2).$$

Para cada valor de θ_1, sea $h^*_{\theta_1}(\underset{\sim}{\theta}_2)$ (i.e., $h^*(\cdot)$ vista como función de $\underset{\sim}{\theta}_2$ solamente) y supongamos que $h^*_{\theta_1}(\underset{\sim}{\theta}_2)$ tiene un mínimo en $\hat{\underset{\sim}{\theta}}_2^* = \hat{\underset{\sim}{\theta}}_2^*(\underset{\sim}{\theta}_1)$.

- La forma exponencial de la aproximación de Laplace produce ahora la siguiente aproximación para $p(\underset{\sim}{\theta}_1)$:

$$\tilde{p}(\underset{\sim}{\theta}_1) \propto |\Sigma^*(\underset{\sim}{\theta}_1)|^{1/2} \cdot \exp\{-n\, h^*_{\theta_1}(\hat{\underset{\sim}{\theta}}_2^*(\underset{\sim}{\theta}_1))\}$$

$$\propto |\Sigma^*(\underset{\sim}{\theta}_1)|^{1/2} \cdot p(\underset{\sim}{\theta}_1, \hat{\underset{\sim}{\theta}}_2^*(\underset{\sim}{\theta}_1)),$$

donde
$$\Sigma^*(\underset{\sim}{\theta}_1) = \Sigma^*_{\theta_1}\left(\hat{\underset{\sim}{\theta}}{}^*_2(\underset{\sim}{\theta}_1)\right),$$

con
$$\Sigma^*_{\theta_1}(\theta_2) = \left\{ \frac{\partial^2 h^*_{\theta_1}(\underset{\sim}{\theta}_2)}{\partial \underset{\sim}{\theta}_2' \, \partial \underset{\sim}{\theta}_2} \right\}^{-1}.$$

- Debe observarse que estas expresiones proporcionan aproximaciones para el kernel de $p(\underset{\sim}{\theta}_1)$ solamente. Las correspondientes constantes de normalización generalmente deben calcularse utilizando algún otro método. Excepto por esto, en general el método de Laplace parece ser más eficaz para aproximar densidades marginales que para aproximar valores esperados. En muchos casos la aproximación de Laplace "captura" la forma de la verdadera densidad marginal, especialmente si tanto la parametrización como la función $h(\cdot)$ se eligen con cuidado.

Ejemplo: (Distribución Logística). Sea
$$p(\theta_1, \theta_2) = \frac{2e^{-\theta_1} e^{-\theta_2}}{(1+e^{-\theta_1}+e^{-\theta_2})^3} \qquad (\theta_1, \theta_2) \in \mathbb{R}^2$$

La densidad marginal de θ_1 está dada por

$$p(\theta_1) = 2e^{-\theta_1} \int_{-\infty}^{\infty} e^{-\theta_2} / (1+e^{-\theta_1}+e^{-\theta_2})^3 \, d\theta_2$$

$$= e^{-\theta_1} / (1+e^{-\theta_1}+e^{-\theta_2})^2 \Big]_{-\infty}^{\infty}$$

$$= e^{-\theta_1} / (1+e^{-\theta_1})^2 .$$

Calculemos ahora la aproximación de Laplace en su forma exponencial. Sea

$$\begin{aligned} h^*_{\theta_1}(\theta_2) &= -\log p(\theta_1, \theta_2) \\ &= -\log 2 + \theta_1 + \theta_2 + 3\log(1 + e^{-\theta_1} + e^{-\theta_2}) \end{aligned}$$

$$\frac{\partial h^*_{\theta_1}(\theta_2)}{\partial \theta_2} = 1 - \frac{3e^{-\theta_2}}{1 + e^{-\theta_1} + e^{-\theta_2}}$$

$$\frac{\partial h^*_{\theta_1}(\hat{\theta}^*_2)}{\partial \theta_2} = 0 \iff e^{-\hat{\theta}^*_2} = \frac{1}{2}(1 + e^{-\theta_1})$$

