Técnicas de Reducción de Varianza: Muestreo por importancia y muestreo condicional

Temas Selectos I

Dr. Rubén Blancas Rivera

Universidad de las Américas Puebla

Otoño 2025

Content

Introducción

- En la simulación de Monte Carlo, una de las principales limitaciones es la **varianza de los estimadores**.
- El teorema del límite central asegura:

$$\text{Error típico} \sim \frac{1}{\sqrt{n}}.$$

- En la práctica: se requieren demasiadas simulaciones para lograr la precisión deseada.
- $\Rightarrow$ Necesidad de métodos que produzcan estimadores más eficientes.

Introducción

- Estrategias para disminuir la **dispersión de los resultados** de Monte Carlo.
- Mantienen la validez estadística de los estimadores.
- Idea central:
  - Aprovechar información adicional.
  - Usar correlaciones o estructuras del problema.
  - Concentrar la variabilidad y acelerar la convergencia.

Muestreo por Importancia

Supongamos nuevamente que el problema es calcular una integral de la forma

$$\theta = \int_a^b g(x)\,dx.$$

Usando el método de la media muestral, si X es una variable aleatoria con función de densidad $f(x)$, cuyo soporte contiene al intervalo (a,b), entonces la integral θ se puede expresar como la siguiente esperanza:

$$\theta = \int_a^b \frac{g(x)}{f(x)} f(x)\,dx = \mathbb{E}\left(\frac{g(X)}{f(X)}\right) \tag{1}$$

Muestreo por Importancia

El método Monte Carlo de la media muestral establece que si $X_1, \dots, X_n$ es una muestra aleatoria de la función de densidad $f(x)$, entonces se propone como estimador para θ a la variable aleatoria

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \frac{g(X_i)}{f(X_i)} \tag{2}$$

Muestreo por Importancia

- Usando (1) es fácil verificar que $\hat{\theta}$ es un estimador insesgado para el valor desconocido θ.
- Esta afirmación es válida para cualquier función de densidad $f(x)$ que se utilice.

Muestreo por Importancia

- El objetivo del **muestreo por importancia** es utilizar la función de densidad $f(x)$ tal que los puntos que se generen de esta función tengan mayor o menor frecuencia según lo indique la función a integrar $g(x)$.
- Esta función de densidad existe y es aquella que hace que la varianza del estimador $\hat{\theta}$ sea mínima.

Muestreo por Importancia

Teorema.

La varianza del estimador $\hat{\theta}$ definido mediante el método de la media muestral (2) satisface

$$\mathrm{Var}(\hat{\theta}) \;\geq\; \frac{1}{n}\left[\left(\int_a^b |g(x)|\,dx\right)^2 - \theta^2\right].$$

Además, el valor mínimo se alcanza cuando se toma

$$f(x) = \frac{|g(x)|}{\int_a^b |g(x)|\,dx}, \qquad a < x < b. \tag{3}$$

Demostración.
Recordemos que la desigualdad de Cauchy–Schwarz para integrales establece que

$$\left(\int_a^b f(x)\, g(x)\, dx\right)^2 \leq \left(\int_a^b f^2(x)\, dx\right) \cdot \left(\int_a^b g^2(x)\, dx\right),$$

con igualdad $\Leftrightarrow$ $f(x)/g(x)$ es constante.
Por lo tanto, usando (2),

$$\mathrm{Var}(\hat{\theta}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i)}{f(X_i)}\right) = \frac{1}{n}\,\mathrm{Var}\left(\frac{g(X)}{f(X)}\right).$$

Muestreo por Importancia

La varianza del cociente g/f puede expresarse como

$$\mathrm{Var}\left(\frac{g(X)}{f(X)}\right) = \mathbb{E}\left(\frac{g^2(X)}{f^2(X)}\right) - \mathbb{E}^2\left(\frac{g(X)}{f(X)}\right)$$

$$= \int_a^b \frac{g^2(x)}{f^2(x)}\, f(x)\, dx \; - \; \theta^2$$

$$= \left(\int_a^b \left(\frac{|g(x)|}{f^{1/2}(x)}\right)^2 dx\right) \cdot \left(\int_a^b \left(f^{1/2}(x)\right)^2 dx\right) - \theta^2$$

$$= \left(\int_a^b \left(\frac{|g(x)|}{f^{1/2}(x)}\right)^2 dx\right) \cdot 1 - \theta^2 \;\; \geq \;\; \left(\int_a^b |g(x)|\, dx\right)^2 - \theta^2. \quad (4)$$

