1. (a) Let Z_i be independent $N(0,1), i = 1, 2, \cdots, n$. Are $\overline{Z} = \frac{1}{n}\sum_{i=1}^n Z_i$ and $S_Z^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z})^2$ independent? Prove your claim.

   (b) Let $X_1, X_2, \cdots, X_n$ be independent identically distributed normal with mean θ and variance θ^2, where $\theta > 0$ is unknown. Let

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

   Are $\overline{X}$ and S^2 independent? Prove your claim. (Hint: you can directly use the result in the first part of this problem.)

   (c) Show that $(\overline{X}, S^2)$ is a sufficient statistic for θ, but it is not complete.

2. (a) Let $X_1, X_2, \cdots, X_n$ be exponentially distributed with density

   $$f(x) = \lambda \exp(-\lambda x), \quad x > 0.$$

   Let $c > 0$ be a constant and if $X_i < c$, we observe X_i, otherwise we observe c.

   $$S_n = \sum_{i=1}^n X_i I(X_i < c), \quad T_n = \sum_{i=1}^n I(X_i > c),$$

   where $I(A) = 1$ if event A occurs and $I(A) = 0$, otherwise. Write down the likelihood function of the observed values in terms of T_n and S_n.

   (b) Show the maximum likelihood estimator of λ is

   $$\widehat{\lambda}_n = \frac{n - T_n}{S_n + cT_n}.$$

1. (a) Let $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$ be an i.i.d. sample. Find the method of moments estimate $\hat{\lambda}_{MOM}$ and the maximum likelihood estimate $\hat{\lambda}_{MLE}$ of λ.

   (b) Is $\hat{\lambda}_{MLE}$ unbiased? Is it efficient?

   (c) Give an example of a distribution where the MOM estimate and the MLE are different.

2. (a) Prove that, for any (possibly correlated) collection of random variables $X_1, \ldots, X_k$,

$$\text{Var}\left(\sum_{i=1}^{k} X_i\right) \leq k \sum_{i=1}^{k} \text{Var}(X_i). \tag{1}$$

   (b) Construct an example with $k \geq 2$ where equality holds in (1).

1. (a) Consider an independent identically distributed sequence $X_1, X_2, \cdots, X_{n+1}$ taking values 0 or 1 with probability distribution

$$P\{X_i = 1\} = 1 - P\{X_i = 0\} = p.$$

Uniformly choose M *fragments* $F_1, F_2, \cdots, F_M$ of length 2 starting in the interval [1, n], that is, $F_i = (X_{j_i}, X_{j_i+1})$ for some $1 \le j_i \le n$. Let $\mathbf{W} = (1, 1)$.

- Let $N_{\mathbf{W}}$ be the number of times the word $\mathbf{W}$ occurs among the M fragments. Calculate $E(N_{\mathbf{W}})$.
- Calculate the probability $P(F_1 = \mathbf{W}, F_2 = \mathbf{W})$.
- Calculate $\mathrm{Var}(N_{\mathbf{W}})$.

(Note: Due to time constraints, you can ignore the boundary effect.)

2. Let T and C be independent Geometric random variables with success probability of r and s, respectively. That is

$$P[T = j] = r(1-r)^{j-1}; j = 1, 2, \cdots,$$
$$P[C = j] = s(1-s)^{j-1}; j = 1, 2, \cdots,$$

Let $X = (\min(T, C), I(T \le C))$. Denote $X_1 = \min(T, C), X_2 = I(T \le C)$, where $I(\cdot)$ is the indicator function.

(a) What is the joint distribution of X?

(b) Calculate $EX = (EX_1, EX_2)$ and the covariance matrix of $X = (X_1, X_2)$.

(c) Let $T_1, T_2, \cdots, T_n$ be a random sample from T, and $C_1, C_2, \cdots, C_n$ be a random sample from C. Define

$$S_1 = \sum_{i=1}^{n} \min(T_i, C_i)$$

$$S_2 = \sum_{i=1}^{n} I(T_i \le C_i).$$

What is the maximum likelihood estimate $(\hat{r}, \hat{s})$ of (r, s), in terms of S_1 and S_2?

Fall 2013 Math 541a Exam

1. For $p \in (0,1)$ unknown, let $X_0, X_1, \ldots$ be independent identically distributed random variables taking values in $\{0,1\}$ with distribution

$$P(X_i = 1) = 1 - P(X_i = 0) = p,$$

and suppose that

$$T_n = \sum_{i=0}^{n-1} I(X_i = 1, X_{i+1} = 1). \tag{1}$$

is observed.

   (a) Calculate the mean and variance of T_n.

   (b) Find a consistent method of moments $\widehat{p}_n = g_n(T_n)$ estimator for the unknown p as a function g_n of T_n that may depend on n, and prove that your estimate is consistent for p.

   (c) Show that T_n is not the sum of independent, identically distributed random variables. Nevertheless, determine the non-trivial limiting distribution of $\widehat{p}_n$, after an appropriate centering and scaling, as if (1) was the sum of i.i.d. variables and has the same mean and variance as the one computed in part (a).

   (d) Explain why you would, or would not, expect $\widehat{p}_n$ to have the same limiting distribution as the one determined in part (c).

2. Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with density given by

$$f_\beta(x) = \frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp(-x/\beta), \text{ for } x > 0$$

where $\alpha > 0$, and is known. Suppose it is desired to estimate β^3.

   (a) Find the Cramer-Rao lower bound for the variance of an unbiased estimator of β^3.

   (b) Find a complete and sufficient statistic for β. Then, compute its k^{th} moment, where k is an positive integer.

   (c) If a UMVUE (uniform minimum variance unbiased estimator) exists, find its variance and compare it to the bound in part (a).

Spring 2014 Math 541a Exam

1. For known values $x_{i,1}, x_{i,2}, i = 1, \ldots, n$ let

$$Z_i = \beta_1 x_{i,1} + \epsilon_i$$

and

$$Y_i = \beta_1 x_{i,1} + \beta_2 x_{i,2} + \epsilon_i \quad i = 1, \ldots, n,$$

where $\epsilon_i, i = 1, 2, \ldots, n$ are independent normal random variables with mean 0 and variance 1.

   (a) Given the data $\mathbf{Z} = (Z_1, \ldots, Z_n)$ compute the maximum likelihood estimate of β_1 and show that it achieves the Cramer-Rao lower bound. Throughout this part and the following, make explicit any non-degeneracy assumptions that may need to be made.