$$\iff \hat{\theta}^*_2(\theta_1) = -\log\left\{\frac{1}{2}(1 + e^{-\theta_1})\right\}$$

$$\frac{\partial^2 h^*_{\theta_1}(\theta_2)}{\partial \theta_2^2} = \frac{3(1 + e^{-\theta_1})e^{-\theta_2}}{(1 + e^{-\theta_1} + e^{-\theta_2})^2}$$

$$\Sigma^*(\theta_1) = 3/2$$

$$\therefore \quad \tilde{p}(\theta_1) \propto p(\theta_1, \hat{\theta}^*_2(\theta_1))$$

$$\propto \frac{e^{-\theta_1}(1 + e^{-\theta_1})}{\left\{\frac{3}{2}(1 + e^{-\theta_1})\right\}^3}$$

$$\propto e^{-\theta_1} / (1 + e^{-\theta_1})^2 \qquad \theta_1 \in \mathbb{R}.$$

Ejemplo: (Distribución Normal/Gamma). Sea

$$p(\mu,\tau) = p(\mu|\tau)\,p(\tau)$$
$$= N(\mu|\,\mu_0,\,v_0/\tau)\cdot Ga(\tau|\,\alpha/2,\,\beta/2)$$

i.e.

$$p(\mu,\tau) \propto \tau^{1/2}\exp\left\{-\frac{\tau}{2v_0}(\mu-\mu_0)^2\right\}\cdot\tau^{\alpha/2-1}\cdot\exp\left\{-\beta\tau/2\right\}$$
$$\propto \tau^{(\alpha+1)/2-1}\cdot\exp\left\{-\tfrac{1}{2}\left[\beta+\tfrac{1}{v_0}(\mu-\mu_0)^2\right]\tau\right\}.$$

La densidad marginal de μ está dada por

$$p(\mu) \propto 2\,\Gamma((\alpha+1)/2)\,/\,\left[\beta+\tfrac{1}{v_0}(\mu-\mu_0)^2\right]^{(\alpha+1)/2}$$
$$\propto 1\,/\,\left[1+\frac{1}{\beta v_0}(\mu-\mu_0)^2\right]^{(\alpha+1)/2}.$$

$$\therefore\quad p(\mu) = St(\mu|\,\mu_0,\,\beta v_0/\alpha,\,\alpha).$$

→ Laplace en τ (forma estándar).

Sean
$$g(\mu,\tau) = \tau^{1/2}\exp\left\{-\frac{\tau}{2v_0}(\mu-\mu_0)^2\right\}$$
y
$$h(\mu,\tau) = h(\tau) = \tfrac{1}{2}\left[\beta\tau-(\alpha-2)\log\tau\right]$$

$$\frac{\partial h(\tau)}{\partial\tau} = \tfrac{1}{2}\left[\beta-(\alpha-2)/\tau\right]$$

$$\frac{\partial h(\hat\tau)}{\partial\tau} = 0 \iff \hat\tau = (\alpha-2)/\beta$$

$$\frac{\partial^2 h(\tau)}{\partial\tau^2} = (\alpha-2)/2\tau^2$$

$\Sigma_\mu(\hat\tau)$ no depende de μ.

$$\therefore \quad \hat{p}(\mu) \propto \exp\left\{-\frac{\hat{\tau}}{2V_0}(\mu-\mu_0)^2\right\}$$

i.e.

$$\hat{p}(\mu) = N(\mu \mid \mu_0, \beta V_0/(\alpha-2))$$

→ Laplace en τ (forma exponencial).

