Exercices corrections

Correction exercice .1. ***(Linear Regression)*** *The linear model* $y = mx + b$ *can be solved using the least squares method. The formulas for* m *and* b *are:*

$$m = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$b = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

where n *is the number of data points,* (x_i, y_i)*. Substitute the given values:*

$$n = 5, \Sigma x = 15, \Sigma y = 35, \Sigma x^2 = 55, \Sigma xy = 135$$

$$m = \frac{(5 \times 135) - (15 \times 35)}{(5 \times 55) - (15)^2} = \frac{375}{20} = 18.75$$

$$b = \frac{(35 \times 55) - (15 \times 135)}{(5 \times 55) - (15)^2} = \frac{1925}{20} = 96.25$$

Therefore, the linear model is $y = 18.75x + 96.25$.

Correction exercice .2. ***(Polynomial Regression)*** *The polynomial model* $y = a_0 + a_1 x + a_2 x^2$ *can be solved using the least squares method. The formulas for* $a_0, a_1,$ *and* a_2 *are obtained by solving a system of linear equations. The system of equations is:*

$$\begin{aligned}
a_0 + a_1(1) + a_2(1)^2 &= 2 \\
a_0 + a_1(2) + a_2(2)^2 &= 5 \\
a_0 + a_1(3) + a_2(3)^2 &= 10 \\
a_0 + a_1(4) + a_2(4)^2 &= 17 \\
a_0 + a_1(5) + a_2(5)^2 &= 26
\end{aligned}$$

Solving this system gives:

$$a_0 = 0.5, \quad a_1 = 0, \quad a_2 = 1$$

Therefore, the polynomial model is $y = 0.5 + x^2$.

Correction exercice .3. ***(Exponential Regression)*** *The exponential model* $y = ae^{bx}$ *can be solved using the least squares method. Take the natural logarithm of both sides to linearize the model:*

$$\ln(y) = \ln(a) + bx$$

This is now in the form $\ln(y) = \alpha + \beta x$*, which is a linear model. Apply linear regression to find the values of* α *and* β*. Solving for* a *and* b *gives:*

$$a = e^{\alpha}, \quad b = \beta$$

Substitute the given values:

$$\alpha \approx 0.69, \quad \beta \approx 0.69$$

Therefore, the exponential model is $y \approx e^{0.69}e^{0.69x}$.

Correction exercice .4. ***(Multiple Linear Regression)*** *The multiple linear regression model* $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ *can be solved using the least squares method. The formulas for* $\beta_0, \beta_1,$ *and* β_2 *are obtained by solving a system of linear equations. The system of equations is:*

$$\begin{aligned} \beta_0 + \beta_1(1) + \beta_2(2) &= 5 \\ \beta_0 + \beta_1(2) + \beta_2(4) &= 8 \\ \beta_0 + \beta_1(3) + \beta_2(6) &= 10 \\ \beta_0 + \beta_1(4) + \beta_2(8) &= 12 \\ \beta_0 + \beta_1(5) + \beta_2(10) &= 14 \end{aligned}$$

Solving this system gives:

$$\beta_0 = 1, \quad \beta_1 = 2, \quad \beta_2 = 1$$

Therefore, the multiple linear regression model is $y = 1 + 2x_1 + x_2$.

Solutions to Exercises on Characteristics of an Estimator

Correction exercice .5. ***(Bias, Mean Squared Error, and Convergence)***

1. *To calculate the bias of* $\hat{\mu}$*, use the following formula:*

$$Bias(\hat{\mu}) = E(\hat{\mu}) - \mu$$

   If $\hat{\mu}$ *is unbiased, the bias should equal zero.*

2. *To calculate the mean squared error (MSE) of* $\hat{\sigma}^2$*, use the following formula:*

$$MSE(\hat{\sigma}^2) = E\left((\hat{\sigma}^2 - \sigma^2)^2\right)$$

3. *To show that* $\hat{\theta}_n$ *converges in probability to* θ *as* n *approaches infinity, you can use the definition of convergence in probability:*

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| < \epsilon) = 1, \text{ for all } \epsilon > 0$$

Correction exercice .6. ***(Fisher Information and Cramer-Rao Bound)***

1. *To derive the Fisher information $I(\theta)$ for the parameter θ, you need to find the second derivative of the log-likelihood function with respect to θ.*

2. *The Cramer-Rao Lower Bound (CRLB) for the variance of any unbiased estimator of θ is given by:*

$$Var(\hat{\theta}) \geq \frac{1}{nI(\theta)}$$

Correction exercice .7. ***(Efficiency)***

1. *Calculate the efficiency of $\hat{\theta}_1$ with respect to $\hat{\theta}_2$ using the formula:*

$$Efficiency(\hat{\theta}_1, \hat{\theta}_2) = \frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$$

An estimator with higher efficiency is more efficient.

Correction exercice .8. ***(Completeness)***

1. *A sufficient statistic T is complete if, for any measurable function $g(t)$, the expectation of $g(T)$ is zero only if $g(t)$ is zero with probability one:*

$$E[g(T)] = 0 \Rightarrow P(g(T) = 0) = 1$$

To show that a statistic is complete, you can use the definition and properties of completeness.

2. *To find a statistic that is minimal sufficient but not complete, consider a distribution where minimal sufficiency is achieved, but the distribution itself does not meet the completeness criterion. You can explore different probability distributions to illustrate this.*