## Advanced theory of statistics – Homework 3, Winter 2025.

---

**Problem 1.** [Sequential testing in Gaussian case] Let $X, X_1, X_2, \dots$ be an i.i.d. sequence of normal variables $N(\mu, 1)$. Consider the simple binary testing problem: $H_0 : \mu = 0$, $H_1 : \mu = \mu_0 > 0$. For simplicity, let us specify equal probabilities of error, that is, $\alpha = 1 - \beta \ll \frac{1}{2}$.

(i) Show that the optimal cut-off for the (non-sequential) likelihood ratio test is to reject the null if $\log \mathrm{LR}_n \geq 0$. Conclude that the number of samples required for a specified size $\alpha$ is

$$n = \left\lceil \frac{2(\Phi^{-1}(1-\alpha))^2}{\mu_0} \right\rceil .$$

(ii) Develop the corresponding SPRT and approximate its expected stopping time using the formulas developed in the lecture. Use both $\mu_0$ and $\alpha$ as parameters that can take some fixed values. What do you observe?

(iii) Propose simulations that allow you to estimate the expected stopping time. What is the estimated variance?

(iv) Suppose that each additional sample incurs a cost $c > 0$. Consider the objective of minimizing the expected total cost, defined as:

$$\mathbb{E}[N] \cdot c + \mathbb{P}_0(\text{reject } H_0) + \mathbb{P}_1(\text{accept } H_0).$$

It is intuitively clear that a modified Wald test should aim to terminate earlier. We now modify the SPRT so that the stopping boundaries are of the form:

$$\gamma_0' = \frac{1-\beta}{1-\alpha} e^{-\lambda c}, \quad \gamma_1' = \frac{\beta}{\alpha} e^{\lambda c},$$

where $\lambda > 0$ is a parameter that accounts for the sampling cost. In simulations analyse the expected total cost for different $c$ and different values of $\lambda \in (0, \frac{1}{2}\mu_0^2)$.

**Problem 2.** (Requires material from later lectures) Let $X \in \mathbb{R}^{n \times d}$ be a random matrix with i.i.d. zero mean entries that are $\sigma$-sub-Gaussian. Denote by $S^{n-1}$ the unit sphere in $\mathbb{R}^n$ and by $\mathbb{B}_2^n$ the corresponding unit ball.

(i) Show that $u^T X v$ is a $\sigma$-sub-Gaussian random variable for any $u \in \mathbb{B}_2^n$, $v \in \mathbb{B}_2^d$.

(ii) The operator norm $\|X\|$ is defined as $\|X\| := \sup_{v \neq 0} \frac{\|Xv\|}{\|v\|}$. Show that

$$\|X\| = \max_{u \in S^{n-1}} \max_{v \in S^{d-1}} u^T X v = \max_{u \in \mathbb{B}_2^n} \max_{v \in \mathbb{B}_2^d} u^T X v.$$

(iii) Using (i) and (ii), show that there exists a constant $C > 0$ such that $\mathbb{E}\|X\| \leq C(\sqrt{n} + \sqrt{d})$.