Cloud Publications
International Journal of Advanced Mathematics and Statistics
2016, Volume 2, Issue 1, pp. 64-67, Article ID Sci-510
ISSN 2348-5175

Methodology Article **Open Access**

The Multivariate Central Limit Theorem and its Relationship with Univariate Statistics

Athanase Polymenis

Department of Economics, University of Patras, University Campus at Rio, 26504 Rio-Patras, Greece

Publication Date: 15 October 2016

Article Link: http://scientific.cloud-journals.com/index.php/IJAMS/article/view/Sci-510

Abstract In the present article the multivariate central limit theorem is revisited. Rather than simply reviewing existing methodology our approach mostly aims at giving particular emphasis on some univariate techniques that support the proof of this theorem. On those grounds all necessary mathematical arguments are duly provided.
Keywords *Characteristic Function; Convergence in Distribution; Normal Distribution*

1. Introduction

The well-known central limit theorem for independent and identically distributed vector random variables is a very important result since it allows for an approximate normal distribution for the mean (and, equivalently, for the sum) of a sequence of a very large number of variables which satisfy the aforementioned properties; thus it has been extensively mentioned and used in existing literature. The idea of the present article emanates from Theorem 29.4 of [2, p. 383], and a subsequent comment, namely that "certain limit theorems can be reduced in a routine way to the one-dimensional case". Based on that comment our intention is therefore to revisit the proof of this theorem in view of providing a detailed description of the way statistical theory that is used in the one-dimensional case (mentioned in the text as the univariate case) applies in a straightforward manner in order to deal with a multivariate problem.

2. The Multivariate Central Limit Theorem

The multivariate central limit theorem we will focus on is called Theorem 1 and is an early result which is very famous and well documented. For instance, it is found as Theorem 3.4.3 in [1, p. 81]. We first state this result.

Theorem 1

Let $X_1, X_2, \ldots$ *be a sequence of p-dimensional independent and identically distributed random vectors with finite mean vector* $E(X_j)=\mu$ *and finite variance-covariance matrix* $E(X_j-\mu)(X_j-\mu)'=\Sigma$. *Then the asymptotic distribution of* $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j-\mu)$ *converges towards a normal* $\mathrm{N}(\underset{\sim}{0},\Sigma)$ *as* $n\to\infty$.

Remark that in the univariate case, i.e. the case where *p*=1, Theorem 1 is just the univariate central limit theorem also called the Lindeberg-Lévy theorem (see for example [3, p. 215]). The proof of the multivariate central limit theorem, which we provide in the sequel, will be essentially based on techniques appearing in the proof of Lindeberg-Lévy's theorem.

Proof of Theorem 1

Like in Theorem 3.4.3 of [1, p. 81], we use $\phi(t,s)=[E(e^{is\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j-E(tXj)]})]$ as characteristic function of $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j-E(tX_j)]$, for real *s*, fixed $t=(t_1,\ldots,t_p)$, and $i^2=-1$. Also letting $X=\sum_{j=1}^{n}tX_j$, with $E(tX_j)=m$ and $Var(tX_j)=\sigma^2$ we obtain $E(X)=m_X=E(\sum_{j=1}^{n}tX_j)=nm$, and $Var(X)=\sigma_X^2=Var(\sum_{j=1}^{n}tX_j)=n\sigma^2$ (or in other words $\sigma_X=\sigma\sqrt{n}$), since the X_j's are independently distributed. We now adopt the same rationale as that used by [3, p. 215] for the proof of the Lindeberg-Lévy theorem. Let $\hat{\phi}(u)$ be the characteristic function of tX_j-m and $\tilde{\phi}(u)$ the characteristic function of $X-m_X$. We have

$\tilde{\phi}(u)=E(e^{iu(X-m)})=E(e^{iu[(tX_1-m)+\ldots+(tX_n-m)]})=e^{-iunm}E(e^{iutX_1}\ldots e^{iutX_n})$,

and since the X_j's are independently distributed we obtain $\tilde{\phi}(u)=e^{-ium}E(e^{iutX_1})\ldots e^{-ium}E(e^{iutX_n})=\hat{\phi}(u)\ldots\hat{\phi}(u)=[\hat{\phi}(u)]^n$ (since $\hat{\phi}(u)=e^{-ium}E(e^{iutX_j})$, for any $j=1,\ldots n$). Then, if $\phi(u)$ is the characteristic function of $(X-m_X)/\sigma_X$, we obtain $\phi(u)=E(e^{iu(\frac{X-m_X}{\sigma_x})})=E(e^{i\frac{u}{\sigma_X}(X-m_X)})=\tilde{\phi}(\frac{u}{\sigma_X})$, with $\tilde{\phi}(\frac{u}{\sigma_X})=[\hat{\phi}(\frac{u}{\sigma_X})]^n$ (since $\tilde{\phi}(u)=[\hat{\phi}(u)]^n$).

