Problem 1

Let $(X_i, i = 1, \dots, n)$ be a normally distributed random sample with $X_i \sim \mathcal{N}(\mu, \sigma^2)$ for $i = 1, \dots, n$. Denote the sample mean by $\bar{X} := 1/n \sum_{i=1}^n X_i$ and the sample variance by $S^2 := 1/(n-1) \sum_{i=1}^n (X_i - \bar{X})^2$.

(a) Show that the family of normal pdfs is an exponential family.

(b) Show that $(\bar{X}, S^2)$ is a minimal sufficient statistic for (μ, σ^2).

(c) Derive the likelihood function for σ^2. Compute the maximum likelihood estimator $\hat{\sigma}^2$ for σ^2, where you are allowed to assume that the maximization with respect to μ can be done independently and the maximum likelihood estimator is known to be $\hat{\mu} = \bar{X}$.

(d) Define the mean squared error of an estimator of a parameter and compute it for S^2 as well as for $\hat{\sigma}^2$. (Hint: You are allowed to use that $\mathsf{Var}(S^2) = 2\sigma^4/(n-1)$.)

(e) Give at least one reason why S^2 should be preferred over $\hat{\sigma}^2$ and at least one reason why $\hat{\sigma}^2$ performs better than S^2.

(23 points)

Problem 2

Let f be a unimodal pdf. If the interval $[a, b]$ satisfies

(i) $\displaystyle\int_a^b f(x)\,dx = 1 - \alpha$,

(ii) $f(a) = f(b) > 0$,

(iii) $a \leq x^* \leq b$, where x^* is a mode of f,

then $[a, b]$ is a shortest among all intervals that satisfy (i).

(a) Prove the claim by contradiction by assuming that there exists a shorter interval $[a', b']$ that satisfies (i) with $a' > a$ and that the claim is already proven for $a' \leq a$.

(b) Give a counterexample to show that the shortest interval that satisfies (i) is not necessarily unique.

(c) Let $(X_i, i = 1, \dots, n)$ be a random sample with $X_i \sim \mathcal{N}(\mu, \sigma^2)$ for $i = 1, \dots, n$. For fixed $\alpha \in (0, 1)$ compute the shortest $1 - \alpha$ confidence interval for the parameter μ.

(13 points)

Problem 3
Let $(X_i, i = 1, \dots, n)$ be a random sample with $X_i \sim \mathcal{N}(\theta, \sigma^2)$ for $i = 1, \dots, n$ and suppose that the prior distribution on θ is $\mathcal{N}(\mu, \tau^2)$, where σ^2, μ, and τ^2 are assumed to be known. In an exercise it was shown that the posterior distribution $\pi(\cdot|\bar{x})$ is normal with

$$\mathbb{E}(\theta|\bar{x}) = \frac{\tau^2}{\tau^2 + \sigma^2/n}\bar{x} + \frac{\sigma^2/n}{\tau^2 + \sigma^2/n}\mu \quad \text{and} \quad \mathsf{Var}(\theta|\bar{x}) = \frac{\tau^2\sigma^2/n}{\tau^2 + \sigma^2/n},$$

where $\bar{X}$ denotes the sample mean of the random sample.

(a) Introduce the Bayesian approach to statistics.

(b) Recall the definition of a conjugate family and give an example.

(c) What is a possible point estimate in the Bayesian framework? Compute this explicitly for the parameter θ.

(d) Construct a Bayesian test for the hypothesis $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$ by deriving an acceptance region.

(e) Give the definition of a credible set and construct for fixed $\alpha > 0$ a $1 - \alpha$ credible set for θ such that it is the highest posterior density region. What is an interpretation of a $1 - \alpha$ credible set?

(16 points)

Problem 4
Let $(W_n, n \in \mathbb{N})$ be a consistent sequence of estimators of a parameter θ and $(a_n, n \in \mathbb{N})$ and $(b_n, n \in \mathbb{N})$ real-valued sequences such that $\lim_{n\to+\infty} a_n = 1$ and $\lim_{n\to+\infty} b_n = 0$.

(a) Give the definition of a consistent sequence of estimators of a parameter θ.

(b) Show that $(a_nW_n + b_n, n \in \mathbb{N})$ is a consistent sequence of estimators of θ.

(c) What does the result imply for observations of finite size and consistency in the sense of this problem?

(8 points)