Lecture 06

Lecture on Estimating Distribution Parameters Using the Method of Moments

Parameter Estimation

in order to fit a probability law to data, one typically has to estimate parameters associated with the probability law from the data.

The Method of Moments

The kth moment of a probability law is defined as

$$\mu_k = E(X^k)$$

where X is a random variable following that probability law (of course, this is defined only if the expectation exists). If $X_1, X_2, \ldots, X_n$ are i.i.d. random variables from that distribution, the kth **sample moment** is defined as

$$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

We can view $\hat{\mu}_k$ as an estimate of μ_k. The method of moments estimates parameters by finding expressions for them in terms of the lowest possible order moments and then substituting sample moments into the expressions.

Suppose, for example, that we wish to estimate two parameters, θ_1 and θ_2. If θ_1 and θ_2 can be expressed in terms of the first two moments as

$$\theta_1 = f_1(\mu_1, \mu_2)$$

$$\theta_2 = f_2(\mu_1, \mu_2)$$

then the method of moments estimates are

$$\hat{\theta}_1 = f_1(\hat{\mu}_1, \hat{\mu}_2)$$

$$\hat{\theta}_2 = f_2(\hat{\mu}_1, \hat{\mu}_2)$$

Example: Normal Distribution
Suppose $\theta_1 := \mu$, and $\theta_2 := \sigma^2$.
The first and second moments for the normal distribution are

$$\mu_1 = E(X) = \theta_1 = \mu$$
$$\mu_2 = E(X^2) = \theta_1^2 + \theta_2 = \mu^2 + \sigma^2$$

Therefore,

$$\theta_1 = \mu_1$$
$$\theta_2 = \mu_2 - \mu_1^2$$

The Method of Moments

The construction of a method of moments estimate involves three basic steps:

1. Calculate low order moments, finding expressions for the moments in terms of the parameters. Typically, the number of low order moments needed will be the same as the number of parameters.
2. Invert the expressions found in the preceding step, finding new expressions for the parameters in terms of the moments.
3. Insert the sample moments into the expressions obtained in the second step, thus obtaining estimates of the parameters in terms of the sample moments.

Normal Distribution

The first and second moments for the normal distribution are

$$\mu_1 = E(X) = \mu$$

$$\mu_2 = E(X^2) = \mu^2 + \sigma^2$$

Therefore,

$$\mu = \mu_1$$

$$\sigma^2 = \mu_2 - \mu_1^2$$

The corresponding estimates of μ and σ^2 from the sample moments are

$$\hat{\mu} = \overline{X}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Gamma Distribution

The first two moments of the gamma distribution are

$$\mu_1 = \frac{\alpha}{\lambda}$$

$$\mu_2 = \frac{\alpha(\alpha + 1)}{\lambda^2}$$

(see Example B in Section 4.5). To apply the method of moments, we must express α and λ in terms of μ_1 and μ_2. From the second equation,

$$\mu_2 = \mu_1^2 + \frac{\mu_1}{\lambda}$$

or

$$\lambda = \frac{\mu_1}{\mu_2 - \mu_1^2}$$

Also, from the equation for the first moment given here,

$$\alpha = \lambda\mu_1 = \frac{\mu_1^2}{\mu_2 - \mu_1^2}$$

The method of moments estimates are, since $\hat{\sigma}^2 = \hat{\mu}_2 - \hat{\mu}_1^2$,

$$\hat{\lambda} = \frac{\overline{X}}{\hat{\sigma}^2}$$

and

$$\hat{\alpha} = \frac{\overline{X}^2}{\hat{\sigma}^2}$$

An Angular Distribution

The angle θ at which electrons are emitted in muon decay has a distribution with the density

$$f(x \mid \alpha) = \frac{1 + \alpha x}{2}, \qquad -1 \leq x \leq 1 \quad \text{and} \quad -1 \leq \alpha \leq 1$$

where $x = \cos\theta$. The parameter α is related to polarization. Physical considerations dictate that $|\alpha| \leq \frac{1}{3}$, but we note that $f(x \mid \alpha)$ is a probability density for $|\alpha| \leq 1$. The method of moments may be applied to estimate α from a sample of experimental measurements, $X_1, \ldots, X_n$. The mean of the density is

$$\mu_1 = \int_{-1}^{1} x \frac{1 + \alpha x}{2} dx = \frac{\alpha}{3}$$

Thus, the method of moments estimate of α is $\widehat{\alpha} = 3\widehat{\mu_1}$.