   (b) Based on $\mathbf{Y} = (Y_1, \ldots, Y_n)$, compute the Cramer-Rao lower bound for the estimation of (β_1, β_2), and in particular compute a variance lower bound for the estimation of β_1 in the presence of the unknown β_2.

   (c) Compare the variance lower bound in (a), which is the same as the one for the model for Y_i where β_2 is known to be equal to zero, to the one in (b), where β_2 is unknown, and show the latter one is always at least as large as the former.

2. Suppose we observe the pair (X, Y) where X has a Poisson(λ) distribution and Y has a Bernoulli$(\lambda/(1+\lambda))$ distribution, that is,

$$P_\lambda(X = j) = \frac{\lambda^j e^{-\lambda}}{j!}, j = 0, 1, 2, \ldots$$

and

$$P_\lambda(Y = 1) = \frac{\lambda}{1+\lambda} = 1 - P_\lambda(Y = 0),$$

with X and Y independent, and $\lambda \in (0, \infty)$ unknown.

   (a) Find a one-dimensional sufficient statistic for λ based on (X, Y).

   (b) Is there a UMVUE (uniform minimum variance unbiased estimator) of λ? If so, find it.

   (c) Is there a UMVUE of $\lambda/(1+\lambda)$? If so, find it.

Fall 2014 Math 541a Exam

1. Let p, q be values in $[0,1]$ and $\alpha \in (0,1]$. Assume α and q known, and that p is an unknown parameter we would like to estimate. A coin is tossed n times, resulting in the sequence of zero one valued random variables $X_1, \ldots, X_n$. At each toss, independently of all other tosses, the coin has probability p of success with probability α, and probability q of success with probability $1-\alpha$.

   (a) Write out the probability function of the observed sequence, and compute the maximum likelihood estimate $\widehat{p}$ of p, when p is considered a parameter over all of $\mathbb{R}$. Verify that when $\alpha = 1$ one recovers the standard estimator of the unknown probability.

   (b) Show $\widehat{p}$ is unbiased, and calculate its variance.

   (c) Calculate the the information bound for p, and determine if it is achieved by $\widehat{p}$.

   (d) If one of the other parameters is unknown, can p still be estimated consistently?

2. Let $\mathbf{X} \in \mathbb{R}^n$ be distributed according the density or mass function $p(\mathbf{x}; \theta)$ for $\theta \in \Theta \subset \mathbb{R}^d$.

   (a) State the definition for $T(\mathbf{X})$ to be sufficient for θ.

   (b) Prove that if the (discrete) mass functions $p(\mathbf{x}; \theta)$ can be factored as $h(\mathbf{x})g(T(\mathbf{x}), \theta)$ for some functions h and g, then $T(\mathbf{X})$ is sufficient for θ.

   (c) Let $X_1, \ldots, X_n$ be independent with the Cauchy distribution $\mathcal{C}(\theta), \theta \in \mathbb{R}$ given by

   $$p(x; \theta) = \frac{1}{\pi(1 + (x-\theta)^2)}.$$

   Prove that the unordered sample $S = \{X_1, \ldots, X_n\}$ can be determined from any $T(\mathbf{X})$ sufficient for θ.(Hint: Produce a polynomial from which S can be determined).

1. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample from the uniform distribution on a disc $X^2 + Y^2 \leq \theta^2$, where $\theta > 0$ is unknown. That is, the probability density function of (X, Y) is

$$f_{(X,Y)}(x, y; \theta) = \frac{1}{\pi\theta^2} 1_{[0,\theta]}(\sqrt{x^2 + y^2})$$

   (a) Find a complete sufficient statistic of θ and its distribution.

   (b) Find the UMVU estimator of θ.

   (c) Find the maximum likelihood estimator of θ.

2. Let $Y_1, \ldots, Y_n$ be independent with $Y_i \sim N(\alpha x_i + \beta \log x_i, \sigma^2)$, where $x_1, \ldots, x_n$ are given positive constants, not all equal, and α, β, σ are unknown parameters.

   (a) Prove that the MLE of β is

$$\widehat{\beta} = \frac{S_{ly} S_{x^2} - S_{lx} S_{xy}}{S_{x^2} S_{l^2} - S_{lx}^2}, \tag{1}$$

   where

$$S_{ly} = \sum_i (\log x_i) Y_i, \quad S_{x^2} = \sum_i x_i^2, \quad S_{l^2} = \sum_i (\log x_i)^2, \quad \text{etc.}$$

   (b) Find the distribution of $\widehat{\beta}$, including giving any parameter values for this distribution. Is $\widehat{\beta}$ unbiased for β? Justify your answers.

   (c) Suppose now that you may choose the values of the x_i, but each one must be either 1 or 10. How many of the n observations should you choose to make at $x_i = 1$ in order to minimize the variance of the resulting $\widehat{\beta}$? You can assume that n is a fixed multiple of 11.

1. Let $f(x;\theta), \theta \in \Theta \subset \mathbb{R}^d$ be a family of density functions, and consider the sequence of moments

$$m_{\theta,k} = E_\theta[X^k] \quad \text{for } k = 1, 2, \ldots$$

when these exist, and where E_θ indicates expectation with respect to $f(x;\theta)$. Let $X_1, \ldots, X_n$ be an i.i.d. sample from $f(x;\theta)$ and

$$\widehat{m_k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k,$$

the sample k^{th} moment. It is desired to estimate some function ψ of θ, which may equal θ itself. We say $\widehat{\psi}$ is a moment estimator of ψ of order k when $m_{\theta,1}, \ldots, m_{\theta,k}$ exist and for some function g,

$$\widehat{\psi} = g(\widehat{m_1}, \ldots, \widehat{m_k}) \quad \text{when} \quad \psi = g(m_{\theta,1}, \ldots, m_{\theta,k}). \tag{1}$$

a. Find a moment estimator of order 2 for the variance of the distribution $f(x;\theta)$, assuming it exists.

b. Let $f_1(x), \ldots, f_d(x)$ be known density functions and

$$f(x;\theta) = \sum_{i=1}^{d} \theta_i f_i(x)$$

where $\theta \in \Theta$ with

$$\Theta = \left\{ \theta \in \mathbb{R}^d; \sum_{i=1}^{d} \theta_i = 1, \theta_i \geq 0 \right\}.$$

Find a moment estimator for θ, and write what moment assumptions are needed. It is not necessary to write the function g in (1) explicitly, but explain how to calculate the moment estimator for θ.

c. Prove that if $\widehat{\psi}$ is a moment estimator of ψ of order k of the form (1) where g is continuous then $\widehat{\psi}$ is consistent for ψ.

d. Assuming sufficient smoothness on the function g, identify the asymptotic distribution of $\widehat{\psi}$, after properly centering and scaling to assure the limit is non trivial. (No need to prove the distributional convergence.)

