Método de la Medida Muestral

Temas Selectos I

Dr. Rubén Blancas Rivera

Universidad de las Américas Puebla

Otoño 2025

Content

Método de la Media Muestral

Como antes, si $g(x)$ es una función integrable, se busca encontrar el valor de

$$\theta = \int_a^b g(x)\, dx, \tag{1}$$

en donde ahora consideraremos que la función $g(x)$ no necesariamente es positiva ni acotada superiormente, como se supuso en el caso del método clásico de integración Monte Carlo. Más aún, el intervalo de integración (a, b) no necesariamente será acotado, puede ser el intervalo completo $(-\infty, \infty)$, por ejemplo.

Método de la Media Muestral

Sea X cualquier variable aleatoria con una función de densidad $f(x)$, cuyo soporte es el intervalo (a, b). Entonces la integral θ se puede escribir como el siguiente valor esperado

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

Método de la Media Muestral

Observe que se puede considerar que el cociente $g(X)/f(X)$ es una variable aleatoria y que un estimador para su media es la media muestral, esto es, si $X_1, \ldots, X_n$ es una muestra aleatoria de la función de densidad $f(x)$, entonces se puede definir el siguiente estimador.

Definition

Sea $g(x)$ una función integrable sobre el intervalo (a, b). La función y el intervalo no son necesariamente acotados. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una densidad $f(x)$ con soporte (a, b). Se define el estimador Monte Carlo de la media muestral para la integral θ como

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

Algoritmo

1. Sea $g(x)$ una función a integrar en el intervalo (a, b).
2. Sea $f(x)$ cualquier función de densidad con soporte (a, b).
3. Sean $x_1, \ldots, x_n$ valores de la distribución $f(x)$.

Entonces

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

Dado que la media muestral es un estimador insesgado para la media de una distribución, tenemos que el estimador $\hat{\theta}$ es insesgado. En efecto,

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{g(X_i)}{f(X_i)}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(\frac{g(X_i)}{f(X_i)}\right) = \frac{1}{n}\sum_{i=1}^{n}\theta = \theta.$$

Ejemplo 1

Supongamos que se desea integrar una función $g(x)$ en el intervalo $(0, \infty)$. Puede usarse la función de densidad exponencial $f(x) = \lambda e^{-\lambda x}$, cuyo soporte es justamente ese intervalo. Por simplicidad tomaremos $\lambda = 1$.

Ejemplo 1

Entonces el método de la media muestral establece que

$$\int_0^\infty g(x)\,dx = \int_0^\infty \frac{g(x)}{f(x)} f(x)\,dx = \mathbb{E}\left(\frac{g(X)}{f(X)}\right) = \mathbb{E}(e^X g(X)),$$

en donde el término $e^X g(X)$ es una nueva variable aleatoria. Si $x_1, \ldots, x_n$ son valores de la distribución exponencial indicada, entonces

$$\int_0^\infty g(x)\,dx \;\approx\; \frac{1}{n}\sum_{i=1}^n e^{x_i} g(x_i).$$

Ejemplo 2

Consideremos ahora que tenemos una función $g(x)$ definida en el intervalo $(-\infty, \infty)$, que es integrable y cuya integral se desea aproximar por el método de la media muestral usando la distribución $\mathcal{N}(0, 1)$.

Ejemplo 2

Si $\phi(x)$ denota la función de densidad $\mathcal{N}(0,1)$, entonces

$$\int_{-\infty}^{\infty} g(x)\,dx = \int_{-\infty}^{\infty} \frac{g(x)}{\phi(x)}\,\phi(x)\,dx = \mathbb{E}\left(\frac{g(X)}{\phi(X)}\right) = \sqrt{2\pi}\,\mathbb{E}\left(e^{X^2/2}g(X)\right),$$

en donde $X \sim \mathcal{N}(0,1)$. Si $x_1,\ldots,x_n$ son valores de la distribución normal estándar, entonces

$$\int_{-\infty}^{\infty} g(x)\,dx \;\approx\; \frac{\sqrt{2\pi}}{n}\sum_{i=1}^{n} e^{x_i^2/2}\,g(x_i).$$