Example

Suppose that X is a discrete random variable with

$$P(X = 0) = \frac{2}{3}\theta$$

$$P(X = 1) = \frac{1}{3}\theta$$

$$P(X = 2) = \frac{2}{3}(1 - \theta)$$

$$P(X = 3) = \frac{1}{3}(1 - \theta)$$

where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observations were taken from such a distribution: (3, 0, 2, 1, 3, 2, 1, 0, 2, 1).

a. Find the method of moments estimate of θ.

b. Find the variance of your estimator.

For the method of moments, we first need to find the expected value of X, which is a discrete random variable with the probability distribution function as defined in the exercise. Using the definition of the expected value of a discrete random variable, we have

$$E(X) = \sum_{k=0}^{3} k \cdot P(X = k) = \frac{\theta}{3} + \frac{4}{3} \cdot (1 - \theta) + (1 - \theta) = \frac{7}{3} - 2\theta$$

From the above expression, we can express θ as

$$\theta = \frac{1}{2} \cdot \left(\frac{7}{3} - E(X)\right) = \frac{7}{6} - \frac{1}{2} \cdot E(X)$$

Now, the method of moments simply suggests writing the sample mean $\widehat{\mu_1}$ in place of $E(X)$, and that would be the method of moments estimate of θ.

So, the desired estimate is

$$\widehat{\theta} = \frac{7}{6} - \frac{1}{2} \cdot \widehat{\mu_1}.$$

The sample mean can be found as

$$\widehat{\mu_1} = \frac{3+0+2+1+3+2+1+0+2+1}{10} = \frac{3}{2},$$

which would yield a method of moment estimate of θ :

$$\widehat{\theta} = \frac{7}{6} - \frac{1}{2} \cdot \frac{3}{2} = \frac{5}{12} = 0.417$$

The variance of X_1 can be found as

$$\mathrm{Var}(X_1) = E(X_1^2) - [E(X_1)]^2$$

Therefore, using the definition of the expected value of a (function of a) discrete random variable, we have

$$E(X_1^2) = \sum_{k=0}^{3} k^2 \cdot P(X = k) = \frac{\theta}{3} + \frac{8}{3} \cdot (1 - \theta) + 3 \cdot (1 - \theta) = \frac{17}{3} - \frac{16}{3} \cdot \theta$$

Finally, the variance of X_1 is

$$\mathrm{Var}(X_1) = \frac{17}{3} - \frac{16}{3} \cdot \theta - \left(\frac{7}{3} - 2 \cdot \theta\right)^2 = -4\theta^2 + 4\theta + \frac{2}{9}$$

which means that the variance of $\widehat{\mu}$ is (here, $n = 10$)

$$\mathrm{Var}(\widehat{\mu}) = \frac{1}{10} \cdot \left(-4\theta^2 + 4\theta + \frac{2}{9}\right) = -\frac{2}{5} \cdot \theta^2 + \frac{2}{5} \cdot \theta + \frac{1}{45}$$

Exercise 1

Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. with density function

$$f(x \mid \theta) = e^{-(x-\theta)}, \quad x \geq \theta$$

and $f(x \mid \theta) = 0$ otherwise.

Find the method of moments estimate of θ.

Hint: Solve

$$\mathrm{E}(X) = \int_{\theta}^{+\infty} x \cdot f(x \mid \theta) dx$$

Exercise 2

Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables on the interval [0, 1] with the density function

$$f(x|\alpha) = \frac{\Gamma(3\alpha)}{\Gamma(\alpha)\Gamma(2\alpha)} x^{\alpha-1}(1-x)^{2\alpha-1}$$

where $\alpha > 0$ is a parameter to be estimated from the sample. It can be shown

that

$$E(X) = \frac{1}{3}$$

$$\mathrm{Var}(X) = \frac{2}{9(3\alpha+1)}$$

Find the method of moments estimate of α.