2. (a) Let $a_1, \ldots, a_n$ be distinct real numbers and let X have the discrete uniform distribution on the set $\{a_1, \ldots, a_n\}$, i.e.,

$$P(X = a_i) = 1/n \quad \text{for all} \quad i = 1, \ldots, n.$$

Find

i. $E(X)$

ii. $\mathrm{Var}(X)$

iii. a median $\mathrm{med}(X)$ of X, i.e., a number minimizing $E|X - \mathrm{med}(X)|$, which is not necessarily unique.

(b) A scientist has a data set $x_1, \ldots, x_{10}$ of distinct real numbers such that $\sum_{i=1}^{10}(x_i - \overline{x})^2 = 110$, where $\overline{x} = (x_1 + \ldots + x_{10})/10$ is the sample mean.

i. The scientist claims that one of the data points exceeds the sample mean by 11 or more. Could this be true? Why or why not?

ii. You further learn that all the data points are positive and $\overline{x} = 1.4$. The scientist claims that two of the data points exceed the sample mean by 7 or more. Could this be true? Why or why not?

iii. Finally the scientist claims that 5 is a sample median m of the data set. Could this be true? Why or why not? *Hint:* Use Jensen's inequality to bound $|\overline{x} - m|$.

1. Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability distributions. A statistic V is called *ancillary* for θ if its distribution does not depend on θ.

   (a) Let $X_1, \ldots, X_n$ have normal distribution $N(\mu, 1)$. Show that $V = X_1 - \bar{X}$ is an ancillary statistic for μ.

   (b) Prove that if T is a complete sufficient statistic for the family $\{P_\theta, \theta \in \Theta\}$, then any ancillary statistic V is independent of T. (This is a theorem due to Basu).
   (Hint: show that for any (measurable) set A,

   $$P_\theta(V \in A | T = t) = P(V \in A | T = t) = P(V \in A),$$

   and derive the conclusion).

2. With $\theta > 0$ unknown, let a sample consist of $X_1, \ldots, X_n$, independent observations with distribution

   $$F(y;\theta) = 1 - \sqrt{1 - \frac{y}{\theta}}, \quad 0 < y < \theta.$$

   (a) Prove that the maximum likelihood estimate of θ based on the sample is the maximum order statistic

   $$X_{(n)} = \max_{1 \le i \le n} X_i.$$

   (b) Determine a sequence of positive numbers a_n and a non-trivial distribution for a randon variable X such that

   $$a_n(\theta - X_{(n)}) \to_d X.$$

   (c) Compare the rate a_n in part b) to the rate of a parametric estimation problem whose regularity would allow the application of the Cramer Rao bound. Comment.

Spring 2015 Math 541a Exam

1. For $n \geq 2$ let $X_1, \cdots, X_n$ be independent samples from P_θ, the uniform distribution $U(\theta, \theta+1), \theta \in \mathbb{R}$. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics of the sample.

   (a) Show that $(X_{(1)}, X_{(n)})$ is a sufficient statistic for θ.

   (b) Is $(X_{(1)}, X_{(n)})$ complete? Prove your claim.

   (c) Find a_n and $b(\theta)$ such that $a_n(b(\theta) - X_{(n)}) \to Z$ in distribution, where Z has an exponential distribution with density $f(x) = e^{-x}, x > 0$.

   (d) What is the maximum likelihood estimate of θ given the sample?

2. Let X and Y be independent random variables with $X \sim \text{exponential}(\lambda)$ and $Y \sim \text{exponential}(\mu)$, where the exponential$(\nu)$ density is given by

$$f(x; \nu) = \frac{1}{\nu} \exp\left(-\frac{x}{\nu}\right).$$

   Let

$$Z = \min\{X, Y\},$$

   and

$$W = 1 \text{ if } Z = X, \text{ and } W = 0 \text{ otherwise.}$$

   (a) Find the joint distribution of Z and W.

   (b) Prove that Z and W are independent.

   (c) Suppose that (X, Y) are not observable. Instead, with $n \geq 2$, we observe $(Z_1, W_1), \cdots, (Z_n, W_n)$, independent samples distributed as (Z, W). Write down the likelihood function in terms of the sample averages $(\overline{Z}, \overline{W})$, and find the maximum likelihood estimate $(\widehat{\mu}_n, \widehat{\lambda}_n)$ of (μ, λ).

   (d) Determining whether $1/\hat{\lambda}_n$ is unbiased for $1/\lambda$, and if not, construct an estimator that is.

1. Let $X_1, \ldots, X_n$ be a random sample from the normal distribution $N(\mu, 1)$. Let $u \in \mathbb{R}$ be a given threshold value, and assume that we want to estimate the probability $p = p(\mu) = P_\mu(X_1 \le u)$.

   (a) Find an unbiased estimator of p.

   (b) Letting $\overline{X} = n^{-1} \sum_{i=1}^n X_i$ denote the sample mean, show that the joint distribution of $\overline{X}$ and $X_1 - \overline{X}$ is bivariate normal and find the parameters of this distribution. Use your answer to demonstrate that $\overline{X}$ and $X_1 - \overline{X}$ are independent.

   (c) Use the estimator from part (1a), along with the Rao-Blackwell theorem and part (1b), to find the uniform minimal variance unbiased estimator (UMVUE) for p.

2. For all $k = 0, 1, \ldots$, we have that

$$\int_{-\infty}^{\infty} x^k e^{-x^4/12} dx = 2^{\frac{k-3}{2}} 3^{\frac{k+1}{4}} \left((-1)^k + 1\right) \Gamma\left(\frac{k+1}{4}\right).$$

   (a) Determine c_1 such that

$$p(x) = c_1 e^{-x^4/12} \tag{1}$$

   is a density function.

   In the following we consider the location model

$$p(x;\theta) = p(x-\theta), \quad \theta \in (-\infty, \infty)$$

   where $p(x)$ is as in (1). Assume a sample $X_1, \ldots, X_n$ of independent random variables with distribution $p(x;\theta)$ has been observed.

