Large-sample approximate sampling distribution of the MLE: Pareto Distribution

Stat 343 Lab: Sampling Distribution of the MLE

The Pareto distribution is often used to model the distribution of variables with long right tail; for example, one recent article used the Pareto distribution (and several other candidate models) to model the distribution of carbon dioxide emissions by countries around the world (Akhundjanov, S. B., Devadoss, S., & Luckstead, J. (2017). Size distribution of national CO2 emissions. Energy Economics, 66182-193).

The Pareto distribution has two parameters, $x_0 > 0$ and $\alpha > 0$. If $X \sim \text{Pareto}(x_0, \alpha)$, then X is a continuous random variable. Its probability density function (pdf) is:

$$f(x|x_0, \alpha) = \frac{\alpha x_0^\alpha}{x^{\alpha+1}} \mathbb{I}_{[x_0, \infty)}(x) = \begin{cases} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & \text{if } x \geq x_0 \\ 0 & \text{otherwise} \end{cases}$$

The mean of X is $E(X) = \frac{\alpha x_0}{\alpha - 1}$ if $\alpha > 1$, and infinity otherwise.

Let's fix $x_0 = 1$ and think about how we might estimate the parameter α.

Strategy 1: Method of Moments

One approach to estimation that we have skipped over so far is the method of moments. The method of moments works by setting the first few moments of a distribution equal to the corresponding moments of a data set, and solving for any unknown parameters. In this case, that means setting the sample mean equal to the mean of the Pareto distribution and solving for α. Go ahead and do that to find an estimator of α, $\hat{\alpha}^{MOM}$. It will be a function of the sample mean $\bar{X}$.

Strategy 2: Maximum Likelihood

I'll let you skip a few steps (but on the other hand, here is a chance to get some practice if the mechanics of these steps are not yet comfortable).

The log-likelihood for a Pareto distribution is given by

$\ell(\alpha|x_1, \ldots, x_n, x_0) = n\log(\alpha) + n\alpha\log(x_0) - (\alpha+1)\sum_{i=1}^{n}\log(x_i).$

The first derivative of the log-likelihood is

$\frac{d}{d\alpha}\ell(\alpha|x_1, \ldots, x_n, x_0) = \frac{n}{\alpha} + n\log(x_0) - \sum_{i=1}^{n}\log(x_i).$

The second derivative of the log-likelihood is

$\frac{d^2}{d\alpha^2}\ell(\alpha|x_1, \ldots, x_n, x_0) = \frac{-n}{\alpha^2}.$

Note that when $x_0 = 1$, $\log(x_0) = 0$.

Find the maximum likelihood estimator

Asymptotic Distribution of the MLE, assuming $\alpha^{true} = 1.1$

Find a normal approximation to the sampling distribution of $\hat{\alpha}^{MLE}$. This will depend on the sample size n. Is there any difference between the observed Fisher information and the Fisher information in this case?