0.1 Expectation and Variance

0.1.1 Abstract

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

0.1.2 Assumption

Now given a bunch of data:

$$X = (x_1, x_2, ..., x_N)^T$$

$$x_i \in \mathcal{R}^p$$

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

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

$$x \backsim N(\mu, \sigma^2)$$

$$\theta = (\mu, \sigma)$$

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:

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

0.1.3 Expectation

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

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

By deriving the formula:

$$\sum_{i=1}^{N} 2(x_i - \mu) = 0$$
$$\sum_{i=1}^{N} x_i - N\mu = 0$$
$$\mu_{MLE} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

0.1.4 Variance

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

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

Similarly, we derive the formula:

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

Finally, we get the estimated value:

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

0.1.5 Bias Estimation

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

μ

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

So μ_{MLE} is an unbiased estimation

σ

First we deform the estimate of σ

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

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

so:

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

similarly:

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

finally, adding f_1 and f_2 , we get:

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

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 $\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_{MLE})^2$