   (b) Prove that the maximizer of the likehood function is unique and can be found by setting the derivative of the log likelihood to zero.

   (c) Determine the maximum likelihood estimate $\widehat{\theta}$ of θ based on the sample in a form as explicit as you can in terms of the sample moments

$$m_k = \frac{1}{n} \sum_{i=1}^n X_i^k.$$

   (d) Determine the information $I_{\mathbf{X}}(\theta)$ for θ in the sample $X_1, \ldots, X_n$, and the non-trivial limiting distribution of $\widehat{\theta}$, when properly scaled and centered. You may assume regularity conditions hold without explicitly noting them.

1. For known $\alpha > 0$ and unknown $\theta \in \mathbb{R}$, we observe the sequence $X_1, \ldots, X_n$ of independent variables with distribution function

$$F(x;\theta) = \begin{cases} 0 & x \le \theta \\ (x-\theta)^\alpha & x \in (\theta, \theta+1] \\ 1 & x > \theta + 1, \end{cases}$$

and consider the estimator of θ given by

$$\widehat{\theta} = \min\{X_1, \ldots, X_n\}.$$

   a. Determine the distribution function of $\widehat{\theta}$.

   b. Determine a non-trivial distribution for a random variable Y and an increasing sequence a_n of real numbers such that

$$a_n(\widehat{\theta} - \theta) \to_d Y.$$

   c. What rates of (distributional) convergence to θ are possible for the estimator $\widehat{\theta}$ as α ranges over $(0, \infty)$?

2. Let $X_1, \ldots, X_n$ be i.i.d. with the $N(\theta, \theta)$ distribution, for some $\theta > 0$.

   (a) Find the MLE of θ.

   (b) Show that the MLE is a consistent estimator of θ.

   (c) Assume that $n = 1$.

      i. Show that $T(X) = |X|$ is a sufficient statistic for θ.

      ii. Note that $\hat{\theta}(X) = X$ is an unbiased estimator of θ, but is not a function of $T(X)$. Hence, it can be improved; apply the Rao-Blackwell theorem to find another estimator θ^* that is a function of $T(X)$ and is unbiased for θ.

1. Let $U_1, \ldots, U_n$ be iid with the uniform distribution $\mathcal{U}[\beta, \beta+1]$ for $\beta \in \mathbb{R}$ and $n \geq 2$, and let $U_{(1)} < \cdots < U_{(n)}$ be the order statistics of the sample.

   a. Show that $(U_{(1)}, U_{(n)})$ is sufficient but not complete.

   b. Show that $\overline{U}_n - 1/2$ is not a uniformly minimum variance estimator of β, where $\overline{U}_n$ is the average of $U_1, \ldots, U_n$.

2. Fix $n \geq 2$.

   (a) Let $y_{(1)} \leq \ldots \leq y_{(n)}$ be some ordered real numbers and define

   $$x_k = \sum_{i>k} y_{(i)} - (n-k)y_{(k)} \quad \text{for} \quad k = 1, \ldots, n, \tag{1}$$

   where an empty sum denotes 0 by convention. Show that x_k is a non-increasing sequence and $x_n = 0$.

   (b) Let $Y_1, \ldots, Y_n$ be independent with $Y_i \sim N(\theta_i, \sigma^2)$ for unknown $\sigma^2 > 0$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ in the $(n-1)$-simplex

   $$\theta_1 + \ldots + \theta_n = 1 \quad \text{and} \quad \theta_i \geq 0 \quad \text{for all } i. \tag{2}$$

   Based on the data $Y_1 = y_1, \ldots, Y_n = y_n$:

   i. find the MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, subject to the constraints (2), by minimizing

   $$\frac{1}{2}\sum_i (y_i - \theta_i)^2 + \lambda \left(\sum_i \theta_i - 1 \right) \tag{3}$$

   over $\theta_i \geq 0$ and λ (Hint: Use part 2a, and consider the smallest $k \in \{1, \ldots, n\}$ such that $x_k < 1$, and $\lambda \in [y_{(k-1)}, y_{(k)})$.)

   ii. find the MLE of σ^2. You can state your answer in terms of $\widehat{\boldsymbol{\theta}}$ even if you don't solve part 2(b)i.

1. Let $\{\mathcal{P}_\theta,\ \theta \in \Theta\}$ be a family of probability distributions, and $X \sim \mathcal{P}_\theta$ for some $\theta \in \Theta$. Let $\mathcal{E}$ be the set of all unbiased estimators of 0 with finite variances, and let T be an unbiased estimator of θ with $\mathrm{Var}(T) < \infty$.

   (a) Assume that T is the uniformly minimum-variance unbiased estimator (UMVUE) of θ. Show that $\mathbb{E}_\theta[T(X)U(X)] = 0$ for any $U \in \mathcal{E}$ and $\theta \in \Theta$.

   (b) Assume that $\mathbb{E}_\theta[T(X)U(X)] = 0$ for any $U \in \mathcal{E}$ and $\theta \in \Theta$. Show that T is the UMVUE.

   (c) Assume now that the family of distributions is parametrized by k parameters $\theta_1, \ldots, \theta_k$, and that T_j is the UMVUE of θ_j, $1 \le j \le k$. Use (a) and (b) to show that $\sum_{j=1}^k \alpha_j T_j$ is the UMVUE of $\sum_{j=1}^k \alpha_j \theta_j$ for any $\alpha_1, \ldots, \alpha_k$.

2. Let $X_1, \ldots, X_n$ be independent with distribution equal to that of X, a random variable with mean $\mu = E[X]$, variance σ^2, and finite sixth moment $E[X^6] < \infty$. Define the skewness parameter

$$\tau = \frac{E[X-\mu]^3}{\sigma^3}.$$

   Assume μ is known, and consider the estimation of τ by

$$\widehat{\tau} = \frac{\widehat{m_3}}{\widehat{\sigma^3}} \quad \text{where} \quad \widehat{m_3} = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^3 \quad \text{and} \quad \widehat{\sigma^3} = \left(\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2\right)^{3/2}.$$

   (a) Show that the parameter τ is invariant with respect to translation and positive scaling of X.

   (b) Determine, and justify, whether or not $\widehat{\tau}$ is UMVU when the X distribution is normal.

