14.381: Statistics

Problem Set 4

due Tuesday, October 8, 2019

Problem 3 is required.

1. Suppose we have an i.i.d. sample (Y_i, X_i) drawn from a model

$$Y_i = X_i'\beta + e_i$$

where $e_i \sim N(0, \sigma^2)$ and independent of X_i. Assume that X_i is a $k \times 1$ vector and its distribution does not depend on β or σ^2 (take it as a known function $f_X(\cdot)$ and assume it has four finite moments). Let β and σ^2 be two unknown parameters ($k \times 1$ and a scalar correspondingly).

(a) Write down the log-likelihood function.

(b) Find the MLEs for both β and σ^2.

(c) Calculate the Fisher information matrix and check that the second informational equality holds.

(d) What is the limit behavior of the MLE estimator in this case?

Now imagine that some of the distributional assumptions we made above are incorrect - we are now working with a pseudo-likelihood. The true data comes from a data generating process (DGP) in which assumptions that $e_i \sim N(0, \sigma^2)$ and e_i is independent of X_i are no longer true. Instead, the conditions $E[e_i X_i] = 0$, $E[e^2] = \sigma^2$ and $E[e^4] < \infty$ hold, but aside from this e_i may depend on X_i, which has some distribution with four finite moments. Denote this joint distribution as Q.

(e) Calculate the covariance matrix of the score of the likelihood derived in (a) at the true β, σ^2.

(f) Calculate the expectation of the negative second derivative of the likelihood derived in (a) at the true β, σ^2. Does the second information equality still hold?

(g) What can we say about the asymptotic behavior of the estimator of β derived in (b). In particular, what is its asymptotic variance?

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.