Properties of Estimators

Proof of the Cramér-Rao bound

Bruce M. Boghosian

Department of Mathematics

Tufts University

Outline

The Cauchy-Schwarz inequality

- Note that

$$\begin{aligned} 0 \leq & \|\vec{X}z + \vec{Y}\|^2 \\ = & \left(\vec{X}z + \vec{Y}\right) \cdot \left(\vec{X}z + \vec{Y}\right) \\ = & \vec{X} \cdot \vec{X}z^2 + 2\vec{X} \cdot \vec{Y}z + \vec{Y} \cdot \vec{Y} \\ = & \|X\|^2 z^2 + 2\vec{X} \cdot \vec{Y}z + \|\vec{Y}\|^2. \end{aligned}$$

- Because the above quadratic in z has at most one real root, its discriminant must be less than or equal to zero, so

$$4\left(\vec{X} \cdot \vec{Y}\right)^2 - 4\|X\|^2\|\vec{Y}\|^2 \leq 0.$$

- From this follows the *Cauchy-Schwarz inequality*,

$$\boxed{\left|\vec{X} \cdot \vec{Y}\right| \leq \|X\|\|Y\|}$$

Standard deviation as a vector norm

- Given random data $\{X_j\}_{j=1}^n$ and $\{Y_j\}_{j=1}^n$ with means

$$\mu_X = \frac{1}{n}\sum_{j=1}^n X_j \qquad \mu_Y = \frac{1}{n}\sum_{j=1}^n Y_j$$

- Define the *deviations from the means*

$$\vec{X} := \{X_j - \mu_X\}_{j=1}^n \qquad \vec{Y} := \{Y_j - \mu_Y\}_{j=1}^n$$

- The standard deviations are then

$$\sigma_X = \sqrt{\frac{1}{n}\sum_{j=1}^n (X_j - \mu_X)^2} = \sqrt{\frac{1}{n}\vec{X}\cdot\vec{X}} = \frac{1}{\sqrt{n}}\|\vec{X}\|$$

$$\sigma_Y = \sqrt{\frac{1}{n}\sum_{j=1}^n (Y_j - \mu_Y)^2} = \sqrt{\frac{1}{n}\vec{Y}\cdot\vec{Y}} = \frac{1}{\sqrt{n}}\|\vec{Y}\|$$

Covariance and correlation as inner products

- The *covariance* between X and Y is then

$$\mathrm{Cov}(X,Y) = \frac{1}{n}\sum_{j=1}^{n}(X_j-\mu_X)(Y_j-\mu_Y) = \frac{1}{n}\vec{X}\cdot\vec{Y}$$

- The *Pearson correlation coefficient* between X and Y is then

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{\frac{1}{n}\vec{X}\cdot\vec{Y}}{\frac{1}{\sqrt{n}}\|\vec{X}\|\frac{1}{\sqrt{n}}\|\vec{Y}\|} = \frac{\vec{X}\cdot\vec{Y}}{\|\vec{X}\|\,\|\vec{Y}\|}$$

- Note that, by the Cauchy-Schwarz inequality, we have

$$|\rho_{X,Y}| = \frac{\left|\vec{X}\cdot\vec{Y}\right|}{\|\vec{X}\|\,\|\vec{Y}\|} \le 1,$$

so $\rho(X,Y)\in[-1,+1]$, or $|\mathrm{Cov}(X,Y)|^2 \le \mathrm{Var}(X)\,\mathrm{Var}(Y)$

Proof of the Cramér-Rao bound I

Bruce M. Boghosian

- Suppose we have random variable X with a one-parameter PDF $f(x;\theta)$.
- We have an estimator $\hat{t}(X)$ whose expectation is $\psi(\theta)$
- Estimators are random variables, so give this one a name

$$T = \hat{t}(X)$$

- Expectation value of T is a function of the parameter θ,

$$E(T) = \int dx\ f(x;\theta)\hat{t}(x) = \psi(\theta).$$

- We want to show that there is a lower bound on

$$\mathrm{Var}(T) = E(T^2) - (E(T))^2 = E(T^2) - [\psi(\theta)]^2$$

Proof of the Cramér-Rao bound II

Bruce M. Boghosian

- Define another random variable

$$V = \frac{\partial}{\partial\theta} \ln f(X;\theta) = \frac{1}{f(X;\theta)} \frac{\partial}{\partial\theta} f(X;\theta)$$

