S&DS 242/542: Homework 8

Due Wednesday, April 2, at 1PM

1. **Migration rates.** To study the rates of migration between cities 1, 2, and 3, the locations of n people living in these cities at the start and end of a 10-year period were recorded. Let N_{11} be the number of people who started and ended in city 1, N_{12} be the number of people who started in city 1 and ended in city 2, etc. We model the counts $(N_{ij})_{1 \leq i,j \leq 3}$ as Multinomial$(n, (p_{ij})_{1 \leq i,j \leq 3})$, where $(p_{ij})_{1 \leq i,j \leq 3}$ is a probability vector with 9 entries summing to 1.

We wish to test the "equilibrium" null hypothesis that $p_{12} = p_{21}$, $p_{13} = p_{31}$, and $p_{23} = p_{32}$. Express the generalized likelihood ratio test statistic for this problem as a simple formula in the observed counts $(N_{ij})_{1 \leq i,j \leq 3}$ and n, and describe how you would carry out a level-α test of this null hypothesis when n is large.

2. **The Laplace distribution.** Suppose that $X_1, \ldots, X_n$ are IID with PDF

$$f(x \mid \mu, b) = \frac{b}{2} \exp\left(-b|x - \mu|\right)$$

where $\mu \in \mathbb{R}$ and $b > 0$.

(a) If (μ, b) are both unknown, what are the MLEs $(\hat{\mu}, \hat{b})$? You may assume that n is odd.

(b) Suppose $\mu = 0$ is known, and for b consider a prior $B \sim \text{Gamma}(\alpha, \beta)$. What is the posterior distribution of B? What is its posterior mean?

3. **Bayesian inference for multinomial proportions.**
The Dirichlet$(\alpha_1, \ldots, \alpha_K)$ distribution for $\alpha_1, \ldots, \alpha_K > 0$ is a joint distribution for random variables $(P_1, \ldots, P_K)$ such that $0 \leq P_1, \ldots, P_K \leq 1$ and $P_1 + \ldots + P_K = 1$. It has PDF[1]

$$f(p_1, \ldots, p_K \mid \alpha_1, \ldots, \alpha_K) = \frac{\Gamma(\alpha_1 + \ldots + \alpha_K)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_K)} p_1^{\alpha_1 - 1} \times \ldots \times p_K^{\alpha_K - 1}.$$

Letting $\alpha = \alpha_1 + \ldots + \alpha_K$, you may use without proof that this distribution satisfies

$$\mathbb{E}[P_i] = \frac{\alpha_i}{\alpha}, \quad \text{Var}[P_i] = \frac{\alpha_i(\alpha - \alpha_i)}{\alpha^2(\alpha + 1)}.$$

[1] This may be interpreted as a joint PDF in the variables $p_1, \ldots, p_{K-1}$, over the domain $p_1, \ldots, p_{K-1} \geq 0$ and $p_1 + \ldots + p_{K-1} \leq 1$, where $p_K = 1 - p_1 - \ldots - p_{K-1}$. For $K = 2$, this is the Beta(α_1, α_2) distribution.

(a) Let $(X_1, \ldots, X_6) \sim \text{Multinomial}(n, (p_1, \ldots, p_6))$. Consider the prior $(P_1, \ldots, P_6) \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_6)$. What is the posterior distribution of $(P_1, \ldots, P_6)$ given $(X_1, \ldots, X_6)$? What is the posterior mean and variance of each parameter $P_1, \ldots, P_6$?

(b) For what choice of prior parameters $\alpha_1, \ldots, \alpha_6$ would the posterior means agree with the MLEs? How might you instead choose $\alpha_1, \ldots, \alpha_6$ to represent a strong prior belief that each $p_1, \ldots, p_6$ is close to 1/6?

4. **Ridge regularization and empirical Bayes.** Suppose $X_1, \ldots, X_n \sim \mathcal{N}(\theta, 1)$, where $\theta \in \mathbb{R}$ is unknown with prior distribution $\Theta \sim \mathcal{N}(0, \tau^2)$.

(a) Show that the posterior mean and posterior mode for Θ coincide, and that both are given by the solution to the optimization problem

$$\arg\min_{\theta \in \mathbb{R}} \sum_{i=1}^{n} (X_i - \theta)^2 + \frac{1}{\tau^2} \cdot \theta^2$$

This is sometimes called a "ridge-regularized" least squares estimator with ridge penalty $\frac{1}{\tau^2}$.

(b) In the joint distribution of $(X_1, \ldots, X_n, \Theta)$, what is the marginal distribution of $\bar{X} = \frac{1}{n}(X_1 + \ldots + X_n)$? [Hint: First express $\bar{X}$ as the sum of Θ and an independent normal random variable.]

(c) In applications where we do not know how to set τ^2, an *empirical Bayes* approach is to estimate τ^2 from the data $X_1, \ldots, X_n$, assuming this Bayesian model. Using your answer in (b), suggest an unbiased estimator $\hat{\tau}^2$ for τ^2. Then give an expression for the resulting "empirical Bayes posterior mean" estimator of θ in part (a) that replaces the prior parameter τ^2 by $\hat{\tau}^2$.