HMMA 307 : Modèles linéaires avancés

REstricted Maximum Likelihood (REML)

Ophélie Coiffier

https://github.com/opheliecoiffier/REML

Université de Montpellier

Table of Contents

Table of Contents

The main issue : biased variance

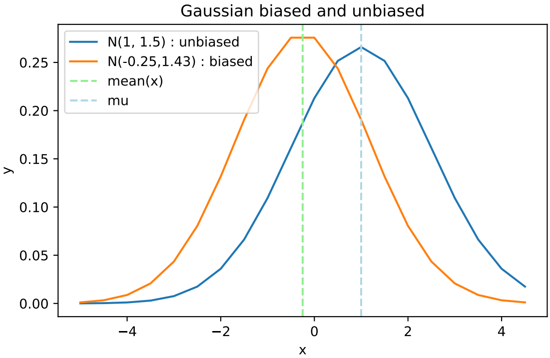

Figure: Difference between a Biased Normal distribution and an unbiased Normal distribution

Characterization of the biased variance

The expected value of the variance's biased estimator is not the variance estimator. That means :

Dimension : 1 the model writes out :

$$y = X\beta + \varepsilon$$

where y $\sim \mathcal{N}(\mu, \sigma^2)$ and $\varepsilon \sim \mathcal{N}(0, 1)$

$$\mathbb{E}[\widehat{\sigma^2}] = \frac{N-1}{N}\sigma^2 \neq \sigma^2$$

Characterization of the biased variance

The expected value of the variance's biased estimator is not the variance estimator. That means :

Dimension : 1 the model writes out :

$$y = X\beta + \varepsilon$$

where y $\sim \mathcal{N}(\mu, \sigma^2)$ and $\varepsilon \sim \mathcal{N}(0, 1)$

$$\mathbb{E}[\widehat{\sigma^2}] = \frac{N-1}{N}\sigma^2 \neq \sigma^2$$

Dimension : k the model writes out :

$$y = X\beta + \varepsilon$$

where y $\sim \mathcal{N}(X\beta, \sigma^2 I_k)$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_k)$

$$\mathbb{E}[\widehat{\sigma^2}] = \frac{N-k}{N}\sigma^2 \neq \sigma^2$$

Characterization of the biased variance

The expected value of the variance's biased estimator is not the variance estimator. That means :

Dimension : 1 the model writes out :

$$y = X\beta + \varepsilon$$

where y $\sim \mathcal{N}(\mu, \sigma^2)$ and $\varepsilon \sim \mathcal{N}(0, 1)$

$$\mathbb{E}[\widehat{\sigma^2}] = \frac{N-1}{N}\sigma^2 \neq \sigma^2$$

Dimension : k the model writes out :

$$y = X\beta + \varepsilon$$

where y $\sim \mathcal{N}(X\beta, \sigma^2 I_k)$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_k)$

$$\mathbb{E}[\widehat{\sigma^2}] = \frac{N-k}{N}\sigma^2 \neq \sigma^2$$

It is an underestimation of the true variance. We need to choose $\widehat{\sigma^2} = \frac{1}{N-k}(y - X\hat{\beta})^T(y - X\hat{\beta})$ to find an unbiased variance.

The REML solution

The idea of the REML method is :

The REML solution

The idea of the REML method is :

1) Integrate the likelihood in relation to μ

The REML solution

The idea of the REML method is :

1) Integrate the likelihood in relation to μ
2) Calculate the log of the previous result using Taylor's development to simplify calculations

The REML solution

The idea of the REML method is :

1) Integrate the likelihood in relation to μ
2) Calculate the log of the previous result using Taylor's development to simplify calculations
3) Recognize the Maximum likelihood solution and the REML approach (the bias, the fee)

The REML solution

The idea of the REML method is :

1) Integrate the likelihood in relation to μ
2) Calculate the log of the previous result using Taylor's development to simplify calculations
3) Recognize the Maximum likelihood solution and the REML approach (the bias, the fee)
4) Determine the variance estimator

Table of Contents

An example : Data

Ind	Resp	Treat
1	10	0
1	25	1
2	3	0
2	6	1

In *Treat* column : 0 means that the individual gets the treatment and 1 means that it doesn't get the treatment.

An example : Codes

First, we use linear regression with Least Squared method (OLS)

```
linear_reg = sm.OLS(df.Resp,df.Treat)
linear_reg_fit = linear_reg.fit()
print(linear_reg_fit.summary())
```

Model : $Y_{Resp} = \mu + \beta X_{Treat} + \varepsilon$

An example : Codes

Then, we use mixed effects model with REML

```
mixed_random = smf.mixedlm("Resp~Treat",df,
            groups=df['Ind'])
mixed_fit = mixed_random.fit()
print(mixed_fit.summary())
```

Model : $Y_{Resp} = \mu + \beta X_{Treat} + \alpha X_{Ind} + \varepsilon$

An example : Results

Method	**Log-likelihood**	σ^2	σ_s^2
OLS linear regression	-14.23	6.95	-
REML mixed effects model	-7.89	6.00	8.15

Both methods have the same value for the coefficients : $\beta_1 = 6.0$ and $\beta_2 = 15.5$

Table of Contents

Conclusion

The REML method allows to estimate the variance without bias.

The log-likelihood is different because we have a second term with the REML method : a fee. But we have the same coefficients β.

Bibliography

- Coiffier, Ophélie. *REML to Python code*. https://github.com/opheliecoiffier/REML. 2020.
- Oskolkov, Nikolay. *Maximum Likelihood (ML) vs. REML*. https://towardsdatascience.com/maximum-likelihood-ml-vs-reml-78cf79bef2cf.
- Salmon, Joseph. *HMMA 307 - Modèles linéaires avancées*. 2019.