Maximum Likelihood Estimation
The Normal Distribution

Gary Schurman MBE, CFA

December, 2022

In this white paper we will use Maximum Likelihood Estimation to estimate the mean and variance of a normal distribution. To that end we will work through the following hypothetical problem...

Our Hypothetical Problem

Five random numbers are pulled from a normal distribution with mean 3.00 and variance 10.00. Using this random number sample, approximate the parameters of the normal distribution that generated these random numbers.

Table 1: Random Sample

Observation	Random Number
x_1	1.12
x_2	1.41
x_3	2.93
x_4	5.40
x_5	2.76

Question: What are the estimated mean and variance of the normal distribution that generated the sample above?

Sample Mean and Variance

We will define the variable N to be sample size, the variable $\bar{m}$ to be sample mean, and the variable $\bar{v}$ to be sample variance. The equations for sample mean and variance are...

$$\bar{m} = \frac{1}{N} \sum_{i=1}^{N} x_i \text{ ...and... } \bar{v} = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{m})^2 \tag{1}$$

Using Equation (1) above the sample mean and variance are...

$$\bar{m} = \frac{1}{5} \times 13.62 = 2.72 \text{ ...and... } \bar{v} = \frac{1}{5} \times 11.50 = 2.30 \tag{2}$$

The Probability Density Function

We will define the variables m and v to be the mean and variance, respectively, of the normal distribution. We will define the function $f(x)$ to be the probability density function of a normally-distributed random number x. Using these parameters the equation for the probability density function of the normal distribution is... [1]

$$f(x_i) = \sqrt{\frac{1}{2\,\pi\,v}} \, \text{Exp}\left\{ -\frac{1}{2\,v} (x_i - m)^2 \right\} \tag{3}$$

The equation for the natural log of Equation (3) is...

$$\ln\Big(f(x_i) \Big) = \frac{1}{2} \ln\left(\frac{1}{2\,\pi\,v} \right) - \frac{1}{2\,v} (x_i - m)^2 \tag{4}$$

The Maximum Log Likelihood Function

Maximum likelihood estimation is a method of estimating the parameters of an assumed probability distribution, given some observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable.

We will define the function ML to be the maximum likelihood function. Given that the random observations in our sample are independent, using Equation (3) above the joint probability of observing all of the data points in our sample is...

$$ML = \prod_{i=1}^{N} f(x_i) \tag{5}$$

We will define the variable MLL to be the maximum log likelihood function. Using Equation (5) above the equation for the maximum log likelihood function is...

$$MLL = \ln\left(\prod_{i=1}^{N} f(x_i)\right) = \sum_{i=1}^{N} \ln(f(x_i)) \tag{6}$$

Using Appendix Equations (17) and (18) below, the equations for the derivatives of Equation (6) above with respect to the mean (m) and variance (v) are...

$$\frac{\delta\, MLL}{\delta\, m} = \sum_{i=1}^{N} \frac{x_i - m}{v} \text{ ...and... } \frac{\delta\, MLL}{\delta\, v} = \sum_{i=1}^{N} \frac{1}{2\,v}\left[\frac{(x_i - m)^2}{v} - 1\right] \tag{7}$$

The estimate the mean and variance of the normal distribution that is consistent with our hypothetical problem above we want to maximize the joint probability of observing all of the data points in our sample. We do this by setting the derivatives in Equation (7) above to zero and then solving for m (mean) and v (variance).

The Solution To Our Hypothetical Problem

We will start by setting the first derivative in Equation (7) above to zero. The equation that we want to solve is...

$$\frac{\delta\, MLL}{\delta\, m} = \sum_{i=1}^{N} \frac{x_i - m}{v} = \frac{1}{v}\left(\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} m\right) = 0 \tag{8}$$

If we multiple both sides of Equation (8) above by v then we can rewrite that equation as...

$$\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} m = 0 \text{ ...such that... } \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} m \tag{9}$$

Using Equation (1) above we can make the following definitions...

$$\sum_{i=1}^{N} x_i = N\,\bar{m} \text{ ...and... } \sum_{i=1}^{N} m = N\,m \tag{10}$$

Using Equations (9) and (10) above the estimated value of the parameter m (mean) is...

$$N\,m = N\,\bar{m} \text{ ...such that... } m = \bar{m} \tag{11}$$

We then set the second derivative in Equation (7) above to zero. The equation that we want to solve is...

$$\frac{\delta\, MLL}{\delta\, v} = \sum_{i=1}^{N} \frac{1}{2\,v}\left[\frac{(x_i - m)^2}{v} - 1\right] = \frac{1}{2\,v}\sum_{i=1}^{N} \frac{(x_i - m)^2 - v}{v} = \frac{1}{2\,v^2}\left(\sum_{i=1}^{N}(x_i - m)^2 - \sum_{i=1}^{N} v\right) = 0 \tag{12}$$

If we multiple both sides of Equation (12) above by $2\,v^2$ then we can rewrite that equation as...

$$\sum_{i=1}^{N}(x_i - m)^2 - \sum_{i=1}^{N} v = 0 \text{ ...such that... } \sum_{i=1}^{N}(x_i - m)^2 = \sum_{i=1}^{N} v \tag{13}$$

Using Equation (1) above we can make the following definitions...

$$\sum_{i=1}^{N}(x_i - m)^2 = N\,\bar{v} \text{ ...and... } \sum_{i=1}^{N} v = N\,v \tag{14}$$

Using Equations (13) and (14) above the estimated value of the parameter m (mean) is...

$$N\,v = N\,\bar{v} \text{ ...such that... } v = \bar{v} \tag{15}$$

Conclusion: Per Equation (11) above the Maximum Likelihood estimate of the normal distribution mean is the sample mean. Per Equation (15) above the Maximum Likelihood estimate of the normal distribution variance is the sample variance. Therefore the answer to our hypothetical problem is...

$$m = 2.72 \text{ (Actual mean} = 3.00) \text{ ...and... } v = 2.30 \text{ (Actual variance} = 3.00) \tag{16}$$

References

[1] Gary Schurman, *The Calculus of the Normal Distribution - The Mathematics,* Octomer, 2010.

[2] Gary Schurman, *The Newton-Raphson Method - Solving Multivariate Equations,* March, 2016.

Appendix

A. The derivative of the Equation (4) above with respect to the variable m (mean) is...

$$\frac{\delta}{\delta m}\left[\frac{1}{2}\ln\left(\frac{1}{2\,\pi\,v}\right) - \frac{1}{2\,v}(x_i - m)^2\right] = 0 - \frac{1}{2\,v}\bigg(2\,(x_i - m) \times -1\bigg) = \frac{x_i - m}{v} \tag{17}$$

B. The derivative of the Equation (4) above with respect to the variable v (variance) is...

$$\frac{\delta}{\delta v}\left[\frac{1}{2}\ln\left(\frac{1}{2\,\pi\,v}\right) - \frac{1}{2\,v}(x_i - m)^2\right] = \left(\frac{1}{2} \times -\frac{1}{v}\right) - \left(-\frac{1}{2\,v^2} \times (x_i - m)^2\right) = \frac{1}{2\,v}\left[\frac{(x_i - m)^2}{v} - 1\right] \tag{18}$$