Statistical Inference II - (B)

Time: 2h

Ex.1 You observe N pairs $(y_1, x_1), \dots, (y_N, x_N)$ assumed to be generated by following linear model

$$y_i = a + bx_i + \epsilon_i, \qquad \forall i = 1, \dots, N \tag{1}$$

where a, b are real parameters, ϵ_i are i.i.d. Gaussian[1] $\mathcal{N}(0, \sigma^2)$ random variables and $x_1, \dots, x_N$ are deterministic (i.e. they are **not** random variables, they are numbers!).

(a) Compute $\mathbf{E}(y_i)$, $\mathbf{Var}(y_i)$ and deduce the distribution of y_i.
(b) Are $y_1, \dots, y_N$ independent? Are they identically distributed?
(c) Write down the log-likelihood of $y_1, \dots, y_N$ and compute the ML estimate $\hat{a}_{ML}$ of a.
(d) By using the previous point, compute the ML estimate $\hat{b}_{ML}$ of b.
(e) What is the relation between the ML estimates $(\hat{a}_{ML}, \hat{b}_{ML})$ and the OLS estimates $(\hat{a}_{OLS}, \hat{b}_{OLS})$?

Ex.2 Given the standard multivariate liner regression model

$$Y = X\beta + \epsilon,$$

where $Y \in \mathbb{R}^N$, $X \in \mathbb{R}^{N \times p}$, $\beta \in \mathbb{R}^p$ and

$$\epsilon \sim \mathcal{N}(0, \sigma^2 I_N),$$

---

[1] Recall that the pdf of a random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ is

$$f(x) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

we saw in class that

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y$$

(a) Compute $\mathbf{E}(\hat{\beta}_{OLS})$. Is it an unbiased estimator?
(b) Compute $\mathbf{Var}(\hat{\beta}_{OLS})$.