Group SBS, Sheet 03, Exercise 01

November 20, 2021

Contributors

Sreyas Sony (sony@uni-potsdam.de), Binoy Chacko (chacko@uni-potsdam.de), Dinesh Kumar (kumar@uni-potsdam.de) Sanika Nair (nair@uni-potsdam.de)

Solution

Given $X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2)$ are independent and identically distributed,

$$f(x_i; \mu, \sigma^2) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left(\frac{-1}{2\sigma^2}(x_i - \mu)^2\right)$$

The Maximum-Likelihood estimator is defined as,

$$\widehat{\theta}_{ML} = \arg\max L(\theta)$$

where, $L(\theta)$ is the Likelihood function defined as,

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) \text{ ....because } X_i \text{ are i.i.d.}$$

Here,

$$\begin{aligned}
L(\mu, \sigma^2) &= \prod_{i=1}^{n} f(x_i; \mu, \sigma^2) \\
&= \frac{1}{(2\pi)^{1/2}\sigma} \exp\left(\frac{-1}{2\sigma^2}(x_1 - \mu)^2\right) \times ... \times \frac{1}{(2\pi)^{1/2}\sigma} \exp\left(\frac{-1}{2\sigma^2}(x_n - \mu)^2\right) \\
&= (2\pi)^{-n/2}\sigma^{-n} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right)
\end{aligned}$$

Taking the log of $L(\mu, \sigma^2)$, because it would be much more convenient to differentiate the sums than products. Since $\log(x)$ is an increasing function, solving $\arg\max(L(\theta))$ and $\arg\max(\log L(\theta))$ would give the same result.

$$\begin{aligned}\log\left(L(\mu,\sigma^2)\right) &= \log\left((2\pi)^{-n/2}\sigma^{-n}\exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right)\\ &= -\frac{n}{2}\log(2\pi) - n\log(\sigma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\end{aligned}$$

1. σ^2 is known and $\mu \in \mathbb{R}$ is unknown

Differentiating $\log(L(\mu, \sigma^2))$ w.r.t. μ,

$$\begin{aligned}\frac{\partial}{\partial\mu}\left(\log(L(\mu,\sigma^2))\right) &= \frac{\partial}{\partial\mu}\left(-\frac{n}{2}\log(2\pi)\right) - \frac{\partial}{\partial\mu}\left(n\log(\sigma)\right) - \frac{1}{2\sigma^2}\frac{\partial}{\partial\mu}\left(\sum_{i=1}^{n}x_i^2 - 2\mu\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\mu^2\right)\\ &= 0 - 0 - 0 + \frac{1}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{1}{\sigma^2}\mu n\end{aligned}$$

Equating $\frac{\partial}{\partial\mu}\left(\log(L(\mu,\sigma^2))\right)$ to 0,

$$\begin{aligned}\frac{1}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{1}{\sigma^2}\mu n = 0\\ \Rightarrow \widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i\end{aligned}$$

To check if $\widehat{\mu}$ maximizes the likelihood, taking the second derivative of $\log(L(\mu, \sigma^2))$ w.r.t. μ,

$$\begin{aligned}\frac{\partial^2}{\partial\mu^2}\left(\log(L(\mu,\sigma^2))\right) &= \frac{\partial}{\partial\mu}\left(\frac{1}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{1}{\sigma^2}\mu n\right)\\ &= 0 - \frac{n}{\sigma^2}\\ &= -\frac{n}{\sigma^2} < 0 \text{ ...because } n \text{ and } \sigma^2 \text{ are always positive.}\end{aligned}$$

Thus, $\widehat{\mu}$ maximizes the likelihood.

2. $\mu \in \mathbb{R}$ **is known and** $\sigma^2 > 0$ **is unknown**

Differentiating $\log(L(\mu, \sigma^2))$ w.r.t. σ^2,

$$
\begin{aligned}
\frac{\partial}{\partial \sigma^2}\left(\log(L(\mu,\sigma^2))\right) &= \frac{\partial}{\partial \sigma^2}\left(-\frac{n}{2}\log(2\pi)\right) - \frac{\partial}{\partial \sigma^2}\left(\frac{n}{2}\log(\sigma^2)\right) - \sum_{i=1}^{n}(x_i-\mu)^2 \frac{\partial}{\partial \sigma^2}\left(\frac{1}{2\sigma^2}\right) \\
&= 0 - \frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

Equating $\frac{\partial}{\partial \sigma^2}\left(\log(L(\mu,\sigma^2))\right)$ to 0,

$$
\begin{aligned}
-\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2 &= 0 \\
-n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2 &= 0 \\
\Rightarrow \widehat{\sigma^2} &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

To check if $\widehat{\sigma^2}$ maximizes the likelihood, taking the second derivative of $\log(L(\mu, \sigma^2))$ w.r.t. σ^2,

$$
\begin{aligned}
\frac{\partial^2}{\partial (\sigma^2)^2}\left(\log(L(\mu,\sigma^2))\right) &= \frac{\partial}{\partial \sigma^2}\left(-\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2\right) \\
&= \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

Substituting the value of $\widehat{\sigma^2}$ in σ^2,

$$
\begin{aligned}
&= \frac{n^3}{2\left[\sum_{i=1}^{n}(x_i-\mu)^2\right]^2} - \frac{n^3\sum_{i=1}^{n}(x_i-\mu)^2}{\left[\sum_{i=1}^{n}(x_i-\mu)^2\right]^3} \\
&= \frac{n^3 - 2n^3}{2\left[\sum_{i=1}^{n}(x_i-\mu)^2\right]^2} \\
&= \frac{-n^3}{2\left[\sum_{i=1}^{n}(x_i-\mu)^2\right]^2} < 0
\end{aligned}
$$

Thus, $\widehat{\sigma^2}$ maximizes the likelihood.