Muestreo por Importancia

Cuando se toma la función de densidad

$$f(x) = \frac{|g(x)|}{\int_a^b |g(x)|\, dx}$$

en la expresión (4), la varianza toma el valor

$$\mathrm{Var}(\hat{\theta}) = \frac{1}{n} \int_a^b \frac{g^2(x)}{|g(x)| / \int_a^b |g(x)|\, dx}\, dx - \frac{\theta^2}{n}$$

$$= \frac{1}{n} \left(\int_a^b |g(x)|\, dx \right)^2 - \frac{\theta^2}{n}.$$

Muestreo por Importancia

- De esta manera, en el sentido de varianza mínima, la mejor función de densidad $f(x)$ que uno puede tomar es la función definida como el valor absoluto de $g(x)$ dividido entre la integral del valor absoluto. Esto es una operación sobre $g(x)$ que la convierte en una función de densidad.

Muestreo por Importancia

- Para que esto sea posible se necesita la hipótesis adicional de que $|g(x)|$ sea integrable en el intervalo (a, b). Por otro lado, el procedimiento de aproximación requiere que se conozca una manera de generar valores de la distribución $f(x)$.

Muestreo por Importancia

Debe observarse que encontrar la función de densidad óptima $f(x)$ para aproximar la integral θ requiere conocer el valor de la integral

$$\theta = \int_a^b |g(x)|\, dx,$$

lo cual es similar al problema original de encontrar θ.
De hecho, el problema es exactamente el mismo en el caso cuando $g(x) \geq 0$.

Muestreo por Importancia

Sin embargo, el teorema anterior sugiere tomar como $f(x)$ a aquella función de densidad que tenga un comportamiento parecido a $|g(x)|$ y de la cual conozcamos una manera de generar valores.

Aproximación discreta

Como una aproximación a la función de densidad óptima $f(x)$ dada en (3), se puede dividir el intervalo (a, b) en m partes:

$$a < u_1 < u_2 < \cdots < u_{m-1} < b,$$

y definir la función constante por pedazos

$$\bar{g}(x) = \begin{cases} g_1, & a < x < u_1, \\ g_2, & u_1 < x < u_2, \\ \vdots & \\ g_m, & u_{m-1} < x < b. \end{cases}$$

Aproximación Discreta

- en donde $g_1, g_2, \ldots, g_m$ son ciertos valores de la función $g(x)$ en cada uno de los m subintervalos, respectivamente.
- Cada valor constante g_i puede ser cualquiera de los valores que toma la función $g(x)$ en el intervalo (u_{i-1}, u_i), con $u_0 = a$ y $u_m = b$.

Aproximación Discreta

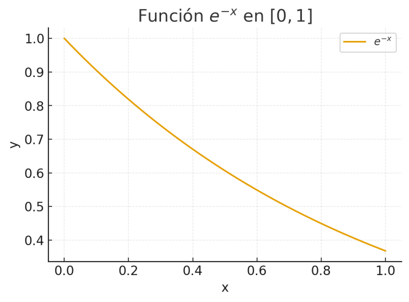
- Tomando valor absoluto de $\bar{g}(x)$ y dividiendo entre la suma de estos valores absolutos, la función $\bar{g}(x)$ se puede transformar en la función de probabilidad $f_d(x)$ de una variable aleatoria discreta y de la cual sabemos una manera de obtener sus valores.
- Si $x_1, \ldots, x_n$ son valores de esta variable aleatoria discreta, entonces la aproximación es

$$\int_a^b g(x)\,dx \approx \frac{1}{n}\sum_{i=1}^{n}\frac{g(x_i)}{f_d(x_i)}.$$

Ejemplo

Queremos calcular

$$\theta = \int_0^1 g(x)\,dx, \qquad g(x) = e^{-x}.$$



Ejemplo

Dividimos $(0, 1)$ en $m = 2$ subintervalos con $0 < u_1 = 0.5 < 1$. Tomamos los puntos medios y definimos

$$g_1 = g(0.25) = e^{-0.25}, \qquad g_2 = g(0.75) = e^{-0.75}.$$

La función por pedazos queda

$$\bar{g}(x) = \begin{cases} g_1, & 0 < x < 0.5, \\ g_2, & 0.5 < x < 1. \end{cases}$$

