hw5-report

Jack McShane

2022-03-26

1. Let $Y \sim Norm(\mu, \sigma^2)$. Write down its pdf.

$$\boxed{f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(y-\mu)^2}{2\sigma^2}}}$$

2. Let $Y_1, Y_2, \ldots, Y_n \sim_{iid} N(\mu, \sigma^2)$, then write down the joint distribution $f(y_1, \ldots, y_n)$ when $Y_1 = y_1, \ldots, Y_n = y_n$ are observed.

Given that the random variables $Y_1, \ldots, Y_n$ are independent and identically distributed, the joint distribution resolves as so:

$$\begin{aligned} f(y_1, y_2, \ldots, y_n) &= f(y_1) * f(y_2) * \cdots * f(y_n) \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_1-\mu)^2}{2\sigma^2}} \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_2-\mu)^2}{2\sigma^2}} \times \cdots \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_n-\mu)^2}{2\sigma^2}} \\ &\boxed{= (\frac{1}{\sqrt{2\pi}\sigma})^n \times e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}} \end{aligned}$$

3. This joint distribution is a function of μ and σ, when $Y_1 = y_1, \ldots, Y_n = y_n$ are observed. It is called the Likelihood function. Write down this Likelihood function $L(\mu, \sigma)$.

Given that we have values for $y_1, y_2, \ldots, y_n$, the equation above can now be written as a function of μ and σ^2, often referred to as the likelihood. This likelihood function will take the following form:

$$\boxed{L(\mu, \sigma^2 | y_1, y_2, \ldots, y_n) = (\frac{1}{\sqrt{2\pi}\sigma})^n \times e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}}$$

4. Write down the log likelihood function, $l(\mu,\sigma) = logL(\mu,\sigma)$, and negative log likelihood function $-l(\mu,\sigma)$.

$$
\begin{aligned}
l(\mu,\sigma) &= logL(\mu,\sigma|y_1,y_2,\ldots,y_n) \\
&= log[(\frac{1}{\sqrt{2\pi}\sigma})^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}] \\
&= nlog(\frac{1}{\sqrt{2\pi}\sigma}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2 \\
&= -nlog(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2 \\
&\boxed{= -nlog(\sqrt{2\pi}) - nlog(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}
\end{aligned}
$$

$$
\begin{aligned}
-l(\mu,\sigma^2) &= -[-nlog(\sqrt{2\pi}) - nlog(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2] \\
&\boxed{= nlog(\sqrt{2\pi}) + nlog(\sigma) + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}
\end{aligned}
$$

5. The maximum likelihood estimator of μ and σ is: $(\hat{\mu},\hat{\sigma}) = argmaxL(\mu,\sigma)$. Explain that it is equivalent to the following equations:

$$(\hat{\mu},\hat{\sigma}) = argmax[l(\mu,\sigma)] = argmin[-l(\mu,\sigma)]$$

The first equation, $(\hat{\mu},\hat{\sigma}) = argmax[l(\mu,\sigma)]$, is equivalent to the original due to the fact that logarithmic functions are monotonically increasing (i.e. the value of the function is forever increasing over its range of x). This property allows us to apply a logarithmic transformation, but preserve the values of the parameters, in this case μ and σ that maximize the likelihood function as they will also be the values that maximize the log likelihood function.

The second function, the negative log likelihood, is simply a reflection across the x-axis, result of which is the previously maximum values of the log likelihood function now represent the minimum values of the *negative* log likelihood function. It therefore follow that the values of μ and σ which minimize the negative log likelihood function are the same values that maximize both the likelihood and the log likelihood functions.

6. Explain that the maximum likelihood estimator and least squared estimator of μ are the same.

With the maximum likelihood estimator, we are trying to find the value of μ that maximizes this expression: $= -nlog(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2$. The variable μ is present only in the latter term of the expression and because we are maximizing the expression in terms of μ, this is the only term we need consider. Ignoring the multiplying constant, this simply leaves us with the term $\sum_{i=1}^{n}(y_i - \mu)^2$ which mirrors the least squared estimator equation. Finally, because of the negative multiplier, we need to minimize the term to maximize the overall expression, leaving us with the least squared estimator, that is, $argmin \sum_{i=1}^{n}(y_i - \mu)^2$.

7. Consider the car price example, and let Y be the price (in hundreds). Assume $Y \sim N(\mu, \sigma^2)$.

a) Read the data in R.

```
head(carprice)
```

```
##   Car Age Miles Price
## 1   1   5    57    85
## 2   2   4    40   103
## 3   3   6    77    70
## 4   4   5    60    82
## 5   5   5    49    89
## 6   6   5    47    98
```

b) Find $\hat{\mu}$ and $\hat{\sigma}$. Show your results.

```
opt.out <- optim(c(100, 30), fn.logL, y=carprice[,4])

mu.hat <- opt.out$par[1]
sigma.hat <- opt.out$par[2]
```

```
## [1] "Mean estimation:  88.6327028130927"
```

```
## [1] "Std. Dev. estimation:  29.7119482759126"
```

8. (Optional) Use calculus to show that for MLE:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

and

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

Show your work.

From the Least Squares equation (you could also take the partial of the log likelihood):

$$\frac{d}{d\mu}\sum_{i=1}^{n}(y_i-\mu)^2 = 0$$

$$\sum_{i=1}^{n}(2)(y_i-\mu)(-1) = 0$$

$$-2\sum_{i=1}^{n}(y_i-\mu) = 0$$

$$\sum_{i=1}^{n}(y_i-\mu) = 0$$

$$\sum_{i=1}^{n} y_i - n\mu = 0$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\boxed{\hat{\mu} = \bar{y}}$$

From the log likelihood equation:

$$\frac{\partial l}{\partial \sigma}[-nlog(\sqrt{2\pi}) - nlog(\sigma) - \frac{1}{2}\sigma^{-2}\sum_{i=1}^{n}(y_i-\mu)^2] = 0$$

$$-\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i-\mu)^2 = 0$$

$$n\sigma^{-1} = \sigma^{-3}\sum_{i=1}^{n}(y_i-\mu)^2$$

$$n\sigma^3\sigma^{-1} = \sigma^3\sigma^{-3}\sum_{i=1}^{n}(y_i-\mu)^2$$

$$\boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})}$$