   (c) Assuming the X distribution is symmetric about its mean, find the mean and variance of $(X_i-\mu)^2$ and $(X_i-\mu)^3$, and their covariance. Apply the multivariate central limit theorem to yield the limiting distribution of the properly scaled and centered (bivariate) sum

$$S_n = \frac{1}{n}\sum_{i=1}^n ((X_i-\mu)^2, (X_i-\mu)^3).$$

   (d) Recall that when a sequence of vectors $\mathbf{Y}_n$ in $\mathbb{R}^d$ converges in distribution to $\mathbf{Y}_0$, that is,

$$\sqrt{n}(\mathbf{Y}_n - E\mathbf{Y}_n) \to_d \mathbf{Y}_0,$$

   and $g : \mathbb{R}^d \to \mathbb{R}^r$ is a nice function, the multivariate delta method yields that

$$\sqrt{n}(g(\mathbf{Y}_n) - g(E\mathbf{Y}_n)) \to_d \dot{g}^T(E(\mathbf{Y}_0))\mathbf{Y}_0,$$

   where $g = (g_1, \ldots, g_r)$ and $\dot{g}$ is the matrix whose columns are the gradients of $g_1, \ldots, g_r$. That is, $\dot{g} = (\nabla g_1, \ldots, \nabla g_r)$. Use this result to derive a properly scaled limiting distribution for $\widehat{\tau}$.

1. Let $X_1, \ldots, X_n$ be independent with distribution $\mathcal{U}[\theta-1, \theta]$, the uniform distribution on the interval $[\theta - 1, \theta]$, for some unknown $\theta \in \mathbb{R}$. Let $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ denote the order statistics.

   (a) Find, with explanation, an unbiased estimator $\widehat{\theta}$, which is a function only of $X_{(1)}$.

   (b) By finding, and proving the superiority (in mean square error) of another unbiased estimator $\widetilde{\theta}$ that dominates $\widehat{\theta}$ in (a), show that $\widehat{\theta}$ is not UMVU. (Hint: To avoid some cumbersome computation, you may use the fact that $X_{(1)}$ and $X_{(n)}$ are not linearly dependent, without proving it.)

   (c) Consider the following statement: Since $\widehat{\theta}$ is not UMVU, $X_{(1)}$ must either be not complete or not sufficient.

      i. From what theorem does the statement follow?

      ii. Give a direct proof that $X_{(1)}$ fails to be (your choice) either complete or sufficient.

2. (a) Let $\{\mathcal{P}_\theta,\ \theta \in \Theta\}$ be a family of probability distributions, and $X \sim \mathcal{P}_\theta$ for some $\theta \in \Theta$. Prove that if $T(X), T'(X)$ are both uniformly minimum-variance unbiased estimators (UMVUEs) of θ, then $T = T'$, $\mathcal{P}_\theta$-almost surely for every $\theta \in \Theta$.

   Hint: Letting $v = \mathrm{Var}(T) = \mathrm{Var}(T')$ denote the minimal variance and $\overline{T} = (T + T')/2$, compare $\mathrm{Var}(\overline{T})$ to v and use the Cauchy-Schwarz inequality $\mathrm{Cov}(T, T') \leq [\mathrm{Var}(T)\mathrm{Var}(T')]^{1/2}$ to show that $\mathrm{Cov}(T, T') = v$. Use this to show that $\mathrm{Var}(T - T') = 0$.

   (b) Let two independent sequences $X_1, \ldots, X_m$ i.i.d. $N(\mu, \gamma\sigma^2)$ and $Y_1, \ldots, Y_n$ i.i.d. $N(\mu, \sigma^2)$ be observed. Here, the X_i are independent of the Y_j, $\gamma > 0$ is known, and μ and $\sigma^2 > 0$ are unknown. Find the UMVUE of μ, and prove that it is UMVUE.

   (c) Let two independent sequences $X_1, \ldots, X_m$ be i.i.d. $N(\mu, \sigma_x^2)$ and $Y_1, \ldots, Y_n$ i.i.d. $N(\mu, \sigma_y^2)$ be observed. Here, the X_i are independent of the Y_j, μ is unknown, and $\sigma_x^2 > 0$ and $\sigma_y^2 > 0$ are unknown and not assumed to be equal. Show that the UMVUE of μ does not exist.

   Hint: Let $\gamma = \sigma_x^2/\sigma_y^2$ and apply part (b) to say what the UMVUE is. Then reverse the roles of the X's and Y's, repeat, and finally apply part (a) for a contradiction.

1. Suppose for some unknown $\theta_0 \in \mathbb{R}$ we observe pairs (x_i, y_i), with x_i non-random and

$$y_i = \theta_0 x_i + \epsilon_i \quad i = 1, 2, \ldots, n$$

where $\epsilon_1, \ldots, \epsilon_n$ are random errors. We estimate θ by least squares, that is, by the value $\widehat{\theta}_n$ minimizing

$$J_n(\theta) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - \theta x_i)^2 .$$

   (a) Compute $\widehat{\theta}_n$. Are any conditions needed for this estimate to exist?

   (b) Find conditions on the distribution of y_i, as minimal as you can, for which $\widehat{\theta}_n$ will be unbiased for θ_0.

   (c) Under the assumptions in (a) and (b), and that the ϵ_i errors each have variance σ^2, find conditions, as simple as possible, under which $\widehat{\theta}_n$ will be consistent.

   (d) Without specifying the precise conditions, find the limiting asymptotic distribution of $\widehat{\theta}_n$ once it has been properly scaled and centered.

2. (a) Let $Z_1, \ldots, Z_n$ be i.i.d. $N(\mu, \sigma^2)$ random variables, for some $\sigma > 0$ and μ. Define $X_i = e^{Z_i}$, $i = 1, \ldots, n$. Find the mean, variance, and median of the distribution of X_1 in terms of μ and σ.

   (b) Let $M_n = (\prod_{i=1}^{n} X_i)^{1/n}$ be the geometric mean of the X_i. In terms of μ and σ, find

      i. the c.d.f. of M_n

      ii. the mean, variance, and median of M_n (*hint:* use part 2a)

      iii. $\lim_{n\to\infty} M_n$

   (c) Show that, for any $n \geq 1$,

$$\lim_{n' \to \infty} M_{n'} \leq EM_n \leq EX_1.$$

1. For $\lambda > 0$, consider observing a single observation $X \sim \mathcal{P}(\lambda)$ from the Poisson distribution, satisfying

$$P_\lambda(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \quad k = 0, 1, 2 \ldots. \tag{1}$$

(For instance, X might be the number of arrivals in the time interval $[0, 1]$, when arrivals follow a Poisson process with rate λ.)