Ejemplo

Normalizamos los valores para obtener una función de probabilidad (discreta) sobre $\{x_1, x_2\} = \{0.25, 0.75\}$:

$$f_d(x_1) = \frac{g_1}{g_1 + g_2}, \qquad f_d(x_2) = \frac{g_2}{g_1 + g_2}.$$

Ejemplo

Si la muestra es $\{0.25, 0.75, 0.25, 0.25, 0.75\}$, entonces

$$\hat{\theta} = \frac{1}{5}\left(\frac{g(0.25)}{f_d(0.25)} + \frac{g(0.75)}{f_d(0.75)} + \frac{g(0.25)}{f_d(0.25)} + \frac{g(0.25)}{f_d(0.25)} + \frac{g(0.75)}{f_d(0.75)}\right),$$

Ejemplo

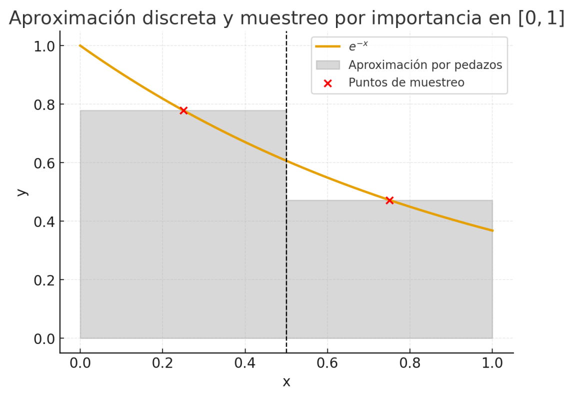
donde

$$g(0.25) = e^{-0.25}, \quad g(0.75) = e^{-0.75}, \quad f_d(0.25) = \frac{e^{-0.25}}{e^{-0.25} + e^{-0.75}}.$$

$$f_d(0.75) = \frac{e^{-0.75}}{e^{-0.25} + e^{-0.75}}.$$

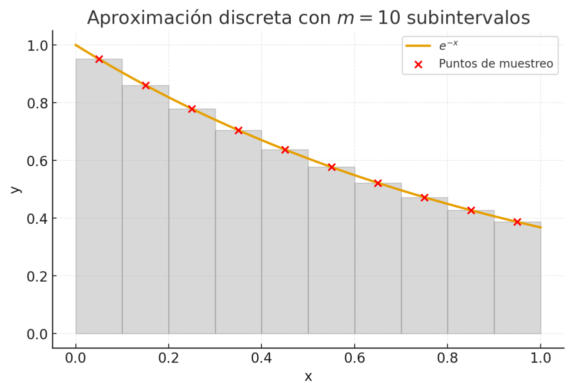
(Obsérvese que $\frac{g(0.25)}{f_d(0.25)} = e^{-0.25}\,\frac{e^{-0.25} + e^{-0.75}}{e^{-0.25}} = 1 + e^{-0.5}$ y análogamente $\frac{g(0.75)}{f_d(0.75)} = 1 + e^{0.5}$.)

Ejemplo



Aproximación discreta y muestreo por importancia en [0, 1]
e^{-x}
Aproximación por pedazos
Puntos de muestreo
1.0
0.8
0.6
0.4
0.2
0.0
y
0.0
0.2
0.4
0.6
0.8
1.0
x

Ejemplo


Aproximación discreta con $m = 10$ subintervalos
e^{-x}
Puntos de muestreo
1.0
0.8
0.6
0.4
0.2
0.0
y
0.0
0.2
0.4
0.6
0.8
1.0
x

Caso Multidimensional

El análisis llevado a cabo permanece válido cuando la función a integrar es función de varias variables. Por ejemplo, en el caso de 2 variables tenemos la integral

$$\theta = \int_{a_1}^{b_1} \int_{a_2}^{b_2} g(x, y)\, dy\, dx.$$

Caso Multidimensional

Como antes, cuando $f(x,y)$ es una función de densidad bivariada, el estimador por el método de la media muestral es

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i,Y_i)}{f(X_i,Y_i)}.$$

Caso Multidimensional

Por el teorema demostrado, la varianza de este estimador satisface

$$\mathrm{Var}(\hat{\theta}) \;\geq\; \frac{1}{n}\left[\left(\int_{a_1}^{b_1}\int_{a_2}^{b_2} |g(x,y)|\,dy\,dx\right)^2 - \theta^2\right],$$

en donde la cota inferior se alcanza cuando se toma

$$f(x,y) = \frac{|g(x,y)|}{\int_{a_1}^{b_1}\int_{a_2}^{b_2} |g(x,y)|\,dy\,dx}, \qquad a_1 < x < b_1, \;\; a_2 < y < b_2.$$

