Problem Set 6 **(Graded)** —*Due Tuesday, December 5, before class starts*

For the Exercise Sessions on Nov 21 and Nov 28

Last name	First name	SCIPER Nr	Points

Problem 1: Property Testing: Variance

A colleague claims to have implemented an algorithm which outputs i.i.d. samples distributed according to a discrete distribution P that has unit variance. Your task is to design a statistic to test whether this is indeed true.

Let Δ_k be the set of probability distributions on the alphabet $\mathcal{X} = \{1, \cdots, k\}$. Assume that $P \in \mathcal{P} \cup \mathcal{Q}$ with $\mathcal{P} := \{P \in \Delta_k : P \text{ has variance } 1\}$ and $\mathcal{Q} := \{P \in \Delta_k : P \text{ has variance} \in [0, 1-\epsilon] \cup [1+\epsilon, \infty)\}$, where $0 < \epsilon < 1$. You are given n samples $\{X_i\}_{i=1}^n$, where the X_i are independent copies sampled according to P.

Remark: For the following three questions we do not ask you to write down a proof (or explicit calculation) that your proposed solution works.

a) We say that an estimator $e : S^n \mapsto \Pi$ on a sample S^n of length n (ϵ, δ)-learns a parameter $p \in \Pi$ if for any $(\epsilon, \delta) \in (0, 1)^2$, given sufficiently many samples n, we have that $\mathbb{P}(\{|e(S^n) - p| > \epsilon\}) < \delta$. Give a brief explanation (one sentence, no calculations) why the empirical estimator of the second moment $\hat{\mu}_{X^2} := \frac{1}{n}\sum_{i=1}^n X_i^2$ can (ϵ, δ)-learn the second moment in our setting.

b) First, assume that a genie tells you that X has zero mean. Design a simple test statistic and give a threshold in order to check for the above mentioned unit variance property.

   Hint: Use the claim in a).

c) Now consider the more general case where X can have arbitrary mean. Again, design a simple test statistic and give a threshold.

   Hint: You can assume that $\hat{\mu}_X^2$ (ϵ, δ)-learns $\mathbb{E}[X]^2$.

Problem 2: MMSE Estimation

Consider the scenario where $p(x|d) = de^{-dx}$, for $x \geq 0$ (and zero otherwise), that is, the observed data x is distributed according to an exponential with mean $1/d$. Moreover, the desired variable d itself is also exponentially distributed, with mean $1/\mu$.

(a) Find the MMSE estimator of d given x, and calculate the corresponding mean-squared error incurred by this estimator.

(b) Find the MAP estimator of d given x.

Problem 3: Parameter Estimation and Fisher Information

The Fisher information $J(\Theta)$ for the family $f_\theta(x), \theta \in \mathbf{R}$ is defined by

$$J(\theta) = \mathbb{E}_\theta \left(\frac{\partial f_\theta(X)/\partial\theta}{f_\theta(X)} \right)^2 = \int \frac{(f_\theta^{'})^2}{f_\theta}$$

Find the Fisher information for the following families:

(a) $f_\theta(x) = N(0,\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}}$

(b) $f_\theta(x) = \theta e^{-\theta x}, x \geq 0$

(c) What is the Cramèr Rao lower bound on $\mathbb{E}_\theta(\hat{\theta}(X) - \theta)^2$, where $\hat{\theta}(X)$ is an unbiased estimator of θ for (a) and (b)?

Problem 4: Missing Data

We are given real-valued data with a single missing sample :

$$X_1, X_2, X_3, X_4, X_5, X_6, ?, X_8, X_9, \ldots \tag{1}$$

where we assume that the data is wide-sense stationary with autocorrelation function $R_X[k] = \alpha^{|k|}$, where $0 < \alpha < 1$. We would like to find a meaningful estimate for the missing sample X_7.

1. As a starting point, let us consider the estimate $\hat{X}_7 = wX_6$, where w is a real number. Find the value of w so as to minimize the mean-squared error $\mathbb{E}[(X_7 - \hat{X}_7)^2]$, and determine the incurred mean-squared error.
2. Now, consider the estimate $\hat{X}_7 = w_1 X_6 + w_2 X_8$. Again, find the values of w_1 and w_2 so as to minimize the mean-squared error $\mathbb{E}[(X_7 - \hat{X}_7)^2]$, and determine the incurred mean-squared error.

Problem 5: Tweedie's Formula

For the special case where $X = D + N$, where N is Gaussian noise of mean zero and variance σ^2, *Tweedie's formula* says that the conditional mean (that is, the MMSE estimator) can be expressed as

$$\mathbb{E}\left[D \middle| X = x\right] = x + \sigma^2 \ell'(x), \tag{2}$$

where

$$\ell'(x) = \frac{d}{dx} \log f_X(x), \tag{3}$$

where $f_X(x)$ denotes the marginal PDF of X. In this exercise, we derive this formula.

(a) Assume that $f_{X|D}(x|d) = e^{\alpha dx - \psi(d)} f_0(x)$ for some functions $\psi(d)$ and $f_0(x)$ and some constant α (such that $f_{X|D}(x|d)$ is a valid PDF for every value of d). Define

$$\lambda(x) = \log \frac{f_X(x)}{f_0(x)}, \tag{4}$$

where $f_X(x)$ is the marginal PDF of X, i.e., $f_X(x) = \int f_{X|D}(x|\delta) f_D(\delta) d\delta$. With this, establish that

$$\mathbb{E}\left[D \middle| X = x\right] = \frac{1}{\alpha} \frac{d}{dx} \lambda(x). \tag{5}$$

(b) Show that the case where $X = D + N$, where N is Gaussian noise of mean zero and variance σ^2, is indeed of the form required in Part *(a)* by finding the corresponding $\psi(d), f_0(x)$, and α. Show that in this case, we have

$$\frac{f_0'(x)}{f_0(x)} = -\frac{x}{\sigma^2}, \tag{6}$$

and use this fact in combination with Part *(a)* to establish Tweedie's formula.