   (a) Based on X, find the UMVU of $\phi(\lambda) = e^{-3\lambda}$ (the probability that there are no arrivals in the interval $[1, 4]$). Hint: Find a function g that satisfies

$$E_\lambda[g(X)] = e^{-3\lambda},$$

   using (1) and the infinite series representation of the exponential function. Does the UMVU exist uniquely? Compute the variance of the resulting estimator.

   (b) Compute the value of the estimator for some small values of X, and comment on any peculiarities you observe.

2. Let $X_1, \ldots, X_n$ be a sample from the $\theta = \theta_0$ density from the family $\{p(x; \theta), \theta \in \Theta\}$ of positive densities, where $\Theta = \{\theta_0, \theta_1, \ldots, \theta_d\}$ is a finite set, with the distributions corresponding to differing parameters being unequal.

   (a) Define the scaled log likelihood function

$$\ell_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log p(X_i; \theta).$$

   Making any necessary first moment assumptions, use the law of large numbers to identify the limit for all $\theta \in \Theta_0$, and prove convergence to it; make sure to state what type of convergence is claimed to hold.

   (b) For p and q any density functions, recall the Kullback-Leibler divergence

$$D(P||Q) = E_p\left[\log\frac{p(Y)}{q(Y)}\right],$$

   where E_p means we take expectation with respect to density p. Use Jensens's inequality to show that $D(P||Q) \geq 0$. Is it the case that the inequality will be strict when the distributions P and Q are not equal?

   (c) Define the maximum likelihood estimate as

$$\widehat{\theta}_n = \operatorname{argmax} \ell_n(\theta),$$

   that is, as the θ value that achieves the maximum log likelihood for parameters in Θ; in the case of a tie, the parameter θ_j with the highest index j shall be chosen. Prove that $\widehat{\theta}_n$ converges to the true value θ_0. If convergence occurs in multiple ways, prove the strongest statement you can.

1. Let $X_1, \dots, X_n$ be i.i.d. observations with normal distribution $N(\mu, 1)$, where it is known that $\mu \geq 0$.

   1.1. Find the maximum likelihood estimator $\widehat{\mu}$ of μ.

   1.2. Derive the asymptotic distribution, as $n \to \infty$ of $\sqrt{n}(\widehat{\mu} - \mu)$, when the true mean satisfies $\mu > 0$; please provide a direct proof not relying on any general result about MLEs.

   1.3. Derive the asymptotic distribution, as $n \to \infty$, of $\sqrt{n}(\widehat{\mu} - \mu)$, that is, the limit of the probability $P(\sqrt{n}(\widehat{\mu} - \mu) \leq x)$ for all values of $x \in \mathbb{R}$, when the true mean μ is equal to 0.

   1.4. Why would typical general results about the asymptotic distribution of maximum likelihood estimators not apply in part (3)?

2. 2.1. Show that if W is UMVU for a parameter $\theta \in \mathbb{R}$, then for any U that is an unbiased estimator of 0 with finite variance, that is, $E_\theta[U] = 0$ and $E_\theta[U^2] < \infty$ for all $\theta \in \mathbb{R}$, we have $E_\theta[WU] = 0$ for all $\theta \in \mathbb{R}$. (Hint: Construct a family of unbiased estimators of θ using W and U.)

   2.2. Let X be a sample from the uniform distribution on the interval $[\theta - 1/2, \theta + 1/2]$, where $\theta \in \mathbb{R}$ is unknown. Find a continuous function $u(\cdot)$ such that $U = u(X)$ is an unbiased estimator of zero and $P_\theta(U \neq 0) = 1$. (Hint: consider periodic functions.)

   2.3. Let $g : \mathbb{R} \to \mathbb{R}$ be a nonconstant differentiable function of $\theta \in \mathbb{R}$. Show that there does not exist a UMVU of $g(\theta)$ of the form $w(X)$, a function of the observation X that is uniformly distributed over the interval $[\theta - 1/2, \theta + 1/2]$. You may assume you are given an unbiased estimator U of zero that satisfies the conditions in part (2), and that $W = w(X)$ with $w(x)$ continuous in x. (Hint: Use part (1) and differentiate the equalities $E_\theta[U] = 0$ and $E_\theta[WU] = 0$ with respect to $\theta \in \mathbb{R}$.)

1. For $\beta \geq 1$ and $\gamma > 0$, let X have the distribution of Y/γ, where Y has cumulative distribution function $F_Y(y) = 1 - (1-y)^\beta$ for $y \in [0,1]$. Let a sample $X_1, \ldots, X_n$ consist of independent random variables having the cumulative distribution function F_X of X.

   (a) Find the density function $p(x;\beta,\gamma)$ of X, and the density $p_n(\boldsymbol{x}_n, \beta, \gamma)$ of the full sample $\boldsymbol{X}_n = (X_1, \ldots, X_n)$.

   (b) Prove that for any fixed $\beta > 0$ and sample $\boldsymbol{X}_n$ there exists a solution $\widehat{\gamma}$ to the likelihood equation $U_n(\gamma, \boldsymbol{X}_n) = 0$ where

   $$U_n(\gamma, \boldsymbol{X}_n) = \frac{\partial}{\partial \gamma} \log p_n(\boldsymbol{X}_n, \beta, \gamma).$$

   Comment on how a solution may be found, the form that solution takes and any other notable features that may cause difficulties in solving this equation

   (c) With $X_{(n)} = \max\{X_1, \ldots, X_n\}$, find the limiting value of

   $$\lim_{n\to\infty} P(n^{1/\beta}(1/\gamma - X_{(n)}) \leq u)$$

   and comment on the feasibility of using a function of $X_{(n)}$ to form an estimate of γ, and what advantages or disadvantates it has over the MLE in part [b] above, and in particular as regards their rates of convergence. (You may wish to use that the cumulative distribution function of $X_{(n)}$ is F_X^n.)

2. Let $Y_1, Y_2, \ldots$ be random variables such that $\sqrt{n}Y_n$ converges in probability as $n \to \infty$ to a mean zero Gaussian random variable with variance 1.

   (a) Let $f(x) := (e^x + e^{-x})/2 = \cosh(x)$ for all $x \in \mathbb{R}$. Prove that

   $$n(f(Y_n) - f(0))$$

   converges in distribution as $n \to \infty$, and describe the limiting distribution explicitly. (Hint: if $f'(0)$ were nonzero, then the delta method implies that $\sqrt{n}(f(Y_n) - f(0))$ converges in distribution as $n \to \infty$, but since $f'(0) = 0$, another version of the delta method needs to be used.)