Esperanza Condicional

- La esperanza condicional de una variable aleatoria X dado otra variable aleatoria Y se denota por el símbolo $E(X \mid Y)$.
- A pesar de su nombre y escritura, no se trata de un número sino de una variable aleatoria.
- Uno de sus posibles valores es el número $E(X \mid Y = y)$, en donde y es cualquier valor de Y.
- A esta última cantidad también se le llama esperanza condicional y se estudia en los cursos básicos de probabilidad.

Varianza Condicional

También se puede definir la varianza condicional de X respecto de Y de la siguiente manera.

Varianza Condicional

Sea X una variable aleatoria con segundo momento finito y sea Y otra variable aleatoria. La varianza condicional de X dado Y es la variable aleatoria

$$\mathrm{Var}(X \mid Y) = \mathbb{E}\big((X - \mathbb{E}(X \mid Y))^2 \,\big|\, Y\big).$$

Esperanza Condicional

Proposición

La varianza condicional satisface las siguientes propiedades:

1. $\mathrm{Var}(X \mid Y) \geq 0$.
2. $\mathrm{Var}(X \mid Y) = \mathbb{E}(X^2 \mid Y) - \mathbb{E}^2(X \mid Y)$.
3. $\mathrm{Var}(X) = \mathbb{E}\big(\mathrm{Var}(X \mid Y)\big) + \mathrm{Var}\big(\mathbb{E}(X \mid Y)\big)$.

Esperanza Condicional

Proposición

Para cualquier variable aleatoria integrable X,

$$\operatorname{Var}\big(\mathbb{E}(X \mid Y)\big) \;\leq\; \operatorname{Var}(X).$$

Muestreo Condicional

- La esperanza condicional $E(X \mid Y)$ se puede considerar como un estimador insesgado para la cantidad $E(X)$ pues $E(E(X \mid Y)) = E(X)$, y
- tiene varianza posiblemente más pequeña que la varianza de X debido a la Proposición anterior.
- A la aplicación de este resultado al problema de estimar una integral usando el método de la media muestral de Monte Carlo se le llama *muestreo condicional*.

Muestreo Condicional

Sea $g(x)$ una función integrable sobre el intervalo (a, b). Consideremos nuevamente el problema de estimar la siguiente integral

$$\theta = \int_a^b g(x)\,dx = \int_a^b \frac{g(x)}{f(x)} f(x)\,dx = \mathbb{E}\left(\frac{g(X)}{f(X)}\right),$$

donde X es una variable aleatoria con función de densidad $f(x)$.

Muestreo Condicional

Si $X_1, \ldots, X_n$ es una muestra aleatoria de esta distribución, entonces el estimador de Monte Carlo de la media muestral para θ es

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} \frac{g(X_i)}{f(X_i)}. \tag{6}$$

Este estimador es insesgado para θ, es decir, $\mathbb{E}(\hat{\theta}_1) = \theta$, y su varianza es

$$\operatorname{Var}(\hat{\theta}_1) = \operatorname{Var}\left(\frac{1}{n} \sum_{i=1}^{n} \frac{g(X_i)}{f(X_i)}\right) = \frac{1}{n^2} \sum_{i=1}^{n} \operatorname{Var}\left(\frac{g(X_i)}{f(X_i)}\right)$$

$$= \frac{1}{n} \operatorname{Var}\left(\frac{g(X)}{f(X)}\right).$$

Muestreo Condicional

Por otro lado, sabemos que algunas probabilidades y esperanzas son más fáciles de calcular cuando se condicionan sobre la ocurrencia de algún evento. En nuestra situación, supongamos que la siguiente esperanza condicional puede ser calculada con facilidad

$$\mathbb{E}\left(\frac{g(X)}{f(X)}\,\middle|\,Y=y\right), \tag{7}$$

donde Y es una variable aleatoria auxiliar que guarda alguna relación con X.

