Method of Moments

February 19, 2008

Let $M_!, M_2, \ldots$ be independent random variables having a common distribution possessing a mean μ. The strong law of large numbers states that the sample means converge to the population mean

$$\bar{M}_n = \frac{1}{n}\sum_{i=1}^{n} M_i \to \mu \quad \text{with probability 1 as } n \to \infty.$$

For a simple random sample $X_1, X_2, \ldots$ chosen according to $P_\theta, \theta \in \Theta$ and m a real valued function. If $g(\theta) = E_\theta m(X_1)$, then

$$\frac{1}{n}\sum_{i=1}^{n} m(X_i) \to g(\theta) \text{ with probability 1 as } n \to \infty.$$

The choices $m(x) = x^m$ is called the **method of moments**.

1 Example

Example 1. *A $\Gamma(\alpha, \beta)$ random variable has mean α/β and variance α/β^2.*

$$E_{(\alpha,\beta)}X_1 = \frac{\alpha}{\beta} \quad \textit{and} \quad E_{(\alpha,\beta)}X_1^2 = Var_{(\alpha,\beta)}(X_1) + E_{(\alpha,\beta)}[X_1]^2 = \frac{\alpha}{\beta^2} + \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha(1+\alpha)}{\beta^2}.$$

So set

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \textit{and} \quad \overline{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

to obtain estimates

$$\bar{X} = \frac{\hat{\alpha}}{\hat{\beta}} \quad \textit{and} \quad \overline{X^2} = \frac{\hat{\alpha}(1+\hat{\alpha})}{\hat{\beta}^2}.$$

Then,

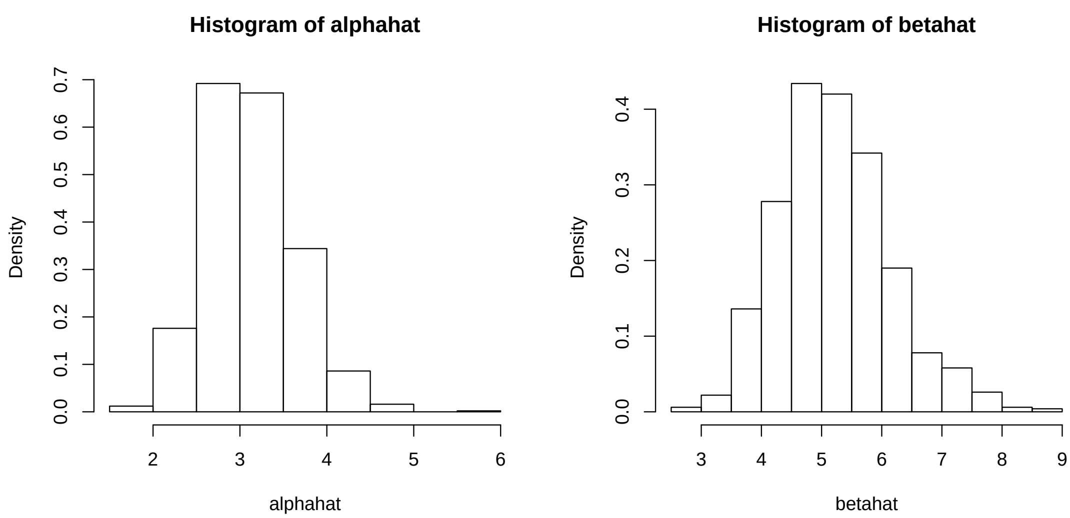
$$\overline{X^2} - (\bar{X})^2 = \frac{\hat{\alpha}}{\hat{\beta}^2} \quad \frac{\bar{X}}{\overline{X^2} - (\bar{X})^2} = \hat{\beta}$$

and

$$\hat{\alpha} = \hat{\beta}\bar{X} = \frac{(\bar{X})^2}{\overline{X^2} - (\bar{X})^2}$$

To investigate the method of moments on simulated data using R*, we consider 1000 repetitions of 100 independent observations of a $\Gamma(3, 5)$ random variables.*

```
> xbar <- rep(0,1000)
> x2bar <- rep(0,1000)
> for (i in 1:1000){x<-rgamma(100,3,5);xbar[i]=mean(x);x2bar[i]=mean(x^2)}
> betahat <- xbar/(x2bar-(xbar)^2)
> alphahat <- betahat*xbar
> mean(alphahat)
[1] 3.121951
> var(alphahat)
[1] 0.2601792
> mean(betahat)
[1] 5.217926
> var(betahat)
[1] 0.8466257
> hist(alphahat,probability=TRUE)
> hist(betahat,probability=TRUE)
```



2 Comparison to Maximum Likelihood Estimation

To obtain the maximum likelihood estimate, write

$$\mathbf{L}(\alpha,\beta|\mathbf{x}) = \left(\frac{\beta^\alpha}{\Gamma(\alpha)} x_1^{\alpha-1} e^{-\beta x_1}\right) \cdots \left(\frac{\beta^\alpha}{\Gamma(\alpha)} x_n^{\alpha-1} e^{-\beta x_n}\right).$$

$$\ln \mathbf{L}(\alpha,\beta|\mathbf{x}) = n(\alpha \ln\beta - \ln\Gamma(\alpha)) + (\alpha-1)\sum_{i=1}^n \log x_i - \beta \sum_{i=1}^n x_i.$$

To determine the parameters that maximize the likelihood, solve the equations

$$\frac{\partial}{\partial \alpha} \ln \mathbf{L}(\alpha, \beta|\mathbf{x}) = n(\ln \beta - \frac{d}{d\alpha} \ln \Gamma(\alpha)) + \sum_{i=1}^{n} \ln x_i = 0, \quad \overline{\ln x} = \frac{d}{d\alpha} \ln \Gamma(\alpha) - \ln \beta$$

and

$$\frac{\partial}{\partial \beta} \ln \mathbf{L}(\alpha, \beta|\mathbf{x}) = n\frac{\alpha}{\beta} - \sum_{i=1}^{n} x_i = 0, \quad \bar{x} = \frac{\alpha}{\beta}.$$

To compute the Fisher information note that

$$I_\alpha(\alpha, \beta) = -\frac{\partial^2}{\partial \alpha^2} \ln \mathbf{L}(\alpha, \beta|\mathbf{x}) = n\frac{d^2}{d\alpha^2} \ln \Gamma(\alpha) \quad \textit{and} \quad I_\beta(\alpha, \beta) = -\frac{\partial^2}{\partial \beta^2} \ln \mathbf{L}(\alpha, \beta|\mathbf{x}) = n\frac{\alpha}{\beta^2}.$$

For $\alpha = 3$ and $\beta = 5$ and $n = 100$,

$$I_\alpha(3, 5) = 100 \times (0.3943) = 39.43, \quad I_\beta(3, 5) = 100 \times \frac{3}{5^3} = 12$$

and

$$Var_{(3,5)}(\hat{\alpha}) \approx 0.043, \quad Var_{(3,5)}(\hat{\beta}) \approx 0.083.$$

Compare this to the empirical values of 0.261 and 0.847 for the method of moments