

# Intervalos de confianza: MonteCarlo vs. Bootstrap

Felipe Gerard

Octubre de 2015

## Antecedentes

Cuando se estima una cantidad de interés, es importante conocer la precisión del estimador que se está utilizando. Cuando la distribución es conocida (o parecida a una conocida), se puede hacer supuestos y calcular explícitamente la distribución del estimador, de donde se puede obtener intervalos de confianza. Sin embargo, cuando es desconocida, usualmente se recurre a propiedades asintóticas de los estimadores. En particular, se supone que el estimador “ya convergió” a una normal (por el TCL) y se utilizan los intervalos de confianza asociados a una normal. La pregunta es entonces si esto es una buena idea o si hay algún método para obtener intervalos de mejor calidad. En este trabajo compararemos los intervalos asintóticos contra los obtenidos con remuestreo bootstrap.

## Planteamiento

Supongamos que tenemos una variable aleatoria  $X \sim f$ , de la cual queremos conocer  $\theta = E[\varphi(X)]$ . Supongamos que tenemos una muestra aleatoria  $\mathcal{L} = \{X_1, \dots, X_N\}$  de  $X$  y que vamos a estimar  $\theta$  con  $\hat{\theta} = \hat{\theta}(\mathcal{L})$ . Aludiendo a Montecarlo, tendríamos

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^N \varphi(X_i) \rightarrow N(\theta, \sigma_{\varphi(X)}^2/N)$$

De este modo, si estimamos  $\theta$  con  $\hat{\theta}$  y  $\sigma_{\varphi(X)}^2$  con la varianza muestral  $S_{\varphi(X)}^2$ , entonces el intervalo de confianza asintótico aproximado para  $\theta$  es

$$\hat{\theta} \pm S_{\varphi(X)}/\sqrt{N}$$

Ahora bien, otra forma de obtener intervalos es utilizando remuestreo bootstrap. Para ello, en lugar de utilizar una muestra real de  $X$  pero con una distribución aproximadamente normal para  $\hat{\theta}$ , el enfoque es obtener una muestra aproximada de  $\hat{\theta}$  y utilizar la distribución real (de la muestra aproximada) para obtener los intervalos. Una remuestra bootstrap de  $\mathcal{L}$  es una muestra con reemplazo de tamaño  $N$ ,  $\mathcal{L}_b = \{X_1^b, \dots, X_N^b\}$ , de la cual podemos obtener

$$\hat{\theta}_b = \frac{1}{N} \sum_{i=1}^N \varphi(X_i^b),$$

que se distribuye aproximadamente como  $\hat{\theta}$ . Si obtenemos  $B$  remuestras bootstrap, tendremos una muestra aleatoria  $\hat{\mathcal{L}}_\theta = \{\hat{\theta}_1, \dots, \hat{\theta}_B\}$  que aproxima la verdadera distribución de  $\hat{\theta}$ . Podemos obtener por ejemplo los intervalos de confianza empíricos simples con significancia  $1 - \alpha$ :

$$[q_{\alpha/2}^B, q_{1-\alpha/2}^B]$$

## Comparación

La comparación teórica de ambos tipos de intervalos es compleja y requiere muchos antecedentes teóricos. En esta sección daremos una breve intuición de lo que sucede. Uno de los problemas que puede tener usar los intervalos normales es que, dependiendo de la distribución de  $\varphi(X)$ ,  $\hat{\theta}$  podría tardar más o menos en converger a una normal y en general es difícil saber cuándo. El problema es aún más grave cuando  $N$  no es muy grande (por ejemplo, en el caso de un fenómeno donde es difícil muestrear o donde es muy caro simular de  $f$ ). En contraste con el método asintótico, lo único que requiere bootstrap para funcionar bien es que  $\mathcal{L}$  sea razonablemente representativa, es decir, que  $\hat{F} \approx F$ , donde  $F$  es la distribución asociada a  $f$  y  $\hat{F}$  es la distribución empírica obtenida de  $\mathcal{L}$ .

El factor peligroso de los intervalos normales asintóticos es que, cuando la distribución del estadístico converge por un lado (p. ej. cuando  $\hat{\theta}$  se distribuye en realidad como una  $\chi^2$ ), aunque converja en el límite, los intervalos estarán consistentemente sesgados y afectarán los resultados a menos que la muestra sea realmente grande. Este hecho puede hacer que, en la práctica, los intervalos no tengan la cobertura que deberían, hecho particularmente riesgoso en fenómenos con mucho sesgo. Una ventaja de bootstrap en este caso es que los intervalos que propone *no están sesgados*, sino que se aproximan más y más a los reales cuando  $\hat{F}_\theta \rightarrow F_\theta$ . Además de ello, se puede

construir los intervalos bootstrap de varias maneras, por ejemplo si se quiere garantizar que sean simétricos.

Otra ventaja es que bootstrap utiliza una sola estimación y no dos para obtener los intervalos. Esto se debe a que para usar el teorema asintótico, como no se conoce  $\sigma_{\varphi(X)}^2$ , se utiliza  $S_{\varphi(X)}^2$ . Es decir, se estima la varianza y luego se obtienen los intervalos (cuantiles) de una aproximación asintótica a la distribución. En contraste, en bootstrap únicamente se estima  $\hat{F}_{\hat{\theta}}$  a partir de  $\mathcal{L}$  y se utilizan sus cuantiles para estimar los reales.

Una manera de ver que los intervalos asintóticos no siempre son adecuados es en experimentos artificiales en los que se conoce la verdadera distribución de  $\hat{\theta}$ . Otra alternativa que permite experimentar con funciones más complejas es aumentar progresivamente el tamaño de muestra y ver si hay un sesgo sistemático de los intervalos asintóticos en comparación con la distribución normal asintótica.

## Bibliografía

DiCiccio, Thomas J.; Efron, Bradley. Bootstrap confidence intervals. *Statist. Sci.* 11 (1996), no. 3, 189–228. URL: [http://projecteuclid.org/download/pdf\\_1/euclid.ss/1032280214](http://projecteuclid.org/download/pdf_1/euclid.ss/1032280214)

Hall, Peter. Bootstrap Lectures. University of California, Davis, CA, USA. URL: <http://anson.ucdavis.edu/~peterh/sta251/bootstrap-lectures-to-may-16.pdf>