Midterm Exam: Statistical Inference

Due: Wed Mar 6 3:50pm

For $\theta \in \Theta = \mathbb{R}_+ = (0, \infty)$ and $x \in \mathbb{N} = \{0, 1, ...\}$ set

$$f(x \mid \theta) = \frac{\theta^x \, e^{-\theta}}{x!} \tag{1}$$

and answer the following questions about this $\mathsf{Po}(\theta)$ distribution.

1. Let X be a single observation from $f(x \mid \theta)$.

   a. Is this an exponential family? If so, write $f(x \mid \theta)$ in standard form
   $$f(x \mid \theta) = e^{\eta(\theta) \cdot T(x) - B(\theta)} \, h(x)$$
   for suitable $q \in \mathbb{N}$, $\eta(\theta) \in \mathbb{R}^q$, $T(x) \in \mathbb{R}^q$, $B(\theta) \in \mathbb{R}$, and $h(x) \geq 0$ (you must specify q, η, T, B, and h); if not, explain why.

   b. Find the Fisher Information $I(\theta)$.

2. Let $\vec{x}_n$ be a random sample $\vec{x}_n = (x_1, ..., x_n)$ of $n \in \mathbb{N}$ independent observations.

   a. Find a one-dimensional sufficient statistic $S(\vec{x}_n)$. Verify that your S is sufficient. Find its mean and variance, as functions of θ.

   b. Find the *observed* information $i(\theta, \vec{x}_n)$. Does it depend on $\vec{x}_n$ at all? If so, only through your sufficient statistic?

3. Again with a random sample $\vec{x}_n = (x_1, ..., x_n)$ from $f(x \mid \theta)$,

   a. Find the maximum likelihood estimator $\hat{\theta}_n(\vec{x}_n)$.

   b. Find the squared-error risk function
   $$R(\theta, \hat{\theta}_n) = \mathsf{E}\big[\, |\hat{\theta}_n - \theta|^2 \mid \theta \,\big].$$
   Is $\hat{\theta}_n$ consistent? Why? Plot $R(\theta, \hat{\theta}_4)$.

   c. Find the (absolute) efficiency
   $$\mathsf{Eff}(\hat{\theta}_n) = n \, I(\theta) \, R(\theta, \hat{\theta}_n)$$
   for each $n \in \mathbb{N}$ and $\theta \in \Theta$. Is $\hat{\theta}_n$ asymptotically efficient?

d. For $n = 4$ find an exact 90% equal-tail confidence interval $\big[L(\vec{x}_4),\ R(\vec{x}_4)\big]$, *i.e.*, find statistics L and R with the properties

$$\mathsf{P}[\theta < L(\vec{x}_4) \mid \theta] \le 0.05 \qquad \mathsf{P}[\theta > R(\vec{x}_4) \mid \theta] \le 0.05$$

Evaluate your interval, first for the data set $\vec{x}_4 = (25, 35, 22, 18)$ and then for $\vec{x}_4 = (0, 0, 0, 0)$.

4. Now let's adopt a Bayesian perspective.

a. For a gamma prior distribution $\pi \sim \mathsf{Ga}(\alpha, \lambda)$, *i.e.*,

$$\pi(\theta) = \frac{\lambda^\alpha\, \theta^{\alpha-1}}{\Gamma(\alpha)}\, e^{-\lambda\theta}, \qquad \theta > 0$$

find the posterior distribution $\pi(\theta \mid \vec{x}_n)$ for a sample of size n and the posterior mean (Bayes estimator) $\bar{\theta}_n^\pi(\vec{x}_n) = \mathsf{E}[\theta \mid \vec{x}_n]$.

b. Find the squared-error risk function

$$R(\theta, \bar{\theta}_n^\pi) = \mathsf{E}\big[\, |\bar{\theta}_n^\pi - \theta|^2 \mid \theta\,\big].$$

Is $\bar{\theta}_n^\pi$ consistent? Why? Plot $R(\theta, \bar{\theta}_4^\pi)$ for $\alpha = 1$, $\lambda = 1$.

c. Find the efficiency

$$\mathsf{Eff}(\bar{\theta}_n^\pi) = n\, I(\theta)\, R(\theta, \bar{\theta}_n^\pi)$$

for each $n \in \mathbb{N}$, $\theta \in \Theta$, $\alpha > 0$, and $\lambda \ge 0$. Is $\bar{\theta}_n^\pi$ asymptotically efficient? Why?

d. Find the Jeffreys prior $\pi_J(\theta)$, the posterior distribution $\pi_J(\theta \mid \vec{x}_n)$, and the posterior mean $\bar{\theta}_n^J(\vec{x}_n) = \mathsf{E}^J[\theta \mid \vec{x}_n]$. Does the Jeffreys posterior distribution coincide with any of the gamma posteriors from part a. above? Which?

e. For $n = 4$ find an exact 90% equal-tail credible interval $\big[L(\vec{x}_4),\ R(\vec{x}_4)\big]$ for the Jeffreys prior distribution, *i.e.*, find statistics L and R with the properties

$$\mathsf{P}^J[\theta < L(\vec{x}_4) \mid \vec{x}_4] \le 0.05 \qquad \mathsf{P}^J[\theta > R(\vec{x}_4) \mid \vec{x}_4] \le 0.05$$

Evaluate your interval for the data sets $\vec{x}_4 = (25, 35, 22, 18)$ and $\vec{x}_4 = (0, 0, 0, 0)$.

