Abstract

In this issue, we mainly study some properties of Gaussian distribution

Assumption

Now given a bunch of data:

$$X = (x_1, x_2, \ldots, x_N)^T \tag{18}$$
$$x_i \in R^p \tag{19}$$

First, we assume our model: the Gauss linear model.

to simplify the derivation of formula, we set p equals 1, so

$$x \backsim N(\mu, \sigma^2) \tag{20}$$
$$\theta = (\mu, \sigma) \tag{21}$$

Next, we use maximum likelihood estimation (MLE) to get the expectation and variance based on this bunch of data

The likelihood function is given below:

$$\begin{aligned} p(X|\theta) &= log(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)) \\ &= \sum_{i=1}^{N} log(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)) \\ &= \sum_{i=1}^{N} log(\frac{1}{\sqrt{2\pi}}) - log(\sigma) - \frac{(x_i-\mu)^2}{2\sigma^2} \end{aligned} \tag{22}$$

Expectation

Next, we first use the maximum likelihood estimation to obtain the estimated value of the expected μ

$$\begin{aligned} \mu_{MLE} &= argmax(p(X|\theta)) \\ &= argmin(\sum_{i=1}^{N} (x_i - \mu)^2) \end{aligned} \tag{23}$$

By deriving the formula:

$$\sum_{i=1}^{N} 2(x_i - \mu) = 0 \tag{24}$$

$$\sum_{i=1}^{N} x_i - N\mu = 0 \tag{25}$$

$$\mu_{MLE} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{26}$$

Variance

Similarly, we use maximum likelihood estimation to estimate the variance σ

$$\begin{aligned} \sigma_{MLE} &= argmax(p(X|\theta)) \\ &= argmin(\sum_{i=1}^{N} log(\sigma) + \frac{(x_i - \mu)^2}{2\sigma^2}) \end{aligned} \tag{27}$$

Similarly, we derive the formula:

$$\sum_{i=1}^{N}[\frac{1}{\sigma} - \frac{(x_i - \mu)^2}{\sigma^3}] = 0 \tag{28}$$

Finally, we get the estimated value:

$$\sigma^2_{MLE} = \Sigma_{MLE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{29}$$

Biased Estimation

To verify whether an estimate is biased or unbiased, we only need to calculate the expectation of the estimate.

μ

$$\begin{aligned} E[\mu_{MLE}] &= E[\frac{1}{N}\sum_{i=1}^{N} x_i] \\ &= \frac{1}{N}\sum_{i=1}^{N} E[x_i] \\ &= \mu \end{aligned} \tag{30}$$

So μ_{MLE} is an unbiased estimate

σ

First we deform the estimate of σ

$$
\begin{aligned}
\sigma^2_{MLE} &= \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_{MLE})^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}(x_i^2 - 2x_i\mu_{MLE} + \mu^2_{MLE}) \\
&= \frac{1}{N}\sum_{i=1}^{N}x_i^2 - 2(\frac{1}{N}\sum_{i=1}^{N}x_i)\mu_{MLE} + \mu^2_{MLE} \\
&= \frac{1}{N}\sum_{i=1}^{N}x_i^2 - 2\mu^2_{MLE} + \mu^2_{MLE} \\
&= \frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2_{MLE} \\
&= (\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2) - (\mu^2_{MLE} - \mu^2)
\end{aligned} \tag{31}
$$

set $f_1 = (\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2)$, $f_2 = (\mu^2_{MLE} - \mu^2)$

so:

$$
\begin{aligned}
E[f_1] &= E[\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2] \\
&= E[\frac{1}{N}\sum_{i=1}^{N}(x_i^2 - \mu^2)] \\
&= \frac{1}{N}\sum_{i=1}^{N}E[x_i^2] - E[\mu^2] \\
&= \frac{1}{N}\sum_{i=1}^{N}E[x_i^2] - \mu^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}E[x_i^2] - (E[x_i])^2 \\
&= \sigma^2
\end{aligned} \tag{32}
$$

similarly:

$$
\begin{aligned}
E[f_2] &= E[\mu^2_{MLE} - \mu^2] \\
&= E[\mu^2_{MLE} - (E[\mu_{MLE}])^2] \\
&= Var[\mu_{MLE}] \\
&= Var[\frac{1}{N}\sum_{i=1}^{N}x_i] \\
&= \frac{1}{N^2}\sum_{i=1}^{N}Var[x_i] \\
&= \frac{1}{N}\sigma^2
\end{aligned} \tag{33}
$$

finally, adding f_1 and f_2 , we get:

$$E[\sigma^2_{MLE}] = \frac{N-1}{N}\sigma^2 \tag{34}$$

So our estimate of σ from the maximum likelihood estimate is slightly smaller than the true value, so it is biased.

The unbiased estimate of σ^2 is $\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_{MLE})^2$