# Homework 3: Mathematical Statistics (MATH-UA 234)

Due 10/06 at the beginning of class on Gradescope

**Problem 1.** *Let*

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \vec{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix}$$

*where $a_{i,j}$ are constants and $x_1, x_2, \dots, x_n \sim N(0,1)$ are independent and identically distributed standard normal random variables.*

*(a) What is the covariance matrix $\vec{\Sigma}$ for $\vec{x}$? (Recall the $(i,j)$-entry of $\vec{\Sigma}$ is $\mathrm{Co}\mathbb{V}[x_i, x_j]$.*

*(b) Show that $\mathbb{E}[\vec{x}^\top \vec{A}\vec{x}] = \mathrm{tr}(\vec{A}) := a_{1,1} + a_{2,2} + \cdots + a_{n,n}$. (hint: write out the expression for $\vec{x}^\top \vec{A}\vec{x}$ as a sum over the entires of $\vec{A}$)*

*(c) Suppose $\vec{A}$ is diagonal; i.e. $a_{i,j} = 0$ for all $i \neq j$. Compute the variance of $\vec{x}^\top \vec{A}\vec{x}$. You may use the fact that $x_i^2$ is a Chi-square random variable with one degree of freedom so that $\mathbb{V}[x_i^2] = 2$.*

*This is an example of "stochastic trace estimation" which is an important algorithmic tool in a number of recent algorithms.*

**Problem 2** (Wasserstein 5.5). *Suppose $X_1, X_2, \dots, X_n \sim \mathrm{Ber}(p)$ are independent and identically distributed. Let $Z_n = (X_1^2 + X_2^2 + \cdots + X_n^2)/n$. Prove that*

*(a) $Z_n$ converges in probability to the constant random variable $p$.*

*(b) $Z_n$ converges in quadratic mean to the constant random variable $p$.*

**Problem 3** (Wasserstein 6.1). *Suppose $X_1, X_2, \dots, X_n \sim \mathrm{Poisson}(\lambda)$ are independent and identically distributed and let $\hat{\lambda}_n = n^{-1}\sum_{i=1}^n X_i$ be our point estimator for $\lambda$. Find the bias $\mathrm{Bias}_n = \mathbb{E}[\hat{\lambda}_n] - \lambda$, standard error $se_n = \sqrt{\mathbb{V}[\hat{\lambda}_n]}$, and mean squared error $\mathrm{MSE}_n = \mathbb{E}[(\hat{\lambda}_n - \lambda)^2]$ of $\hat{\lambda}_n$.*

**Problem 4** (Wasserstein 6.2). *Suppose $X_1, X_2, \dots, X_n \sim \mathrm{Unif}(0,\theta)$ are independent and identically distributed and let $\hat{\theta}_n = \max\{X_1, X_2, \dots, X_n\}$ be our point estimator for $\theta$.*

*(a) Write down the distribution for $\hat{\theta}_n$ (hint: we've already done a similar problem)*

*(b) Find the bias $\mathrm{Bias}_n = \mathbb{E}[\hat{\theta}_n] - \theta$, standard error $se_n = \sqrt{\mathbb{V}[\hat{\theta}_n]}$, and mean squared error $\mathrm{MSE}_n = \mathbb{E}[(\hat{\theta}_n - \theta)^2]$ of $\hat{\theta}_n$.*

**Problem 5** (Wasserstein 6.3). *Suppose $X_1, X_2, \dots, X_n \sim \mathrm{Unif}(0,\theta)$ are independent and identically distributed and let $\hat{\theta}_n = 2\overline{X}_n$ be our point estimator for $\theta$. Find the bias $\mathrm{Bias}_n = \mathbb{E}[\hat{\theta}_n] - \theta$, standard error $se_n = \sqrt{\mathbb{V}[\hat{\theta}_n]}$, and mean squared error $\mathrm{MSE}_n = \mathbb{E}[(\hat{\theta}_n - \theta)^2]$ of $\hat{\theta}_n$.*

**Problem 6.** *Suppose $X_1, X_2, \dots, X_n \sim F_{\mu,\sigma^2}$ are independent and identically distributed samples from some distribution $F_{\mu,\sigma^2}$ with mean $\mu$ and variance $\sigma^2$.*

---

problems with a textbook reference are based on, but not identical to, the given reference

*Recall that*

$$\hat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \qquad \hat{T}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

*are both point estimators for the parameter* $\sigma^2$.

(a) *Compute* $\mathbb{E}[\hat{S}_n^2]$ *and* $\mathbb{E}[\hat{T}_n^2]$.

(b) *Which point estimator has smaller bias?*

(c) *Which point estimator has smaller standard error?*

**Problem 7.** *Describe of point estimation of a parameter which you noticed in a different part of* ***your life*** *(e.g. in other classes, on the subway, at the park, in the dorm, etc.).*