   (b) Does

   $$nY_n^3$$

   converges in distribution as $n \to \infty$? If so, what does it converge to? If not, justify your reasoning, perhaps with a counterexample.

   (c) Determine the set of all real numbers $k \in \mathbb{R}$ such that

   $$n^k Y_n^3$$

   converges in distribution as $n \to \infty$.

1. Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R},\ \sigma^2 > 0$, with both parameters unknown.

   (a) Find the complete sufficient statistics for (μ, σ^2), and the uniform minimum variance unbiased estimator $s_{0,n}^2$ of σ^2.

   (b) Find the maximum likelihood estimator s_n^2 of σ^2. Is it biased?

   (c) Which of the two estimators $s_{0,n}^2,\ s_n^2$ has smaller mean squared error? Recall that the MSE of an estimator $\widehat{\theta}$ of parameter θ is defined as $\mathrm{MSE}(\widehat{\theta}) = \mathbb{E}(\widehat{\theta} - \theta)^2$. Based on your answer, explain which estimator you would prefer and why.
   (hint: you may use the fact that $\sum_{j=1}^n (X_j - \bar{X}_n)^2$ has χ^2 distribution with $n-1$ degrees of freedom, and that the variance of a χ^2 random variable is twice the number of degrees of freedom)

2. Let $X_1, X_2, \ldots$ be i.i.d random variables, each with density given by f_θ, where $\theta \in \Theta \subseteq \mathbb{R}$ is an unknown parameter. For any $n \geq 1$, a maximum likelihood estimator Y_n satisfies

$$\prod_{i=1}^n f_{Y_n}(X_i) = \sup_{\theta \in \mathbb{R}} \prod_{i=1}^n f_\theta(X_i).$$

   (a) State (without the proof) the most general version of the maximum likelihood estimator (MLE) consistency theorem familiar to you which concludes that $Y_1, Y_2, \ldots$ converges in probability to the true value of θ, under some assumptions.

   (b) Consider the density of the uniform distribution on $[0, 1]$, that is $f(x) = 1$ for all $x \in [0, 1]$ and $f(x) = 0$ otherwise, and define $f_\theta(x) = f(x - \theta)$, $\theta \in \mathbb{R}$. Find an MLE Y_n of θ in this case. Is it unique?

   (c) Is the MLE that you found consistent? If not, explain why. If it is consistent, can it be proven using the result that you stated in part (a)? If not, explain why and establish consistency in a different way.

1. Let $X_1, \ldots, X_n$ be i.i.d. random variables, so that X_1 has probability density function $f_\theta \colon \mathbb{R} \to [0, \infty)$, where $\theta > 0$ is an unknown parameter and

$$f_\theta(x) := \begin{cases} \frac{2x}{\theta^2}, & 0 \le x \le \theta, \\ 0, & \text{else.} \end{cases}$$

   (a) Find any method of moments estimator $\widehat{\theta}_n$ of θ. Is $\widehat{\theta}_n$ unbiased?

   (b) Show that $\widehat{\theta}_n$ converges in probability as $n \to \infty$.

   (c) Show that $\widehat{\theta}_n$ converges in distribution as $n \to \infty$, and identify the limiting distribution.

   (d) Prove or disprove the following statement: let $W_1, W_2, \ldots$ be real random variables that converge in distribution to W. Let $Z_1, Z_2, \ldots$ be real random variables that converge in distribution to Z. Then $W_1 + Z_1, W_2 + Z_2, \ldots$ converges in distribution to $W + Z$.

2. Assume that $Y_1, \ldots, Y_n$ are independent and generated from a linear model $Y_j = \alpha + \beta x_j + \varepsilon_j$, where $\alpha, \beta \in \mathbb{R}$ are unknown, x_j's are not all equal, and $\varepsilon_j,\ j = 1., \ldots, n$ are i.i.d. $N(0,1)$ random variables.

   (a) Write down the likelihood function and find a complete, sufficient statistic for (α, β). Justify both sufficiency and completeness.

   (b) Find the maximum likelihood estimator (MLE) of the pair (α, β).

   (c) Show that the MLE is unbiased.

   (d) Show that the MLE has the smallest variance among all unbiased estimators.

1. Let $X_1, \ldots, X_n$ be i.i.d. with Poisson distribution $P(\lambda)$ where $\lambda \in (0, \infty)$, that is, for all $i = 1, \ldots, n$ that $P(X_i = k) = e^{-\lambda}\lambda^k/k!, k = 0, 1, \ldots$ and are independent.

   (a) Show that this family of distributions has the monotone likelihood ratio property with respect to an appropriately chosen statistic $T(X_1, \ldots, X_n)$.

   (b) For the statistical model described, give an example of a hypothesis testing problem $H_0 : \theta \in \Theta_0,\ H_a : \theta \in \Theta_a$ where Θ_0, Θ_a are two subsets of $(0, \infty)$ that satisfy $\Theta_0 \cap \Theta_a = \emptyset,\ \Theta_0 \cup \Theta_a = (0, \infty)$, that admits a uniformly most powerful test of any size $\alpha \in [0, 1]$. Justify your answer.

   (c) For the statistical model described above, give an example of a hypothesis testing problem $H_0 : \theta \in \Theta_0,\ H_a : \theta \in \Theta_a$, with $\Theta_0,\ \Theta_a$ satisfying the same properties as in part (b), that **does not** admit a uniformly most powerful test of given size $\alpha \in (0, 1)$. Justify your answer.

2. Let $X = (X_1, \ldots, X_n)$ be a random sample of size n from a family of probability densities $\{f_\theta : \theta \in \mathbb{R}\}$, so that $f_\theta : \mathbb{R}^n \to (0, \infty)$ for any $\theta \in \mathbb{R}$, and $X_1, \ldots, X_n$ are i.i.d. Fix $\theta_0 \in \mathbb{R}$. Suppose we test the hypothesis H_0 that $\{\theta = \theta_0\}$ versus the alternative $\{\theta \neq \theta_0\}$. Let $Y = Y_n$ denote the MLE of θ, and assume that under P_{θ_0} that

$$Y_n \to_p \theta_0 \quad \text{and} \quad \sqrt{n}(Y - \theta_0) \to_d \mathcal{N}(0, I_{X_1}^{-1})$$

   where $\to_p$ and $\to_d$ denote convergence in probability and in distribution respectively, $\mathcal{N}$ denote the normal distribution and I_X the Fisher information, which we assume exists.

