Advanced theory of statistics – Homework 2, Winter 2025.

Problem 1. (15 points) Suppose $Y \sim N_n(\mu, I_n)$. Let $\mathbf{X} \in \mathbb{R}^{n\times p}$ be a fixed matrix with full column rank. In this exercise we consider estimators of the form $\hat{\mu} = \mathbf{X}\hat{\beta}$ for some estimator $\hat{\beta}$.

(i) Let $\widehat{\mu}^{\mathrm{LS}} = \mathbf{X}\widehat{\beta}^{\mathrm{LS}}$ be the corresponding estimator of the mean of Y based on the least squares estimator $\widehat{\beta}^{\mathrm{LS}} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top Y$. Find the risk (in terms of the square loss) of this estimator directly in terms of $C_0 = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$, μ, n, and p only.

(ii) Suppose $\mu = \mathbf{X}\beta^*$ for some $\beta^* \in \mathbb{R}^p$. Show that $R(\mu, \hat{\mu}^{\mathrm{LS}}) = p$.

(iii) Consider the ridge estimator $\hat{\mu}^{\mathrm{ridge}} = \mathbf{X}(\mathbf{X}^\top\mathbf{X} + \delta I_p)^{-1}\mathbf{X}^\top Y$ and the corresponding matrix $C_\delta = \mathbf{X}(\mathbf{X}^\top\mathbf{X} + \delta I_p)^{-1}\mathbf{X}^\top$, where $\delta \geq 0$. Show that $R(\mu, \hat{\mu}^{\mathrm{ridge}}) = \mathrm{tr}(C_\delta^2) + \|(I_n - C_\delta)\mu\|^2$.

(iv) Suppose that $\mathbf{X}^\top\mathbf{X} = I_p$ and that $\mu = \mathbf{X}\beta^*$ for some β^*. Show that for every $\beta^* \neq 0$ there exists $\delta > 0$ such that $R(\mathbf{X}\beta^*, \hat{\mu}^{\mathrm{ridge}}) < R(\mathbf{X}\beta^*, \hat{\mu}^{\mathrm{LS}})$.

Problem 2. (15 points) Consider $2m$ independent coins $X_i \sim \mathrm{Bern}(\theta_i)$ $i = 1, \ldots, m$. Suppose that $\theta_i = \frac{1}{2} + \frac{i}{2m+1}$ for $i = 1, \ldots, m$ and $\theta_i = \frac{1}{2}$ for $i = m+1, \ldots, 2m$. For each of the coins we make n independent tosses.

(a) For each i find the most powerful test for testing $\theta = \frac{1}{2}$ against $\theta > \frac{1}{2}$ for level $\alpha \leq 0.1$ (finding the test of size exactly α may be hard).

(b) Consider now the multiple testing problem for all m coins. We proved that Bonferroni and the Holm procedures both control the FWER. Describe how Bonferroni will look in this case. We want FWER≤ 0.1.

(c) Take a look at the power of Bonferroni. Denote by $\mathbb{P}_i$ the distribution $\mathrm{Bern}(\theta_i)$ for $i = 1, \ldots, m$ (these are the coins for which the null does not hold). Try to find some sufficient conditions to bound the probability of the type II error for each of the false nulls by, say, 0.05.

(d) Provide some simulations to see the difference in power between the Bonferroni and Holm procedures. For fixed m consider $n = m, 10m, 100m$ to see how the answer depends on the ratio between n and m.

(some parts of this problem can be approached in various ways so there is no one correct solution)