STATS 200: Homework 7

Due Wednesday, November 30, at 5PM

1. **The Laplace distribution.** The Laplace (or double-exponential) distribution with mean μ and scale b is a continuous distribution over $\mathbb{R}$ with PDF

$$f(x|\mu, b) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right).$$

It is sometimes used as an alternative to the normal distribution to model data with heavier tails, as this PDF decays exponentially in $|x-\mu|$ rather than in $(x-\mu)^2$.

(a) What are the MLEs $\hat{\mu}$ and $\hat{b}$ given data $X_1, \ldots, X_n$? Why is this MLE $\hat{\mu}$ more robust to outliers than the MLE $\hat{\mu}$ in the $\mathcal{N}(\mu, \sigma^2)$ model?

You may assume that n is odd and that the data values $X_1, \ldots, X_n$ are all distinct. (Hint: The log-likelihood is differentiable in b but not in μ. To find the MLE $\hat{\mu}$, you will need to reason directly from its definition.)

(b) Suppose it is known that $\mu = 0$. In a Bayesian analysis, let us model the scale parameter as a random variable B with prior distribution $B \sim \text{InverseGamma}(\alpha, \beta)$, where $\alpha, \beta > 0$. If $X_1, \ldots, X_n \sim \text{Laplace}(0, b)$ when $B = b$, what are the posterior distribution and posterior mean of B given the data $X_1, \ldots, X_n$?

(The InverseGamma(α, β) distribution is a continuous distribution on $(0, \infty)$ with PDF

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}$$

and with mean $\frac{\beta}{\alpha-1}$ when $\alpha > 1$.)

(c) Still supposing it is known that $\mu = 0$, what is the MLE $\hat{b}$ for b in this sub-model? How does this compare to the posterior mean from part (b) when n is large?

2. **Bayesian inference for multinomial proportions.** The Dirichlet$(\alpha_1, \ldots, \alpha_K)$ distribution with parameters $\alpha_1, \ldots, \alpha_K > 0$ is a continuous joint distribution over K random variables $(P_1, \ldots, P_K)$ such that $0 \le P_i \le 1$ for all $i = 1, \ldots, K$ and $P_1 + \ldots + P_K = 1$. It has (joint) PDF

$$f(p_1, \ldots, p_k|\alpha_1, \ldots, \alpha_K) \propto p_1^{\alpha_1 - 1} \times \ldots \times p_K^{\alpha_K - 1}.$$

Letting $\alpha_0 = \alpha_1 + \ldots + \alpha_K$, this distribution satisfies

$$\mathbb{E}[P_i] = \frac{\alpha_i}{\alpha_0}, \quad \text{Var}[P_i] = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}.$$

(a) Let $(X_1, \ldots, X_6) \sim \text{Multinomial}(n, (p_1, \ldots, p_6))$ be the numbers of 1's through 6's obtained in n rolls of a (possibly biased) die. Let us model $(P_1, \ldots, P_6)$ as random variables with prior distribution $\text{Dirichlet}(\alpha_1, \ldots, \alpha_6)$. What is the posterior distribution of $(P_1, \ldots, P_6)$ given the observations $(X_1, \ldots, X_6)$? What is the posterior mean and variance of P_1?

(b) How might you choose the prior parameters $\alpha_1, \ldots, \alpha_6$ to represent a strong prior belief that the die is close to fair (meaning $p_1, \ldots, p_6$ are all close to 1/6)?

(c) How might you choose an improper Dirichlet prior to represent no prior information? How do the posterior mean estimates of $p_1, \ldots, p_6$ under this improper prior compare to the MLE?

3. **GLRT and the t-test.** Let $X_1, \ldots, X_n \stackrel{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, where both μ and σ^2 are unknown. Consider the problem of testing

$$H_0 : \mu = 0$$
$$H_1 : \mu \neq 0$$

Show that the generalized likelihood ratio test statistic for this problem simplifies to

$$\Lambda(X_1, \ldots, X_n) = \left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2 + n\bar{X}^2} \right)^{n/2}.$$

Letting $S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ and $T = \sqrt{n}\bar{X}/S_X$ (the usual one-sample t-statistic for this problem), show that $\Lambda(X_1, \ldots, X_n)$ is a monotonically decreasing function of $|T|$, and hence the generalized likelihood ratio test is equivalent to the two-sided t-test which rejects for large values of $|T|$.

4. **Migration rates.** To study the rates of migration between 3 cities A, B, and C, 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 A, N_{12} be the number of people who started in city A and ended in city B, etc. Let us model the vector of 9 counts $(N_{ij})_{1 \leq i,j \leq 3}$ as $\text{Multinomial}(n, (p_{ij})_{1 \leq i,j \leq 3})$, 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.