   Let

$$\lambda(X) := \frac{\sup_{\theta \in \mathbb{R}} f_\theta(X)}{f_{\theta_0}(X)}$$

   denote the generalized likelihood ratio statistic. If H_0 is true, we will ask you to show that $-2\log\lambda(X)$ converges in distribution as $n \to \infty$ to a chi-squared random variable with one degree of freedom, under some additional assumptions.

   Fix $x \in \mathbb{R}^n$ and denote $\ell_n(\theta) := \log f_\theta(x)$. Below you may assume whatever smoothness conditions you require for your argument, but please note them when they are applied, in the form of, say: here we assume that a second order Taylor expansion for $\ell_n(\theta)$ holds with appropriate form of remainder.

   (a) Using Taylor series, show that

$$-2\log\lambda(X) = -\ell_n''(\widehat{Y})(\theta_0 - Y)^2$$

   where $\widehat{Y}$ is some point in an interval with endpoints Y and θ_0.

   (b) Using the weak law of large numbers, show that, for any $\theta \in \mathbb{R}$, $\frac{1}{n}\ell_n''(\theta)$ converges in probability to the constant $I_{X_1}(\theta)$ as $n \to \infty$, and that the same conclusion holds for $\frac{1}{n}\ell_n''(\widehat{Y})$.

   (c) Combining the above observations, conclude that $-2\log\lambda(X)$ converges in distribution to a chi-squared random variable with one degree of freedom.

1. Let $X_1, ..., X_n$ be an i.i.d. sample from the exponential distribution $\mathsf{Exp}(\lambda)$ with probability density function
$$f(x; \lambda) = \lambda e^{-\lambda x}, \quad x > 0,$$
parametrized with $\lambda > 0$ (or the "rate parameter") which we are tasked to estimate. Let $n > 1$.

   (a) Show that the maximum likelihood estimator of λ is $\widehat{\lambda}_n = \frac{n}{\sum_{j=1}^n X_j}$.

   (b) Show that $\widehat{\lambda}_n$ is biased. On average, does it overestimate or underestimate λ? You can assume that $\mathbb{E}\widehat{\lambda}_n$ is finite.
   (hint: the function $f(x) = \frac{1}{x}$ has some properties that could be useful in this question)

   (c) Show that (for example, using δ-method)
$$\sqrt{n}(\widehat{\lambda}_n - \lambda) \overset{n\to\infty}{\longrightarrow} N(0, \sigma_\infty^2)$$
   for some $\sigma_\infty^2 > 0$, where $\overset{n\to\infty}{\longrightarrow}$ denotes convergence in distribution. Find σ_∞^2.

2. Answer the following questions (a correct answer with incorrect justification is worth 0 points):

   (a) Let $X_1, \ldots, X_n$ be i.i.d. random variables generated according to some distribution from the family $P_\theta,\ \theta \in \Theta \subseteq \mathbb{R}$, where the value of the parameter θ is unknown. Give a definition of a sufficient statistic. Does a sufficient statistic always exist?

   (b) Prove or disprove: a Uniform Minimal Variance Unbiased estimator always exists.
   (hint: consider a sample of size 1 from the Bernoulli distribution with parameter $p \in [0, 1]$)

   (c) Prove or disprove: a maximum likelihood estimator is always unbiased.

Spring 2024 Math 541a Exam

Problem 1. Suppose that $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) random variables with the common distribution $N(\theta, \theta^2)$, where $\theta \in (0, \infty)$ is the unknown parameter. Let $W_n = n^{-1} \sum_{i=1}^n X_i^2$.

Fact: if $X \sim N(\mu, \sigma^2)$, then $\mathbb{E}X^3 = \mu^3 + 3\mu\sigma^2$ and $\mathbb{E}X^4 = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4$.

1. Find a sequence of constants a_n and a function $b(\theta)$ such that $a_n(W_n - b(\theta)) \xrightarrow{d} T$, where T is a non-degenerate random variable with mean zero. What is the variance of T? Here, a non-degenerate random variable means its variance is strictly positive.
2. Find the MLE $\hat{\theta}$ of θ.
3. Find a sequence of constants c_n and a function $d(\theta)$ such that $c_n(\hat{\theta} - d(\theta)) \xrightarrow{d} U$, where U is a non-degenerate random variable with mean zero. What is the variance of U? [Hint: consider the multivariate Δ-method.]
4. Find a non-degenerate variance stabilizing transformation h of W_n, i.e., the asymptotic non-zero variance of $h(W_n)$ does not depend on θ.

Problem 2. The Information Inequality. Let $\boldsymbol{X}$ be an observation in $\mathbb{R}^n$ with density $p(\boldsymbol{x}, \theta)$ for $\theta \in \Theta \subset \mathbb{R}^p$, and $T = T(\boldsymbol{X})$ an estimator.

1. Use the Cauchy Schwarz inequality to prove the information inequality

$$\mathrm{Var}(T) \geq \frac{[\dot{g}(\theta)]^2}{I(\theta)}$$

when $p = 1$, that is, when the parameter space is a subset of the real line, and where $g(\theta) = E_\theta[T]$ and $\dot{g}$ is its derivative with respect to θ, and $I(\theta) = \mathrm{Var}_\theta(U(\theta, \boldsymbol{X})$ where

$$U(\theta, \boldsymbol{x}) = \frac{\partial}{\partial \theta} \log p(x, \theta).$$

2. Consider now the Information Inequality when $\theta \in \mathbb{R}^d$ and $T \in \mathbb{R}^{1 \times r}$ is an estimate of r functions of θ. Under regularity on the density, we obtain

$$\mathrm{Var}_\theta(T(\boldsymbol{X})^\top) \geq \dot{g}(\theta)^\top \mathbf{I}(\theta)^{-1} \dot{g}(\theta),$$

where each column of $\dot{g}$ is the gradient with respect to θ of the row entry of g, and where the information matrix is now the 'variance-covariance' matrix of the score function.

Let $\boldsymbol{X}$ be a multivariate normal observation with unknown mean μ and known invertible covariance matrix Σ, that is, with density

$$p(x, \mu) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)' \Sigma^{-1} (x - \mu) \right)$$

Compute the information lower bound for the unbiased estimation of

$$g(\mu) = \frac{1}{2} \sum_{i=1} \mu_i^2$$