14.381: Statistics

Problem Set 4

Problem 3 is required.

1. Suppose we have a response variable s that can take k possible values - for convenience labeled 1,2,... k with probability distribution

$$P\{s = i\} = \theta_i, \quad \sum_{i=1}^{k} \theta_i = 1$$

Assume that we have a sample $s_1, ..., s_N$ of such responses.

   (a) Find a minimal sufficient statistics.

   (b) Let X_i stay for the number of i's observed in the sample (we refer to X_i's as counts form from the sample). What is the distribution of vector $(X_1, ..., X_k)$?

   (c) Find the MLE of $(\theta_1, ..., \theta_k)$.

   (d) Since the dimensionality of the parameter space is $k-1$, let us re-parameterize the problem by considering as a new parameter $\widetilde{\theta}$ only first $k-1$ components: $\widetilde{\theta} = (\theta_1, ..., \theta_{k-1})$. (*Remember that $\theta_k = 1 - \sum_{1}^{k-1} \theta_i$.*) Find out the Fisher information matrix for $\widetilde{\theta}$.

   (e) What is the limit behavior of the MLE estimator in this case? Take a simpler case when $k = 3$ and calculate the asymptotic variance of $\xi_i = \sqrt{n}(\widehat{\theta}_{i,ML} - \theta_{i,0})$, and the asymptotic covariance of ξ_1 and ξ_2. Could you explain the sign of the covariance?

2. Let $X_1, \ldots, X_n$ be iid Poisson (λ).

   (a) Find the UMP test for $H_0 : \lambda \leq \lambda_0$ vs. $H_1 : \lambda > \lambda_0$

(b) Consider the specific case $H_0 : \lambda \leq 1$ vs. $H_1 : \lambda > 1$. Determine the sample size n so that the UMP satisfies two conditions:

$$P_{\lambda=1}(\text{reject } H_0) \approx 0.05$$

$$P_{\lambda=2}(\text{reject } H_0) \approx 0.9$$

Here "$\approx$" stays for " approximately equal". Please, use the CLT as approximation device.

3. *(Required problem)* There is a theory that people can postpone their death until after an important event. To test the theory, Phillips and King (1988) collected data on deaths around the Jewish holiday Passover. Of 1919 deaths, 922 died the week before the holiday and 997 dies the week after. Think of this as a binomial and test the null hypothesis that $\theta = 1/2$. Report and interpret the p-value. There are a number of different test you may suggest, any would be fine.

4. Suppose that we have two independent samples: $X_1, \ldots, X_n$ are iid exponential(θ) and $Y_1, \ldots, Y_m$ are iid exponential(μ). Both θ and μ are unknown. We want to test $H_0 : \mu = \theta$ vs $H_1 : \mu \neq \theta$. The goal of this problem is to write down a LR test statistic.

   (a) Write down the likelihood function and find the unrestricted ML estimates of μ and θ

   (b) Find the restricted ML estimate (via imposing the null)

   (c) Write down LR test statistic

   (d) Show that it's a function of test statistic $\frac{\sum_i X_i}{\sum_i X_i + \sum_j Y_j}$