Sean $g^*(\mu,\tau) \equiv 1$ y

$$h^*(\mu,\tau) = h^*_\mu(\tau) = \frac{1}{2}\left\{\left[\beta + \frac{1}{V_0}(\mu-\mu_0)^2\right]\tau - (\alpha-1)\log\tau\right\}$$

$$\frac{\partial h^*_\mu(\tau)}{\partial\tau} = \frac{1}{2}\left\{\left[\beta+\frac{1}{V_0}(\mu-\mu_0)^2\right] - (\alpha-1)/\tau\right\}$$

$$\frac{\partial h^*_\mu(\hat{\tau}^*)}{\partial\tau} = 0 \iff \hat{\tau}^*(\mu) = (\alpha-1)/\left[\beta+\frac{1}{V_0}(\mu-\mu_0)^2\right]$$

$$\frac{\partial^2 h^*_\mu(\tau)}{\partial\tau^2} = (\alpha-1)/2\tau^2$$

$$\Sigma^*(\mu) \propto \left[\beta+\frac{1}{V_0}(\mu-\mu_0)^2\right]^{-2}$$

$$\therefore \quad \tilde{p}_1(\mu) \propto \left[\beta+\frac{1}{V_0}(\mu-\mu_0)^2\right]^{-(\alpha-1)/2-1} \times \exp\{-(\alpha-1)/2\}$$

$$\propto 1\Big/\left[1+\frac{1}{\beta V_0}(\mu-\mu_0)^2\right]^{(\alpha+1)/2}$$

i.e.

$$\tilde{p}_1(\mu) = St(\mu \mid \mu_0, \beta V_0/\alpha, \alpha)$$

<u>Nota</u>: en este caso la aproximación de Laplace produce la distribución marginal verdadera.

→ Laplace en $\lambda = \log \tau$ (forma exponencial)

$$p(\mu) \propto \int_{-\infty}^{\infty} \exp\left\{ \lambda(\alpha+1)/2 - \tfrac{1}{2}\left[\beta + \tfrac{1}{V_0}(\mu-\mu_0)^2\right] e^{\lambda} \right\} d\lambda$$

Sean $g^*(\mu,\lambda) \equiv 1$ y

$$h^*(\mu,\lambda) = h^*_\mu(\lambda) = \tfrac{1}{2}\left\{\left[\beta + \tfrac{1}{V_0}(\mu-\mu_0)^2\right] e^{\lambda} - (\alpha+1)\lambda\right\}$$

$$\frac{\partial h^*_\mu(\lambda)}{\partial\lambda} = \frac{1}{2}\left\{\left[\beta + \frac{1}{V_0}(\mu-\mu_0)^2\right] e^{\lambda} - (\alpha+1)\right\}$$

$$\frac{\partial h^*_\mu(\hat{\lambda}^*)}{\partial\lambda} = 0 \iff e^{\hat{\lambda}^*} = (\alpha+1)\Big/\left[\beta + \tfrac{1}{V_0}(\mu-\mu_0)^2\right]$$

$$\frac{\partial^2 h^*_\mu(\lambda)}{\partial\lambda^2} = \frac{1}{2}\left[\beta + \frac{1}{V_0}(\mu-\mu_0)^2\right] e^{\lambda}$$

$$\Sigma^*(\mu) = 2/(\alpha+1) \qquad \text{(no depende de } \mu)$$

$$\therefore \quad \tilde{p}_2(\mu) \propto \left[\beta + \tfrac{1}{V_0}(\mu-\mu_0)^2\right]^{-(\alpha+1)/2} \cdot \exp\left\{-(\alpha+1)/2\right\}$$

$$\propto 1\Big/\left[1 + \tfrac{1}{\beta V_0}(\mu-\mu_0)^2\right]^{(\alpha+1)/2}$$

i.e.

$$\tilde{p}_2(\mu) = St(\mu \mid \mu_0, \beta V_0/\alpha, \alpha)$$

<u>Nota:</u> en este caso la aproximación de Laplace también produce la distribución marginal verdadera.

6.2. Integración Numérica.

Los métodos de integración numérica (también conocida como cuadratura) permiten calcular eficientemente algunas características de la distribución final de $\underset{\sim}{\theta}$ cuando la dimensión de éste es pequeña.

