ECMA 31000: Problem Set 3

Joe Hardwick

Due October 21 by 11:59PM

Question 1 Show that if X_n is a sequence of random variables such that for some $r > 0$ and some random variable Y,

$$n^r (X_n - X) \xrightarrow{d} Y,$$

then $X_n \xrightarrow{p} X$.

Question 2 a) We stated in class that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$ does not imply

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Prove that, if $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ are independent sequences with $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, and X, Y are independent random variables, then in fact

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

b) Use this result to show that if $\{X_i\}_{i\geq 1}$ and $\{Y_i\}_{i\geq 1}$ are iid sequences that are independent of each other, $\mathrm{E}(X_i) = \mu_X$, $\mathrm{E}(Y_i) = \mu_Y$, and $Var(X_i) = Var(Y_i) = 1$, then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \begin{pmatrix} X_i - \mu_X \\ Y_i - \mu_Y \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Find constants $r \geq 0$ and c such that

$$n^r \left(\frac{\bar{X}_n}{\bar{Y}_n} - c \right)$$

has a non-degenerate limiting distribution in the following two cases: (i) $\mu_X \in \mathbb{R}$, $\mu_Y \neq 0$, and (ii) $(\mu_X, \mu_Y) = (0, 0)$.

Question 3 a) Show that if $X_n \sim Binomial(n, p)$ then

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} \mathcal{N}(0, 1).$$

b) Show that if $X_n \sim \chi^2_n$ then

$$\frac{X_n - n}{\sqrt{2n}} \xrightarrow{d} \mathcal{N}(0,1).$$

Hint: Find a way to use the CLT.

Question 4 Let X_n be a $(K \times 1)$ sequence of random vectors converging in distribution to a random vector X.

a) Show that for any component $X_{n,i}$ of X_n, $X_{n,i} \xrightarrow{d} X_i$, where X_i is the $i-$th component of X.

b) Suppose

$$X_n \xrightarrow{d} \mathcal{N}(0,\Sigma)$$

for some non-singular variance matrix Σ. Show that if A_n is a sequence of $(K \times K)$ matrices such that $A_n \xrightarrow{p} A$ for a constant matrix A satisfying $A\Sigma A' = I_K$, where I_K is the $(K \times K)$ identity matrix, then

$$\|A_n X_n\|^2 \xrightarrow{d} \chi^2_K,$$

where χ^2_K denotes a chi-square distribution with k degrees of freedom, and $\|\cdot\|$ denotes the euclidean norm on $\mathbb{R}^K$.

(Hint: If $Y_1, Y_2, \ldots, Y_K$ are independent standard normal random variables, $\sum_{i=1}^K Y_i^2 \sim \chi^2_K$).

Question 5 (Stochastic Order Relations) a) Prove that $X_n = o_p(1) \implies X_n = O_p(1)$.

b) Prove that $o_p(1) + o_p(1) = o_p(1)$.

c) Prove that $o_p(1) + O_p(1) = O_p(1)$.

Question 6 Let $\{X_i\}_{i \geq 1}$ be an iid sequence with finite fourth moments and $\mathrm{E}(X_i) = \mu$. Find constants a, b and $r > 0$ such that

$$n^r \begin{pmatrix} \left(\bar{X}_n - \mu\right) - a \\ \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - b \end{pmatrix}$$

has an asymptotically normal distribution. Is it always true that the limit distribution is non-degenerate?

Question 7 (Computational Question) Let $\{X_i\}_{i \geq 1}$ be an iid sequence such that $X_i \sim \mathit{Bernoulli}(p)$. That is:

$$X_i = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1 - p. \end{cases}$$

a) Show that

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

and that

$$\mathrm{P}\left(\bar{X}_n - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n\left(1-\bar{X}_n\right)}{n}} < p < \bar{X}_n + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n\left(1-\bar{X}_n\right)}{n}}\right) \to 1-\alpha,$$

where $z_{1-\frac{\alpha}{2}}$ is the $1-\alpha$ quantile of the normal distribution.

b) In problem set 2, we constructed a confidence interval of width 2ϵ which contained p with probability at least $1 - \frac{1}{4n\epsilon^2}$. We now consider the confidence interval

$$\left[\bar{X}_n - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n\left(1-\bar{X}_n\right)}{n}}, \bar{X}_n + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}_n\left(1-\bar{X}_n\right)}{n}}\right]$$

based on the normal approximation in part a). How large must n be to ensure the width of this confidence interval is at most 0.2, when $1-\alpha = 0.95$?

c) Simulate n iid draws from this distribution with $p = 0.4$, for each of $n = 25, 50, 100$. Compute the confidence intervals for each n based on your simulated data. Does the true value of p lie inside the confidence interval? Repeat this exercise 1000 times for each value of n, (though you don't need to display the results of each replication). For each value n, report the proportion of your replications for which the true value of p lies in your confidence interval. Does this vary much with n? Why? What is the advantage of using this confidence interval instead of one where the width is selected based on Chebyshev's inequality?

d) In this part we will show that asymptotic confidence intervals do not guarantee finite sample coverage probabilities. Repeat part c) but with $p = 0.99$, and for sample sizes $n = 25, 50, 100, 250, 500, 1000, 2000$. Does the asymptotic 95% confidence interval contain the true parameter with approximately the right probability? Why is this happening? Interpret your simulations in light of the result derived in part a).