Problem Set 1

Applied Statistics and Econometrics II
Spring 2018, NYU
Ercan Karadas

(Due: February 8, in class)

[1] Suppose $X_n \sim b(n,p)$ (X_n is the number of successes throughout n independent repetitions of a random experiment with the success probability p). The ratio X_n/p is called the relative frequency of success. Show that

$$\text{plim}\, \frac{X_n}{n} \xrightarrow{p} p$$

a) By using Chebyshev's inequality

b) By using WLL

[2] Let $X_1, X_2, \ldots, X_n$ be iid random variables with common pdf[1]

$$f(x) = \begin{cases} e^{-(x-\theta)} & \theta \leq x,\ -\infty < \theta \leq \infty \\ 0 & \text{otherwise} \end{cases}$$

Suppose that lower bound of the distribution θ is unknown. In this exercise, we are going to try to obtain a good estimator for θ.

a) I argue that the minimum of the sample

$$Y_n = \min\{X_1, \ldots, X_n\}$$

is a good estimator for θ when n is large. Why?

b) Find the CDF of Y_n.

c) Using the CDF of Y_n show that it is a consistent estimator of θ:

$$Y_n \xrightarrow{p} \theta$$

d) Using part (b) find pdf of Y_n and then show that

$$E\left(Y_n\right) \neq \theta$$

Thus, conclude that Y_n is a biased estimator of θ.

e) Obtain an unbiased estimator of θ.

[3] Let $X_1, X_2, \ldots$ be i.i.d. with $E(X_i) = \mu$, $V(X_i) = \sigma^2$, and $V(X_i^2) = \nu^2 < \infty$, and define

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

[1] This pdf is called the *shifted exponential* distribution.

a) Show that

$$S_n^2 \xrightarrow{p} \sigma^2$$

b) Show that

$$\sqrt{n}\left(S_n^2 - \sigma^2\right) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2 - \sigma^2\right)$$

c) Using the results in part (a) and (b) show that

$$\sqrt{n}\left(S_n^2 - \sigma^2\right) \xrightarrow{d} N(0, \nu^2)$$

[4] a) Let $X_1, X_2, \ldots, X_n$ be iid random variables from Bernoulli distribution with the parameter p, i.e. $X_i \sim \text{Ber}(p)$. Show that

$$\sqrt{n}\left(\bar{X}_n - p\right) \xrightarrow{d} N\left(0, p(1-p)\right)$$

b) Now suppose that $X_n \sim \text{Bin}(n, p)$. Show that

$$\sqrt{n}\left(\frac{X_n}{n} - p\right) \xrightarrow{d} N\left(0, p(1-p)\right)$$

c) Now suppose that $X_n \sim \text{Bin}(k_n, p)$, where $k_n \to \infty$ as $n \to \infty$. Show that

$$X_n \overset{a}{\sim} N\left(k_n p, k_n p(1-p)\right) \qquad (\overset{a}{\sim}\text{: approx. distributed})$$

[5] For estimating p^2, suppose we have the choice between the following two:

(i) Sample n binomial trials with probability p^2 of success and use X/n as an estimator for p^2, where X is the number of successes in n trials.

(ii) Sample n binomial trials with probability p of success and use $(Y/n)^2$ as an estimator for p^2, where Y is the number of successes in n trials.

a) Show that

$$\sqrt{n}\left(\frac{X}{n} - p^2\right) \xrightarrow{d} N\left(0, p^2(1-p^2)\right)$$

b) Show that

$$\sqrt{n}\left(\left(\frac{Y}{n}\right)^2 - p^2\right) \xrightarrow{d} N\left(0, p(1-p)\cdot 4p^2\right)$$

c) Discuss why (X/n) is a better estimator when $p > 1/3$.

[6] Let $X_n \sim \text{Poisson}(n\lambda)$ where the positive integer n is large and $\lambda > 0$.

a) Find the limiting distribution of

$$\sqrt{n}\left(\frac{X_n}{n} - \lambda\right) \xrightarrow{d} \ ?$$

b) Find the limiting distribution of

$$\sqrt{n}\left(\sqrt{\frac{X_n}{n}} - \sqrt{\lambda}\right) \xrightarrow{d} \; ?$$

[7] In the lecture we proved the Delta-Method under the assumption that g is continuous and $g'(\theta) \neq 0$. Let us now try to obtain a similar result when $g'(\theta) = 0$ but $g''(\theta) \neq 0$:. In this exercise, we are going to show the following:

If

$$\sqrt{n}\,(X_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$

and g is continuous, $g'(\theta) = 0$, $g''(\theta) \neq 0$. Then

$$n\,[g(X_n) - g(\theta)] \xrightarrow{d} \frac{1}{2}\sigma^2 g''(\theta)\chi_1^2$$

a) Carry out a second-order Taylor expansion of g about θ and using the assumptions above show that it can be written as

$$g(X_n) - g(\theta) = \frac{1}{2} g''(\theta)\,(X_n - \theta)^2 + \text{Remainder Term}$$

b) At this step we multiplied both sides by $\sqrt{n}$ to prove the delta method in the lecture, but note that it would not work here. Argue that we should instead multiply both sides by n in order to be able to use the main hypothesis above. Finally, using the fact that $\chi_1^2 = Z^2$, where $Z \sim N(0, 1)$, complete the proof.

[8] Suppose that $X_n \sim \text{Bin}(n, p)$ where the positive integer n is large and $0 < p < 1$. Let $g(x) = x^3 - x$. Find the limiting distribution of

$$n\left[g\left(\frac{X_n}{n}\right) - c\right]$$

for an appropriate constant c when $p = 1/\sqrt{3}$ (Hint: Use the previous exercise).