* Caso univariado.

Sea $f: \mathbb{R} \longrightarrow \mathbb{R}$ una función suave.
Consideremos la integral

$$I = \int_a^b f(\underset{\sim}{\theta})\, d\underset{\sim}{\theta}.$$

Una regla de integración está definida por un conjunto de nodos $\{\eta_i\}_{i=1}^N$ y un conjunto asociado de pesos o ponderaciones $\{u_i\}_{i=1}^N$ tales que

$$I \approx \sum_{i=1}^N u_i f(\eta_i).$$

A continuación describiremos algunas reglas simples utilizadas comúnmente.

- Sean $a = \theta_0 < \theta_1 < \cdots < \theta_N = b$ $(N+1)$ puntos distribuidos sobre el intervalo $[a,b]$. En particular, si los puntos son equidistantes entonces

$$\theta_i = \theta_0 + i\,h \qquad (i=1,\ldots,N)$$

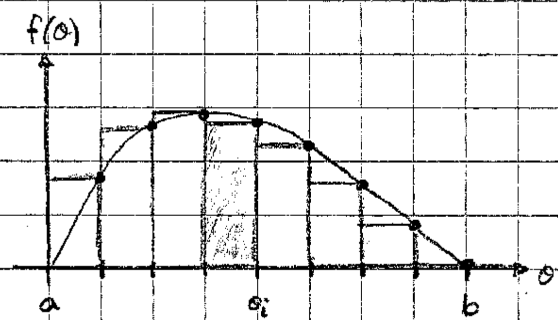
con $h = (b-a)/N$. En otras palabras

$$\theta_i = a + i(b-a)/N \qquad (i = 0, 1, \ldots, N).$$

- Regla del límite superior.

Está dada por

$$\hat{I}_{LS} = \sum_{i=1}^{N} f(\theta_i)\,(\theta_i - \theta_{i-1}). \quad ⊕$$



Si los puntos son equidistantes entonces

$$\hat{I}_{LS} = h \sum_{i=1}^{N} f(\theta_i)$$

$$= \frac{(b-a)}{N} \cdot \sum_{i=1}^{N} f\left(a + i(b-a)/N\right).$$

⊕ // En este caso

$$\eta_i = \theta_i \quad y \quad u_i = (\theta_i - \theta_{i-1})$$

$(i = 1, \ldots, N)$. //

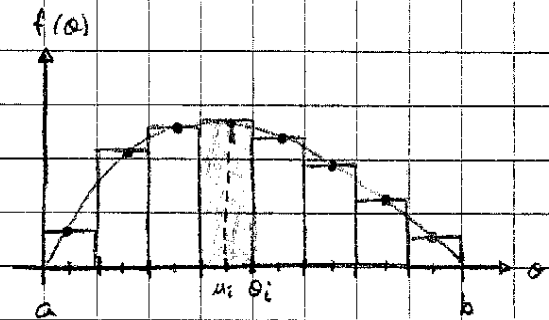
Nota: La regla del límite inferior es análoga.

- Regla del punto medio.

Está dada por

$$\hat{I}_{pM} = \sum_{i=1}^{N} f(\mu_i)(\theta_i - \theta_{i-1}),$$

con $\mu_i = (\theta_i + \theta_{i-1})/2 \qquad (i=1,\dots,N).$ ⊕



En particular, si los puntos son equidistantes entonces

$$\mu_i = (2\theta_0 + ih + (i-1)h)/2$$

$$= \theta_0 + (2i-1)h/2$$

$$= a + (2i-1)(b-a)/2N \qquad (i=1,\dots,N).$$

Por lo tanto

$$\hat{I}_{pM} = \frac{(b-a)}{N} \cdot \sum_{i=1}^{N} f\left(a + (2i-1)(b-a)/2N\right).$$

