Finding the MLE for the mean and standard deviation given a normal distribution

Wednesday, November 3, 2021 8:07 AM

Let's assume, we have N samples - $\{x_1, x_2, \ldots, x_N\}$. The likelihood function is defined by

$$L(x_1, x_2, \ldots, x_N; m, \sigma) = \prod_{i=1}^{N} f(x_i) = f(x_1) \times f(x_2) \times \cdots \times f(x_N)$$

find the MLE for m and σ.

where $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$

(the normal distribution).

To make finding the maximums for μ and σ easier, we can take the log of the function L:

$$\ln(L(x_1, x_2, \ldots, x_N; \mu, \sigma)) = \sum_{i=1}^{N} \ln(f(x))$$

$$= \sum_{i=1}^{N} \ln\left((2\pi\sigma^2)^{-1/2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)$$

$$= \sum_{i=1}^{N} \left(-1/2 \ln(2\pi\sigma^2) + \ln\left(e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)\right)$$

$$= \sum_{i=1}^{N} \left(-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$= -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{N} (x_i-\mu)^2$$

Now we can take the derivative with respect to μ and set it equal to 0 to find where we will have our maximum μ:

$$\frac{\partial \ln(L(x_1, x_2, \ldots, x_n; \mu, \sigma))}{\partial \mu} = -\frac{1}{2\sigma^2} \sum_{i=1}^{N} \frac{\partial}{\partial \mu} (x_i-\mu)^2 = 0$$

$$-\frac{1}{2\sigma^2} \sum_{i=1}^{N} 2(x_i-\mu) \cdot -1 = 0$$

$$\frac{1}{\sigma^2} \sum_{i=1}^{N} (x_i-\mu) = 0$$

$$\sum_{i=1}^{N} (x_i-\mu) = 0$$

$$-N\mu + \sum_{i=1}^{N} x_i = 0$$

$$i=1$$

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Similarly, we can take the derivative with respect to σ^2 and set it equal to zero to find the equation for σ. To make this easier, we will write σ^2 as Z.

$$\frac{\partial \ln(L(x_1, x_2, \ldots, x_N; \mu, \sigma))}{\partial Z} \frac{\partial}{\partial Z}\left(-\frac{N}{2}\ln(2\pi Z) - \frac{1}{2Z}\sum_{i=1}^{N}(x_i-\mu)^2\right) = 0$$

$$= -\frac{N}{2} \cdot \frac{2\pi}{2\pi Z} - \frac{1}{2}\frac{\partial}{\partial Z}\left((Z)^{-1}\sum_{i=1}^{N}(x_i-\mu)^2\right) = 0$$

$$= -\frac{N}{2Z} - \left(-\frac{1}{2} \cdot (Z)^{-2}\sum_{i=1}^{N}(x_i-\mu)^2\right) = 0$$

$$= \frac{-N}{2Z} + \frac{1}{2Z^2}\sum_{i=1}^{N}(x_i-\mu)^2 = 0$$

$$= -N + \frac{1}{Z}\sum_{i=1}^{N}(x_i-\mu)^2 = 0$$

$$N = \frac{1}{Z}\sum_{i=1}^{N}(x_i-\mu)^2$$

$$Z = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$$

Substitute $Z = \sigma^2$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$$

For an example of these equations,

I used numpy to generate these 10 "random" numbers that follow the normal distribution:

[-4.57, -4.47, 1.32, -1.30, 3.63, 2.07, -1.70, .80, 2.93, 1.21]

To find the mean (μ):

$$\mu = \frac{1}{N}\sum_{i=1}^{N}(x_i)$$

$$= \frac{1}{10}\left(-4.57 + {}^{-}4.47 + 1.32 - 1.30 + 3.63 + 2.07 - 1.70 + 0.80 + 2.93 + 1.21\right)$$

$$= \frac{1}{10}\cdot(-0.08)$$

$$= -0.008$$

To find the standard deviation (σ):

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

To make this easier to write, I am going to solve for variance first (σ^2)

$$\sigma^2 = \frac{1}{10}\Big((-4.57 + 0.008)^2 + (-4.47 + 0.008)^2 + (1.32 + 0.008)^2 + (-1.30 + 0.008)^2 + (3.63 + 0.008)^2 + (2.07 + 0.008)^2 + (-1.70 + 0.008)^2 + (0.80 + 0.008)^2 + (2.93 + 0.008)^2 + (1.21 + 0.008)^2$$

$$= \frac{1}{10}\Big((-4.562)^2 + (-4.462)^2 + (1.328)^2 + (-1.292)^2 + (3.638)^2 + (2.078)^2 + (-1.692)^2 + (0.808)^2 + (2.938)^2 + (1.218^2)\Big)$$

$$= \frac{1}{10}(20.81 + 19.91 + 1.76 + 1.67 + 13.24 + 4.32 + 2.86 + .65 + 8.63 + 1.48)$$

$$= \frac{1}{10}(75.33)$$

$$\sigma^2 = 7.533$$

$$\sigma = \sqrt{7.533} = 2.74$$

For context, when generating these numbers, I used $N(\mu = 0, \sigma = 2)$.