Muestreo Condicional

Si contamos con una expresión para esta esperanza condicional para cada posible valor y, entonces se puede encontrar una expresión para la variable aleatoria

$$\mathbb{E}\left(\frac{g(X)}{f(X)}\,\middle|\,Y\right), \tag{8}$$

Muestreo Condicional

Observemos que (7) es un número que depende del valor y, mientras que (8) es una variable aleatoria, a la que hemos llamado *esperanza condicional*. Esto sugiere que si $Y_1, \ldots, Y_n$ es una muestra aleatoria de la misma distribución que Y, entonces se puede definir el siguiente estimador por *muestreo condicional*

$$\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left(\frac{g(X)}{f(X)} \,\middle|\, Y_i \right). \tag{4.9}$$

Muestreo Condicional

- Observe que no es necesario tomar el mismo tamaño de muestra n para definir a $\hat{\theta}_1$ y a $\hat{\theta}_2$, pero lo hemos hecho así a fin de comparar sus varianzas.
- Observemos también que cuando X y Y son independientes, el muestreo sobre Y no es provechoso pues $\hat{\theta}_2$ se reduce a la cantidad desconocida

$$\theta = \mathbb{E}\left(\frac{g(X)}{f(X)}\right).$$

- Las propiedades del estimador $\hat{\theta}_2$ y su comparación con $\hat{\theta}_1$ se muestran a continuación.

Muestreo Condicional

Los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ definidos en (6) y (9) son ambos insesgados para θ, sin embargo,

$$\mathrm{Var}(\hat{\theta}_2) \;\leq\; \mathrm{Var}(\hat{\theta}_1).$$

Ejemplo

Queremos calcular

$$\theta = \int_0^1 \int_0^1 (x+y)\, dx\, dy.$$

La solución exacta:

$$\theta = \int_0^1 \int_0^1 (x+y)\, dx\, dy = \int_0^1 \left(\frac{1}{2} + y\right) dy = 1.$$

Ejemplo

1 Si generamos $(X_i, Y_i) \sim U([0,1]^2)$, el estimador es

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} (X_i + Y_i).$$

Ejemplo 1

Dado $Y = y$, la esperanza condicional es

$$\mathbb{E}[X + Y \mid Y = y] = \mathbb{E}[X] + y = \frac{1}{2} + y.$$

Entonces, generando $Y_i \sim U(0, 1)$,

$$\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} \left(\tfrac{1}{2} + Y_i\right).$$

Ambos estimadores son insesgados, pero $\hat{\theta}_2$ tiene menor varianza porque aprovecha la información condicional.

Ejemplo 2

Sea (X, Y) un vector aleatorio con función de distribución $F(x, y)$. Supongamos que deseamos estimar la siguiente cantidad

$$\theta = P(X + Y \leq u) = E(\mathbf{1}_{(X+Y \leq u)}),$$

para algún número real u fijo.

Ejemplo 2

Si $(X_1, Y_1), \ldots, (X_n, Y_n)$ es una muestra aleatoria de la distribución $F(x, y)$, entonces el estimador para θ por el método directo de Monte Carlo es

$$\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(X_i+Y_i \leq u)}.$$

Ejemplo 2

Alternativamente, podemos usar muestreo condicional y escribir

$$\begin{aligned}\theta &= E\big(\mathbf{1}_{(X+Y\le u)}\big)\\ &= E\big(E(\mathbf{1}_{(X+Y\le u)} \mid Y)\big)\\ &= E\big(E(\mathbf{1}_{(X\le u-Y)} \mid Y)\big)\\ &= E\big(F_X(u-Y)\big),\end{aligned}$$

en donde $F_X(x)$ es la función de distribución marginal de X.

Ejemplo 2

Por lo tanto, si $Y_1, \ldots, Y_n$ es una muestra aleatoria de la distribución $F_Y(y)$ (la función de distribución marginal de Y), entonces el estimador por muestreo condicional para θ es

$$\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} F_X(u - Y_i),$$

en donde sabemos que la varianza de este estimador es menor o igual a la varianza de $\hat{\theta}_1$.

Ejemplo 2

- Es interesante observar que los elementos de la suma que aparece en $\hat{\theta}_1$ son 0 ó 1,
- mientras que los elementos de la suma en $\hat{\theta}_2$ son valores en el intervalo $(0, 1)$.
- Intuitivamente, esto explica el hecho de que los valores de $\hat{\theta}_2$ presentan menor dispersión.

Ejemplo 3

Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias independientes con idéntica distribución $F(x)$. Consideremos una variable aleatoria N con valores $0, 1, \ldots$ e independiente de $X_1, X_2, \ldots$. Defina la suma aleatoria

$$S_N = \sum_{i=1}^{N} X_i.$$

Ejemplo 3

- Observe que el número de sumandos es aleatorio.
- A las distribuciones de este tipo de sumas se les conoce como *distribuciones compuestas*.
- Por ejemplo, se conoce como *modelo Poisson compuesto* a aquella suma en donde N tiene distribución Poisson (λ) y los sumandos X_i tienen cualquier distribución general.