- Note that this has zero mean

$$\begin{aligned}
E(V) &= \int dx\ f(x;\theta) V \\
&= \int dx\ f(x;\theta) \frac{1}{f(x;\theta)} \frac{\partial}{\partial\theta} f(x;\theta) \\
&= \int dx\ \frac{\partial}{\partial\theta} f(x;\theta) \\
&= \frac{\partial}{\partial\theta} \int dx\ f(x;\theta) \\
&= 0.
\end{aligned}$$

Proof of the Cramér-Rao bound III

- Now consider the covariance of V and T,

$$\begin{aligned}
\text{Cov}(V, T) =& E\left[(T - \psi(\theta))\left(\frac{1}{f(X;\theta)}\frac{\partial}{\partial\theta}f(X;\theta)\right)\right] \\
=& E\left[T\left(\frac{1}{f(X;\theta)}\frac{\partial}{\partial\theta}f(X;\theta)\right)\right] \\
=& \int dx\ f(X;\theta)\hat{t}(x)\left(\frac{1}{f(X;\theta)}\frac{\partial}{\partial\theta}f(X;\theta)\right) \\
=& \frac{\partial}{\partial\theta}\int dx\ f(X;\theta)\hat{t}(x) \\
=& \psi'(\theta)
\end{aligned}$$

Proof of the Cramér-Rao bound IV

Bruce M. Boghosian

- To recap, starting with random variable X,
  - We constructed estimator $T = \hat{t}(X)$ with expectation $E(T) = \psi(\theta)$
  - We defined $V = \frac{\partial}{\partial\theta} \ln f(X;\theta)$ with expectation $E(V) = 0$
  - We found $\mathrm{Cov}(V,T) = \psi'(\theta)$
- Now, by the Cauchy-Schwarz inequality, we have

$$\mathrm{Var}(T)\,\mathrm{Var}(V) \geq |\mathrm{Cov}(V,T)|^2 = |\psi'(\theta)|^2,$$

- and from this it follows that

$$\mathrm{Var}(T) \geq \frac{|\psi'(\theta)|^2}{\mathrm{Var}(V)} = \frac{|\psi'(\theta)|^2}{E\left[n\left(\frac{\partial}{\partial\theta}\ln f(X_j;\theta)\right)^2\right]}$$

Proof of the Cramér-Rao bound V

- We have found that

$$\mathrm{Var}(T) \geq \frac{|\psi'(\theta)|^2}{\mathrm{Var}(V)} = \frac{|\psi'(\theta)|^2}{E\left[n\left(\frac{\partial}{\partial\theta}\ln f(X_i;\theta)\right)^2\right]}$$

- In the event that the estimator $\hat{t}$ is for θ itself, and is unbiased so that $E(T) = \psi(\theta) = \theta$, the above result becomes

$$\mathrm{Var}(T) \geq \frac{1}{E\left[n\left(\frac{\partial}{\partial\theta}\ln f(X_i;\theta)\right)^2\right]}$$

- This gives us the first-derivative form of the theorem.

Proof of the Cramér-Rao bound VI

- Finally, note that

$$
\begin{aligned}
E\left[\left(\frac{\partial}{\partial\theta}\ln f(X_i;\theta)\right)^2\right] &= \int dx\ f(x;\theta)\left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)\left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)\\
&= \int dx\ f(x;\theta)\frac{1}{f(x;\theta)}\left(\frac{\partial f(x;\theta)}{\partial\theta}\right)\left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)\\
&= \int dx\ \frac{\partial f(x;\theta)}{\partial\theta}\left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)\\
&= -\int dx\ f(x;\theta)\left(\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta)\right)\\
&= E\left[-\frac{\partial^2}{\partial\theta^2}\ln f(X_i;\theta)\right].
\end{aligned}
$$

- This gives us the second-derivative form of the theorem.

Summary

- We learned about and proved the Cauchy-Schwarz inequality.
- We used the Cauchy-Schwarz inequality to prove both forms of the CR bound.