⊕ // Aquí

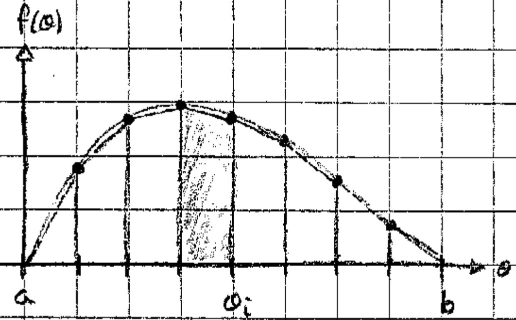
$$\eta_i = \mu_i \qquad y \qquad u_i = (\theta_i - \theta_{i-1})$$

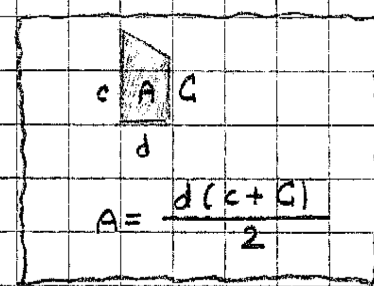
$(i=1,\dots,N).$ //

- Regla trapezoidal.

Esta regla está definida por

$$\hat{I}_T = \frac{1}{2} \sum_{i=1}^{N} \{ f(\theta_{i-1}) + f(\theta_i) \} (\theta_i - \theta_{i-1}). \quad (+)$$





Si los puntos son equidistantes entonces

$$\hat{I}_T = \frac{h}{2} \sum_{i=1}^{N} \{ f(\theta_{i-1}) + f(\theta_i) \}$$

$$= \frac{h}{2} \left\{ f(\theta_0) + 2 \sum_{i=1}^{N-1} f(\theta_i) + f(\theta_N) \right\}$$

$$= h \left\{ [f(\theta_0) + f(\theta_N)]/2 + \sum_{i=1}^{N-1} f(\theta_i) \right\}$$

$$= \frac{(b-a)}{N} \left\{ [f(a)+f(b)]/2 + \sum_{i=1}^{N-1} f(a + i(b-a)/N) \right\}.$$

(+) // En este caso

$$\eta_i = \theta_i \qquad (i = 0, 1, \ldots, N)$$

y

$$u_i = \begin{cases} (\theta_1 - \theta_0)/2 & i = 0 \\ (\theta_{i+1} - \theta_{i-1})/2 & i = 1, \ldots, N-1 \\ (\theta_N - \theta_{N-1})/2 & i = N \end{cases}. \quad //$$

- **Nota:** En la mayoría de las aplicaciones se requiere integrar sobre regiones no acotadas, tales como $(-\infty, \infty)$ o $(0, \infty)$, en cuyo caso las reglas descritas anteriormente no pueden aplicarse directamente. Una alternativa consiste en elegir límites a y b $(a<b)$ de manera que el valor de $f(\theta)$ sea despreciable fuera del intervalo $[a, b]$. Como otra posibilidad, notemos que a través de un cambio de variable apropiado la integral requerida puede reescribirse como

$$\int_a^b f^*(\phi)\, d\phi$$

para alguna variable ϕ y algunos límites $a<b$.

La siguiente regla de integración permite calcular de manera exacta la integral de ciertas funciones definidas sobre $\mathbb{R}$, y produce buenas aproximaciones para una clase muy amplia de funciones.

- Regla de Gauss-Hermite

Supongamos que $f(\theta)$ es de la forma

$$f(\theta) = q(\theta) \exp\{-\theta^2\}$$

donde $q(\theta)$ es una función suave de θ.

La regla de Gauss-Hermite está definida por

$$\hat{I}_{GH} = \sum_{i=1}^{N} w_i \, q(r_i),$$

donde

$$r_i = \text{i-ésima raíz de } H_N(t),$$

$$w_i = \frac{2^{N-1} \, N! \, \sqrt{\pi}}{N^2 \cdot H_{N-1}(r_i)^2}$$

y $H_N(t)$ denota al polinomio de Hermite de orden N. Ver Abramowitz & Stegun (1965), donde también pueden encontrarse tablas de $\{r_i\}$ y $\{w_i\}$ para distintos valores de N.