Ejemplo 3

- En general, el poder encontrar la distribución de S_N es un problema complicado.
- Consideremos entonces el problema de estimar, usando el método de Monte Carlo, la siguiente probabilidad, en donde x es cualquier número real fijo,

$$\theta = P(S_N \leq x) = E\big(\mathbf{1}_{(S_N \leq x)}\big).$$

Ejemplo 3

Método Monte Carlo directo. Si $(N_1, X_1^{(1)}, \dots, X_{N_1}^{(1)}), \dots, (N_n, X_1^{(n)}, \dots, X_{N_n}^{(n)})$ es una muestra aleatoria de tamaño n del vector $(N, X_1, \dots, X_N)$, entonces podemos estimar θ mediante

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(S_{N_i}^{(i)} \leq x)}.$$

en donde $S_N^{(i)} = X_1^{(i)} + \cdots + X_{N_i}^{(i)}$ para $i = 1, \dots, n$.

Ejemplo 3

Sabemos que este es un estimador insesgado para $\hat{\theta}$ y su varianza es

$$\operatorname{Var}(\hat{\theta}_1) = \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}\left(\mathbf{1}_{(S_N^{(i)}\leq x)}\right) = \frac{1}{n}\theta(1-\theta).$$

Ejemplo 3

Método Monte Carlo condicional. Para $n \geq 0$ entero fijo, la función de distribución de S_N con $N = n$ es la n-ésima convolución de $F(x)$. Esta es una nueva función de distribución denotada por $F^{*n}(x)$. Representa la función de distribución de la suma de n variables aleatorias independientes $X_1, \ldots, X_n$ con idéntica distribución $F(x)$, es decir,

$$F^{*n}(x) = P(X_1 + \cdots + X_n \leq x), \qquad -\infty < x < \infty.$$

Ejemplo 3

Cuando $n \geq 2$ podemos escribir esta probabilidad de la siguiente forma:

$$
\begin{aligned}
F^{*n}(x) &= P\left(\sum_{i=1}^{n} X_i \leq x\right) \\
&= P\left(X_1 \leq x - \sum_{i=2}^{n} X_i\right) \\
&= F\left(x - \sum_{i=2}^{n} X_i\right) \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F\left(x - \sum_{i=2}^{n} X_i \,\middle|\, X_2 = x_2, \ldots, X_n = x_n\right) \\
&\quad \times f(x_2, \ldots, x_n)\, dx_2 \cdots dx_n
\end{aligned}
$$

Ejemplo 3

$$
\begin{aligned}
F^{*n}(x) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F\left(x - \sum_{i=2}^{n} x_i\right) f(x_2, \ldots, x_n)\, dx_2 \cdots dx_n \\
&= E\left[F\left(x - \sum_{i=2}^{n} X_i\right)\right].
\end{aligned}
$$

Ejemplo 3

Como esta igualdad se cumple para cualquier entero $n \geq 1$ (cuando $n = 0$ y $n = 1$ la suma es vacía y se define como cero, de modo que la esperanza es $F(x)$), concluimos la validez de la expresión general,

$$F^{*N}(x) = E\left(F\left(x - \sum_{i=2}^{N} X_i\right)\right).$$

Ejemplo 3

De esta manera, la esperanza de la variable $\mathbf{1}_{(S_N \leq x)}$ se ha simplificado un poco al condicionar sobre $Y = (X_2, \ldots, X_N)$. Por lo tanto, si

$$(N_1, X_1^{(1)}, \ldots, X_{N_1}^{(1)}), \ldots, (N_n, X_1^{(n)}, \ldots, X_{N_n}^{(n)})$$

es una muestra aleatoria de tamaño n del vector $(N, X_1, \ldots, X_N)$.

Ejemplo 3

Se propone el estimador insesgado

$$\hat{\theta}_2 = \frac{1}{n}\sum_{k=1}^{n} F\left(x - \sum_{i=2}^{N_k} X_i^{(k)}\right),$$

el cual tiene varianza menor que el estimador que se obtiene del método de Monte Carlo directo. En este caso no se usan las primeras variables $X_1^{(k)}$ y no es necesario generarlas.