MATH562 – Fall 2025

Problem Set: Week 7

1. If $Y_1, \dots, Y_n \overset{i.i.d.}{\sim}$ Poiss(λ), justify why $S = \sum_{j=1}^n Y_j$ is a complete minimal sufficient statistic for λ, and state its distribution.

   (a) What is the conditional distribution of $Y_1, \dots, Y_n$ given that $S = s$?

   (b) One strategy to find an optimal unbiased estimator of some function $\psi(\lambda)$ is to find any unbiased estimator T of $\psi(\lambda)$, and then to compute $h(S) = \mathrm{E}(T \mid S)$. Another strategy is to find the function $h(s)$ that satisfies $\mathrm{E}\{h(S)\} = \psi(\lambda)$ for all λ. Will these give the same estimator?

   (c) Find minimum variance unbiased estimators of (i) $e^{-\lambda}$ and (ii) $e^{-2n\lambda}$. Do you think these are reasonable? If not, suggest better estimators.

2. Let $M = \max(Y_1, \dots, Y_n)$ and $\bar{Y}$ be the maximum and average of $Y_1, \dots, Y_n \overset{i.i.d.}{\sim} U(0, \theta)$, respectively. Recall from the lectures that M is a complete minimal sufficient statistic.

   (a) Show that $U = 2\bar{Y}$ is unbiased for θ and compute its variance.

   (b) Use the Rao–Blackwell theorem to get a better unbiased estimator. Compute its variance.

*3. Observations $\dots, Y_1, \dots, Y_n, \dots$ arise in time order.

   (a) Starting from

$$f_{Y_1,\dots,Y_n}(y_1, \dots, y_n) = f_{Y_n \mid Y_1,\dots,Y_{n-1}}(y_n \mid y_1, \dots, y_{n-1}) f_{Y_1,\dots,Y_{n-1}}(y_1, \dots, y_{n-1}),$$

   establish the *prediction decomposition*

$$f_{Y_1,\dots,Y_n}(y_1, \dots, y_n) = f_{Y_1}(y_1) \prod_{j=2}^{n} f_{Y_j \mid Y_1,\dots,Y_{j-1}}(y_j \mid y_1, \dots, y_{j-1}).$$

   (b) A stationary first-order Gaussian autoregressive process satisfies

$$Y_j \mid Y_1 = y_1, \dots, Y_{j-1} = y_{j-1} \sim \mathbb{N}\{\mu + \alpha(y_{j-1} - \mu), \sigma^2\}, \quad j = 1, \dots, n,$$

   where $|\alpha| < 1$, $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Find the log likelihood for data $y_0, y_1, \dots, y_n$ from this model if the initial value y_0 is treated (i) as a known constant and (ii) as coming from the stationary distribution, $\mathbb{N}\{\mu, \sigma^2/(1 - \alpha^2)\}$.

   (c) Give a minimal sufficient statistic for $\theta = (\mu, \sigma^2, \alpha)$ in (b). Does the model with parameter vector θ form an exponential family?

4. Consider discrete data Y with density $f(y; \theta)$ defined for $y \in \mathcal{Y}$ and let $T = t(Y)$ be a statistic based on Y. Define the sets $\mathcal{T} = \{t(y) : y \in \mathcal{Y}\}$ and $\mathcal{C}_s = \{y \in \mathcal{Y} : t(y) = s\}$ for $s \in \mathcal{T}$.

   (a) Show that the phrase '$y \sim y'$ if and only if $y, y' \in \mathcal{C}_s$' defines an equivalence relation, and that the same equivalence relation is given by taking any bijective function of $t(y)$. Deduce that the equivalence classes form a partition $\mathcal{P}_T$ of $\mathcal{Y}$.

   (b) If T is sufficient, show that the conditional distribution of Y given that $Y \in \mathcal{C}_s$ does not depend on θ; then $\mathcal{P}_T$ is called a *sufficient partition*.

   (c) If Y consists of n independent Poisson variables with mean θ, show that $T = (Y_1, Y_2 + \cdots + Y_n)$ is sufficient and give $\mathcal{Y}$, $\mathcal{T}$ and the $\mathcal{C}_s$. Find a coarser sufficient partition, and check if it is minimal.