Notemos que

$$\sum_{i=1}^{N} w_i \, q(r_i) = \sum_{i=1}^{N} u_i \, f(\eta_i),$$

con $\eta_i = r_i$ y $u_i = w_i \cdot \exp\{r_i^2\}$.

Esta regla elige los nodos y los pesos de manera que la aproximación sea exacta si $q(\theta)$ es un polinomio de orden a lo más $(2N-1)$ en θ. En general, la aproximación es buena aun si $q(\theta)$ es sólo aproximadamente un polinomio (lo cual es común en algunas aplicaciones estadísticas).

Más generalmente, si $f(\theta)$ es de la forma

$$f(\theta) = q(\theta) \cdot N(\theta \mid \mu, \sigma^2)$$

entonces

$$\hat{I}_{GH} = \pi^{-1/2} \sum_{i=1}^{N} w_i \cdot q(\mu + \sqrt{2\sigma^2}\, r_i)$$

o, equivalentemente,

$$\hat{I}_{GH} = \sum_{i=1}^{N} \tilde{u}_i \, f(\tilde{\eta}_i),$$

con $\tilde{u}_i = \sqrt{2\sigma^2} \exp\{r_i^2\} \cdot w_i$ y $\tilde{\eta}_i = \mu + \sqrt{2\sigma^2}\, r_i$.

* Caso multivariado.

Supongamos por el momento que $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$, y notemos que

$$\int_a^b \int_{c(\theta_1)}^{d(\theta_1)} f(\theta_1, \theta_2)\, d\theta_2\, d\theta_1 = \int_a^b \left\{ \int_{c(\theta_1)}^{d(\theta_1)} f(\theta_1, \theta_2)\, d\theta_2 \right\} d\theta_1$$

$$= \int_a^b f_1(\theta_1)\, d\theta_1$$

donde

$$f_1(\theta_1) = \int_{c(\theta_1)}^{d(\theta_1)} f(\theta_1, \theta_2)\, d\theta_2 .$$

En principio es posible utilizar alguna regla de integración univariada

$$\int_a^b f_1(\theta_1)\, d\theta_1 \approx \sum_{i=1}^{N_1} w_{1i}\, f_1(\eta_{1i}) \qquad (13)$$

para integrar $f_1(\theta_1)$, pero para ello es necesario que $f_1(\theta_1)$ sea una función conocida. Dado que $f_1(\theta_1)$ está definida por una integral, es posible aproximar $f_1(\eta_{1i})$ para cada η_{1i} usando una segunda regla de integración univariada

$$f_1(\eta_{1i}) = \int_{c(\eta_{1i})}^{d(\eta_{1i})} f(\eta_{1i}, \theta_2)\, d\theta_2$$

$$\approx \sum_{j=1}^{N_2} w_{2j} \cdot f(\eta_{1i}, \eta_{2j}). \qquad (14)$$

La regla resultante en dos dimensiones está dada por

$$\hat{I} = \sum_{i=1}^{N_1} \cdot \sum_{j=1}^{N_2} w_{1i} \cdot w_{2j}\; f(\eta_{1i}, \eta_{2j}).$$

Podemos pensar ahora en esta regla como una sola regla bidimensional con $N_1 N_2$ nodos $\{\underset{\sim}{\eta}\} = \{(\eta_{1i}, \eta_{2j})\}$ y pesos $\{\underset{\sim}{w}\} = \{(w_{1i}, w_{2j})\}$. Esta regla se conoce como el producto cartesiano de las dos reglas unidimensionales (13) y (14).

- La extensión al caso de d dimensiones es directa. Cabe mencionar que, en general, estas reglas son ineficientes si d es grande.