Assignment 4

(1) When X has a Binomial $\mathrm{Bin}(n, p)$ distribution with $0 < p < 1$, we say that $g(p) = p/(1-p)$ is the 'odds ratio'. For instance, when $p = 2/3$ the odds ratio is 2, corresponding to 'odds' of '2 to 1' for success. Show that there is no unbiased estimator for the odds ratio $g(p)$.

(2) Let $X \sim \mathcal{P}(\lambda)$ be an observation from the Poisson distribution satisfying

$$P_\lambda(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}. \tag{1}$$

(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) Find an unbiased estimate 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.

(b) Compute the variance of the resulting estimator.

(c) Are there any other unbiased estimators of $g(\lambda)$. Is the unbiased estimator that was derived in part a) UMVU?

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

(e) What types of inferential or statistical conclusions can be drawn from this example?

(3) Let $X_1, \ldots, X_n$ be i.i.d. with the $\mathcal{U}[0, \theta]$ distribution, with $\theta \in \Theta \subset (0, \infty)$, unknown. Propose a prior distribution $\pi(\theta)$ for θ and compute its minimum mean square estimate.

(4) Consider a graph $\mathcal{G}$ with vertex set $[n] = \{1, \ldots, n\}$ and some given edge set $\mathcal{E}$, the set of all pairs $\{i, j\}$ for which vertices i and j share an edge; we let $i \sim j$ denote that $\{i, j\} \in \mathcal{E}$, and we exclude the case of 'self edges' that would connect any vertex to itself.

For $\beta \in \mathbb{R}$ consider the distribution on $\boldsymbol{X} \in \{-1, 1\}^n$ given by

$$P_\beta(\boldsymbol{X} = \boldsymbol{x}) = \frac{\exp\left(\beta \sum_{i \neq j} X_i X_j\right)}{Z_\beta}$$

where Z_β is a normalizing constant. Prove that

$$P_\beta(X_i | X_j \neq i) = P_\beta(X_i | X_j, j \sim i).$$

(5) Let $p(\boldsymbol{x}; \theta), \theta \in \Theta \subset \mathbb{R}$ be family of density or probability mass functions.

(a) Propose a set of conditions on $p(\boldsymbol{x}; \theta), \theta \in \Theta \subset \mathbb{R}$, and use the dominated convergence theorem, or any result of your choice, to prove that they are sufficient to permit interchanging integration over $\mathbb{R}$ and derivative with respect to θ. Show these conditions are satisfied by the normal family with known variance, and θ equal to the mean.

(b) Does the set of proposed conditions also permit that same interchange for the Cauchy location family, given by

$$p(x; \mu) = \frac{1}{\pi(1 + (x - \mu)^2)} \quad \mu \in \mathbb{R}. \tag{2}$$

(c) Propose a set of conditions on the family of densities $p(\boldsymbol{x}; \theta)$ such that

$$I_{\boldsymbol{X}}(\theta) = -E_\theta[\partial_\theta^2 \log p(\boldsymbol{X}, \theta)],$$

(6) Compute the information $I_X(\mu)$ for μ in the Cauchy location family (2), and, via invoking an appropriate theorem, provide the asymptotic distribution of the maximum likelihood estimator for μ.

(7) Prove the multivariate information bound: Hint: compute the covariance matrix of $T(\boldsymbol{X})' - \dot{g}(\theta)' \boldsymbol{I}(\theta)^{-1} U(\theta, \boldsymbol{X})$).

(8) Compute the information matrix for the vector μ contained in one observation from a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$ with non-singular covariance matrix Σ.

(9) For what classical families of distributions, (such as the Poisson, Gamma, Beta, and Binomial) can one apply directly apply the central limit theorem to verify that the maximum likelihood estimate, centered and properly scaled, converges in distribution to a mean zero normal distribution with variance (or covariance matrix) given by the inverse Information?

(10) Let $X_1, \ldots, X_n$ be i.i.d from $p(\boldsymbol{x}, \theta), \theta \in \Theta \subset \mathbb{R}^d$ a three times differentiable density function in θ and positive definite Fisher information matrix I_θ that can be expressed as $-E_\theta[\partial_\theta^2 \log p(x, \theta)]$, and let $U(\boldsymbol{X}_n, \theta)$ be the score function for the data $\boldsymbol{X}_n = (X_1, \ldots, X_n)^\top$, and $U_n(\boldsymbol{X}_n, \theta) = U(\boldsymbol{X}_n, \theta)/n$

Expanding the score function over the neighborhood $\|\theta - \theta_0\| \leq \delta$ of θ_0, for a judiciously chosen $\delta > 0$, yields

$$U_n(\boldsymbol{X}_n, \theta) = U_n(\boldsymbol{X}_n, \theta_0) + \dot{U}_n(\boldsymbol{X}_n, \theta_0)(\theta - \theta_0) + R_n \tag{3}$$

where $R_n = o(\|\theta - \theta_0\|^2)$. By the law of large numbers, as $n \to \infty$,

$$U_n(\boldsymbol{X}_n, \theta_0) \to_p 0 \quad \text{and} \quad \dot{U}_n(\boldsymbol{X}_n, \theta_0) \to_p -I_\theta.$$

Consider the large sample version of (3), ignoring the lower order remainder,

$$U_n(\boldsymbol{X}_n, \theta) = -I_\theta(\theta - \theta_0) \quad \text{for} \quad \|\theta - \theta_0\| \leq \delta \tag{4}$$

(a) Show that in the one dimensional situation, that is, for $d = 1$, (4) must have a root in the given neighborhood of radius δ. Hint: Use the intermediate value theorem.

(b) In multidimension, one needs a deeper argument. If U_n has no root in $\|\theta - \theta_0\| \leq \delta$ then the function f given by

$$f(\boldsymbol{t}) = \frac{U_n(\boldsymbol{X}_n, \theta_0 + \delta \boldsymbol{t})}{\|U_n(\boldsymbol{X}_n, \theta_0 + \delta \boldsymbol{t})\|} \quad \text{for } \boldsymbol{t} \in \boldsymbol{R}^d \text{ with } \|\boldsymbol{t}\| \leq 1$$

is a continuous mapping from the the unit ball to itself and by the Brouwer fixed point theorem must have a fixed point τ, that is, a value such that

$$f(\tau) = \tau \quad \text{and in particular} \quad \|\tau\| = \|f(\tau)\| = 1.$$

Use this result to obtain a contradiction, and show (4) has root within a δ neighborhood of the true parameter. (Hint: Multiply on the left of both sides of (4) by $(\theta - \theta_0)^\top$. To make the notation lighter, you can, by reparameterizing, assume $\theta_0 = 0$ without of generality.)

(11) For the maximum likelihood estimate $X_{(n)}$ of θ from an independent sample of size n from the $U[0, \theta]$ distribution, show that $n(X_{(n)} - \theta)$ converges to a non-trivial distribution E, and find the density function of the limit E.