Caso Uniforme

Un caso sencillo y conveniente se obtiene cuando se toma a $f(x)$ como la función de densidad unif(a, b), suponiendo que este intervalo es acotado. En esta situación, la esperanza a calcular es

$$\theta = \mathbb{E}\left(\frac{g(X)}{f(X)}\right) = (b-a)\,\mathbb{E}(g(X)). \qquad (2)$$

Así, tomando una muestra aleatoria $X_1, \ldots, X_n$ de la distribución unif(a, b), tenemos el estimador

$$\hat{\theta}_u := (b-a)\,\frac{1}{n}\sum_{i=1}^{n} g(X_i).$$

Caso Uniforme

$$
\begin{aligned}
\mathrm{Var}(\hat{\theta}_u) &= \mathrm{Var}\left((b-a)\,\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right) \\
&= (b-a)^2\,\frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(g(X_i)) \\
&= (b-a)^2\,\frac{1}{n}\left[\mathbb{E}(g^2(X)) - \mathbb{E}^2(g(X))\right] \\
&= (b-a)^2\,\frac{1}{n}\left[\frac{1}{b-a}\int_a^b g^2(x)\,dx - \frac{\theta^2}{(b-a)^2}\right] \\
&= \frac{1}{n}\left[(b-a)\int_a^b g^2(x)\,dx - \theta^2\right].
\end{aligned}
$$

Caso uniforme: comparación de estimadores

Aplicar tanto el método clásico como el método de la media muestral en su versión uniforme para aproximar la integral de $g(x)$. Los correspondientes estimadores son:

$$\hat{\theta}_1 = c(b-a)\,\frac{N}{n} \qquad \text{Método clásico}$$

$$\hat{\theta}_2 = (b-a)\,\frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

Donde $\hat{\theta}_2$ es el estimador usando el método de la media muestral con $f(x)$ la función de densidad unif(a,b).

Caso uniforme: comparación de estimadores

Proposición

Para los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$, se cumple

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

Caso uniforme: comparación de estimadores

Proof.

Por las hipótesis y los cálculos anteriores,

$$\mathrm{Var}(\hat{\theta}_2) = \frac{1}{n}\left[(b-a)\int_a^b g(x)\cdot g(x)\,dx - \theta^2\right]$$

$$\leq \frac{1}{n}\left[(b-a)\,c\int_a^b g(x)\,dx - \theta^2\right] = \frac{1}{n}\left[c(b-a)\theta - \theta^2\right] = \mathrm{Var}(\hat{\theta}_1).$$

□

Así, en el sentido de varianza menor, el estimador $\hat{\theta}_2$ es mejor que el estimador $\hat{\theta}_1$. Recordemos que para el estimador $\hat{\theta}_2$ se tomó la distribución uniforme. Surge entonces la pregunta:

- ¿existe una función de densidad $f(x)$ que produzca un estimador para la integral θ que tenga varianza mínima?

Caso Multidimensional

El método de la media muestral en una dimensión puede extenderse fácilmente al caso de dimensiones mayores. En el siguiente recuadro se explica el procedimiento en el caso bidimensional.

Integración Monte Carlo: método de la media muestral (dimensión 2)

1. Sea $g(x, y)$ la función a integrar en el rectángulo $(a_1, b_1) \times (a_2, b_2)$.
2. Sea $f(x, y)$ una función de densidad con soporte $(a_1, b_1) \times (a_2, b_2)$.
3. Sean $(x_1, y_1), \ldots, (x_n, y_n)$ puntos al azar de acuerdo a la función de densidad $f(x, y)$.

Caso Multidimensional

Entonces

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} g(x,y)\, dy\, dx \;\approx\; \frac{1}{n} \sum_{i=1}^{n} \frac{g(x_i, y_i)}{f(x_i, y_i)}.$$

Caso Multidimensional

En particular, cuando el rectángulo $(a_1, b_1) \times (a_2, b_2)$ es acotado y se toma a $f(x, y)$ como la función de densidad uniforme, se tiene la aproximación

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} g(x, y)\, dy\, dx \;\approx\; (b_1 - a_1)(b_2 - a_2)\, \frac{1}{n} \sum_{i=1}^{n} g(x_i, y_i).$$