PRML (1.58) : Why $\mathbb{E}[\sigma^2_{sample}] = \frac{N-1}{N}\sigma^2$?

Shoichiro Yamanishi

March 31, 2020

Abstract

This equation (1.58) in PRML[1] is apparently the first problem many readers including me encounter. It's in pp27, Chapter 1, Introduction. It is not apparent to me why it holds, but no proof is given in the book. The following is my take on the proof why the sample variance is biased.

1 Because $\sigma_{x_i x_j} = 0$, if x_1 and x_2 are independent.

Assume x is a random variable with distribution $p(x)$, and $x_1, x_2, \cdots, x_N$ are i.i.d. samples of x, μ_s and σ^2_s denote the sample mean and the sample variance respectively. In the following argument all the expectations are over $p(x)$.

$$\begin{aligned}
\mathbb{E}[\sigma^2_s] &= \mathbb{E}[\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_s)^2] \\
&= \frac{1}{N}\mathbb{E}[\sum_{i=1}^{N} x_i^2] - \frac{2}{N}\mathbb{E}[\mu_s \sum_{i=1}^{N} x_i] + \frac{1}{N}\mathbb{E}[\sum_{i=1}^{N}\mu_s^2] \\
&= \frac{1}{N}\sum_{i=1}^{N}(\mathbb{E}[x_i^2]) - \frac{2}{N}\mathbb{E}[\mu_s N \mu_s] + \frac{1}{N}\mathbb{E}[N\mu_s^2] \\
&= \mathbb{E}[x^2] - \mathbb{E}[\mu_s^2].
\end{aligned} \tag{1}$$

The conversion from the 2nd to the 3rd row in the first term is due to i.i.d. on x_i.

From the formula of variance,

$$\begin{aligned}
\sigma^2_x &= \mathbb{E}[x^2] - \mathbb{E}[x]^2 \\
&= \mathbb{E}[x^2] - \mu^2
\end{aligned} \tag{2}$$

and,

$$\begin{aligned}
\sigma^2_{\mu_s} &= \mathbb{E}[\mu_s^2] - \mathbb{E}[\mu_s]^2 \\
&= \mathbb{E}[\mu_s^2] - \mu_s^2 \\
&= \mathbb{E}[\mu_s^2] - \mu^2
\end{aligned} \tag{3}$$

as $\mathbb{E}[\mu_s] = \mu$ (no bias on the mean).

From equation 2 and 3,

$$\sigma_x^2 - \sigma_{\mu_s}^2 = \mathbb{E}[x^2] - \mathbb{E}[\mu_s^2]. \tag{4}$$

From equation 1 and 4,

$$\mathbb{E}[\sigma_s^2] = \sigma_x^2 - \sigma_{\mu_s}^2. \tag{5}$$

Here by definition $\sigma_{\mu_s}^2 = VAR[\frac{1}{N}\sum_{i=1}^{N} x_i] = \frac{1}{N^2}VAR[\sum_{i=1}^{N} x_i]$, and from i.i.d. assumption,

$$\sigma_{\mu_s}^2 = \frac{1}{N^2}VAR[\sum_{i=1}^{N} x] = \frac{1}{N}VAR[x] = \frac{1}{N}\sigma_x^2. \tag{6}$$

From equation 5 and 6,

$$\mathbb{E}[\sigma_s^2] = \frac{N-1}{N}\sigma_x^2. \tag{7}$$

Q.E.D.

References

[1] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1 edition, 2007.