Example: DC Level in White Gaussian Noise

Seungjin Choi
Department of Computer Science
POSTECH, Korea

Example 1: Consider the multiple observations

$$x_t = \theta + \epsilon_t, \quad t = 1, 2, \ldots, N, \tag{1}$$

where θ is a constant (DC level that we try to estimate) and ϵ_t is white Gaussian noise, $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$. We choose two different estimators, $\hat{\theta}$ and $\check{\theta}$, given by

$$\hat{\theta} = \frac{1}{N} \sum_{t=1}^{N} x_t, \tag{2}$$

$$\check{\theta} = x_1. \tag{3}$$

The first estimator $\hat{\theta}$ is the sample mean and the second estimator $\check{\theta}$ takes the value of the first observation x_1, so it does not make use of all the data. Our interests are:

- How close will estimators be to θ?
- Which one of these two estimators will be better?

Assessment: We first show that the mean of each estimator is true value,

$$\begin{aligned}
\left\langle \hat{\theta} \right\rangle &= \left\langle \frac{1}{N} \sum_{t=1}^{N} x_t \right\rangle \\
&= \frac{1}{N} \sum_{t=1}^{N} \langle x_t \rangle \\
&= \theta, \\
\left\langle \check{\theta} \right\rangle &= \langle x_1 \rangle \\
&= \theta.
\end{aligned}$$

Thus, on the average the estimators produce the true value.

Second, we compute the variances of estimators,

$$\begin{aligned}\mathrm{var}\left(\hat{\theta}\right) &= \mathrm{var}\left(\frac{1}{N}\sum_{t=1}^{N} x_t\right) \\ &= \frac{1}{N^2}\sum_{t=1}^{N}\mathrm{var}\left(x_t\right) \\ &= \frac{1}{N^2}N\sigma^2 \\ &= \frac{\sigma^2}{N},\end{aligned}$$

and

$$\begin{aligned}\mathrm{var}\left(\check{\theta}\right) &= \mathrm{var}\left(x_1\right) \\ &= \sigma^2.\end{aligned}$$

Thus one can see that $\mathrm{var}\left(\hat{\theta}\right) > \mathrm{var}\left(\check{\theta}\right)$, which implies that $\hat{\theta}$ is better than $\check{\theta}$.

Finally, we show that $\hat{\theta}$ is the maximum likelihood estimator (MLE). Since ϵ_t is a white Gaussian sequence, a single factor of the likelihood is given by

$$p(x_t;\theta) = \mathcal{N}(\theta,\sigma^2). \tag{4}$$

Thus the log-likelihood is given by

$$\begin{aligned}\mathcal{L} &= \log\left(\prod_{t=1}^{N} p(x_t;\theta)\right) \\ &= \sum_{t=1}^{N}\log p(x_t;\theta) \\ &= -\frac{N}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{t=1}^{N}(x_t-\theta)^2. \qquad (5)\end{aligned}$$

The MLE θ_{ml} is given by

$$\theta_{ml} = \arg\max_{\theta}\mathcal{L}. \tag{6}$$

It follows from $\frac{\partial\mathcal{L}}{\partial\theta} = 0$ that we have

$$\theta_{ml} = \frac{1}{N}\sum_{t=1}^{N} x_t. \tag{7}$$