Problem 1

Assume that the observation is

$$Y \mid \mu \sim \mathcal{N}(\mu, 1), \quad \mu \sim \mathcal{N}(0, 1),$$

and $\tau \in \{-1, 1\}$ independent of μ and Y, with equal probability. We consider the following two forecasts

$$\begin{cases} \text{Unfocused forecast:} & F_1 = \frac{1}{2}\left(\mathcal{N}(\mu, 1) + \mathcal{N}(\mu + \tau, 1)\right), \\ \text{Sign-reversed forecast:} & F_2 = \mathcal{N}(-\mu, 1). \end{cases}$$

(a) Show that the unfocused forecast is probabilistically calibrated but not marginally calibrated and the sign-reversed forecast is marginally calibrated but not probabilistically calibrated.

(b) Design and implement a simulation study in which you show probability integral transform histograms and assess marginal calibration for the two forecasts.

(c) Quantify and discuss the sharpness of the forecasts.

Problem 2

Let p be a forecast density and let y be the observed outcome. The linear score is defined as:

$$\text{LinS}(p, y) = -p(y).$$

That is, it evaluates the forecast p by penalising the value it assigns to the observed outcome y. At first glance, this score may seem intuitive: assigning more probability to the true outcome gives a smaller (better) score.

Is the linear score a proper scoring rule? Discuss the propriety of the linear score by considering the example where the true density is the standard normal and the predictive distribution of a forecast is the uniform on $(-\epsilon, \epsilon)$, where $\epsilon > 0$.

Problem 3

Let F be the cumulative distribution function of a probability distribution on $\mathbb{R}$, and let $y \in \mathbb{R}$ be an observation. The continuous ranked probability score (CRPS) is defined as:

$$\text{CRPS}(F, y) = \int_{-\infty}^{\infty} \{F(z) - \mathbb{I}\{y \le z\}\}^2 \, dz.$$

(a) Starting from this definition, we aim to show that the CRPS can be equivalently written as:

$$\text{CRPS}(F, y) = \text{E}_F(|X - y|) - \tfrac{1}{2}\text{E}_F(|X - X'|),$$

where $X, X' \sim F$ independently.

- Show that the expected absolute error between a random variable $X \sim F$ and a fixed point $y \in \mathbb{R}$ can be written as:

$$\text{E}(|X - y|) = \int_{-\infty}^{y} F(z) \, dz + \int_{y}^{\infty} (1 - F(z)) \, dz.$$

Hint: Split the expectation into two integrals over $[-\infty, y]$ and $[y, \infty]$, and change the order of integration.

- Show that the expected absolute difference between two independent random variables $X, X' \sim F$ can be written as:

$$\mathrm{E}|X - X'| = 2 \int_{-\infty}^{\infty} F(z)(1 - F(z))\, dz.$$

Hint: Express $\mathrm{E}(|X - X'|)$ *as a double integral over* $\mathbb{R}^2$ *and switch to a single integral via symmetry.*

- Using these two results, show the desired expression of the CRPS.

(b) Let F be a Gaussian mixture of the form

$$F = \sum_{i=1}^{m} w_i \mathcal{N}(\mu_i, \sigma_i^2), \quad \text{with } \sum_{i=1}^{m} w_i = 1, \quad w_i \geq 0.$$

Show that the CRPS for this Gaussian mixture can be expressed as

$$S\left(\sum_{i=1}^{m} w_i \mathcal{N}(\mu_i, \sigma_i^2), y\right) = \sum_{i=1}^{m} w_i A(y - \mu_i, \sigma_i^2) - \tfrac{1}{2} \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j A(\mu_i - \mu_j, \sigma_i^2 + \sigma_j^2),$$

where

$$A(\mu, \sigma^2) = 2\sigma\phi\left(\tfrac{\mu}{\sigma}\right) + \mu\left(2\Phi\left(\tfrac{\mu}{\sigma}\right) - 1\right),$$

and ϕ and Φ denote the standard normal PDF and CDF, respectively.

(c) Briefly interpret each term in the expression for the CRPS of a Gaussian mixture. What do they represent intuitively in terms of sharpness and calibration?