5. Let $\vec{x}_n$ be a random sample and let $\bar{x}_n$ be the sample mean.

   a. Find a Method of Moments estimator $T_n^M(\vec{x}_n)$ of θ based on the sample mean $\bar{x}_n$.

   b. Show that it is consistent in probability, *i.e.*, show that

   $$\mathsf{P}\big[|T_n^M(\vec{X}_n) - \theta| > \epsilon\big] \to 0$$

   as $n \to \infty$ for any $\epsilon > 0$.

   c. Does T_n^M coincide with the MLE $\hat{\theta}_n$, or with the Bayes estimator $\bar{\theta}_n^\pi$ for any Gamma prior $\pi \sim \mathsf{Ga}(\alpha, \lambda)$, or with the limit of such Bayes estimators as α or λ tend to extremes? If so, which? If not, evaluate the risk function $R(\theta, T_n^M)$ for each $n \in \mathbb{N}$ and $\theta \in \Theta$.

6. In the past three problems you have computed four estimators for θ above ($\hat{\theta}_n$, $\bar{\theta}_n^\pi$, $\bar{\theta}_n^J$, T_n^M) and their risk functions $R(\theta, \cdot)$. Let $n = 4$.

   a. For which θ is T_J better than $\hat{\theta}$? For which θ is $\bar{\theta}_n^\pi$ better than $\hat{\theta}$, for $\alpha = 1$ and $\lambda = 1$? For which θ is T_n^M better than $\hat{\theta}$?

   b. In a **brief** paragraph or so, how would you choose among these estimators for a particular problem? Which would you recommend, and why? If any features of the problem would be relevant, identify them and explain.

7. a. Does the MLE $\hat{\theta}_n$ have an asymptotically normal distribution for this distribution? If so, find the mean $\mu_n(\theta)$ and variance $\sigma_n^2(\theta)$ of $\hat{\theta}_n$ and give statistics $\hat{\mu}_n(\vec{x}_n)$ and $\hat{\sigma}_n^2(\hat{x}_n)$ that estimate μ_n and σ_n^2.

   b. Assuming asymptotic normality find an asymptotic 90% equal-tail confidence interval based on these estimators, *i.e.*, find **statistics** $L_n(X)$, $R_n(X)$ satisfying $\mathsf{P}\big[\theta \in [L_n, R_n]\big] \approx 0.90$ because

   $$\mathsf{P}[\theta < L_n(\vec{x}_n) \mid \theta] \approx 0.05 \qquad \mathsf{P}[\theta > R_n(\vec{x}_n) \mid \theta] \approx 0.05$$

   for large enough n.

   c. Calculate the interval and comment on its properties for the $n = 4$ data set $\vec{x}_4 = (25, 35, 22, 18)$. Is $[L, R]$ reasonable?

   d. Calculate the interval and comment on its properties for the $n = 4$ data set $\vec{x}_4 = (0, 0, 0, 0)$. Is $[L, R]$ reasonable?

Summary of some `S-Plus` Commands

S-Plus provides a suite of functions for each of the commonly used probability distributions. For example, for the Poisson distribution

$$X \sim \mathsf{Po}(\lambda): \qquad \mathsf{P}[X \in A] = \sum_{x \in A \cap \mathbb{N}} \frac{\lambda^x}{x!} e^{-\lambda}$$

and Gamma distribution

$$Y \sim \mathsf{Ga}(\alpha, \lambda): \qquad \mathsf{P}[Y \in A] = \int_{A \cap \mathbb{R}_+} \frac{\lambda^\alpha \, y^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda y} \, dy$$

we have the probability mass function (pmf) and density function (pdf)

$$\texttt{dpois(x,lambda)} = \frac{\lambda^x}{x!} e^{-\lambda}$$

$$\texttt{dgamma(y,alpha,lambda)} = \frac{\lambda^\alpha \, y^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda y}$$

and the cumulative distribution functions (CDF's)

$$\texttt{ppois(z,lambda)} = \mathsf{P}[X \le z] = \sum_{x=0}^{z} \frac{\lambda^x}{x!} e^{-\lambda}$$

$$\texttt{pgamma(z,alpha,lambda)} = \mathsf{P}[Y \le z] = \int_0^z \frac{\lambda^\alpha \, y^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda y} \, dy$$

and the inverse CDF's, the quantile functions

$$\texttt{qpois(q,lambda)} = z \Leftrightarrow \texttt{ppois(z,lambda)} = q$$

$$\texttt{qgamma(q,alpha,lambda)} = z \Leftrightarrow \texttt{pgamma(z,alpha,lambda)} = q$$

These are related (in class we used a fish ("poisson") process to motivate this) through (for every $k \in \mathbb{N}$, $t \ge 0$, $\lambda > 0$)

$$\texttt{ppois(k,lambda*t)} = 1 - \texttt{pgamma(t,k+1,lambda)}$$

There are also functions `rpois(n,lambda)` and `rgamma(n,alpha,lambda)` to generate random samples of any size $n \in \mathbb{N}$ from the distributions.

Warning: The open-source work-alike product `R` has similar functions and syntax, but `R` parameterizes the Gamma distribution a little differently from S-Plus— it uses a *scale* parameter $\beta = 1/\lambda$ instead of a *rate* parameter λ. If you use `R`, just remember to replace `lambda` with `beta=1/lambda` in the arguments for `dgamma`, `pgamma`, `qgamma`, and `rgamma`.