Thus $\phi(u)=[\hat{\phi}(\frac{u}{\sigma\sqrt{n}})]^n$.

On the other hand $\hat{\phi}(u)=\int_{-\infty}^{\infty}e^{iuy}f(y)dy$, where $Y_j=tX_j-E(tX_j)$, and by second order Taylor's expansion around 0 we obtain

$$\hat{\phi}(u)=\hat{\phi}(0)+(u-0)\hat{\phi}'(0)+[(u-0)^2/2]\hat{\phi}''(0)+o(u-0)^2=\hat{\phi}(0)+u\hat{\phi}'(0)+(u^2/2)\hat{\phi}''(0)+o(u^2)$$

where $\hat{\phi}'(u) = i\int_{-\infty}^{\infty} ye^{iuy} f(y)dy$ and $\hat{\phi}''(u) = i^2\int_{-\infty}^{\infty} y^2e^{iuy} f(y)dy = -\int_{-\infty}^{\infty} y^2e^{iuy} f(y)dy$.

Hence we have $\hat{\phi}(0) = \int_{-\infty}^{\infty} f(y)dy = 1$, $\hat{\phi}'(0) = i\int_{-\infty}^{\infty} yf(y)dy = iE(Y_j) = 0$, and

$\hat{\phi}''(0) = -\int_{-\infty}^{\infty} y^2 f(y)dy = -E(Y_j^2)$, and so we obtain $\hat{\phi}(u) = 1-[(u)^2/2]E(Y_j^2)+o(u^2)$.

Finally, since $E(Y_j^2) = Var(tX_j) = \sigma^2$, we obtain $\hat{\phi}(u) = 1-(\sigma^2 u^2/2)+o(u^2)$.

Hence $\phi(u) = [\hat{\phi}(\frac{u}{\sigma\sqrt{n}})]^n = [1-(\sigma^2 u^2/2n\sigma^2)+\eta(n,u)/n]^n = (1-u^2/2n+\eta(n,u)/n)^n$

where, for every fixed u, $\eta(n,u) \to 0$ as $n \to \infty$.

We now let $x = -u^2/2n+\eta(n,u)/n$; then $\lim_{n\to\infty}(x) = 0$ and

$\lim_{n\to\infty}(1-u^2/2n+\eta(n,u)/n)^n = \lim_{x\to 0}(1+x)^{-u^2/2x}\lim_{x\to 0}(1+x)^{\eta(n,u)/x}$

$= \{\lim_{x\to 0}[(1+x)^{1/x}]^{-u^2/2}\}\{\lim_{x\to 0}[(1+x)^{1/x}]^{\eta(n,u)}\}$.

Since a well known result is that $\lim_{x\to 0}[(1+x)^{1/x}] = e$, and on the other hand $n\to\infty$ as $x\to 0$ imply that $\eta(n,u) \to 0$ as $x\to 0$, for fixed u, we obtain $\lim_{n\to\infty}(1-u^2/2n+\eta(n,u)/n)^n = e^{-u^2/2}$, and thus $\lim_{n\to\infty}\phi(u) = e^{-u^2/2}$, for any u, $e^{-u^2/2}$ being continuous at $u=0$, with $\lim_{n\to\infty}\phi(0) = 1$, and so we deduce that ϕ is itself a characteristic function (in other words we are not in presence of a pathological situation such as that described in exercise 5.12.35 and solution of [5, p. 266]). It results from these arguments that $\phi(u)$ is indeed the characteristic function of $(X-m_X)/\sigma_X$. On the other hand $e^{-u^2/2}$ is the characteristic function of a standard normal random variable and hence, by the continuity theorem for characteristic functions concerning the univariate case (see for instance [4, p. 190]), $(X-m_X)/\sigma_X$ converges in distribution towards a N(0,1) random variable as $n\to\infty$. This is equivalent to saying that $(X-nm)/\sqrt{n}$ converges in distribution towards N(0, σ^2), with $\sigma^2 = Var(tX_j) = t\Sigma t'$ (where prime denotes transpose), and thus $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j - E(tX_j)]$ converges in distribution towards a N(0, $t\Sigma t'$) random variable (see also proof of Theorem 3.4.3 of [1, p. 81]). The characteristic function of the latter variable is as given by [4, p. 187], Example (5), and reduces to $e^{\frac{-1}{2}s^2 t\Sigma t'}$ for our case since the mean is 0. Thus we have that for each $(t_1,...,t_p)$

$$\lim_{n\to\infty}\phi(t,s) = \lim_{n\to\infty}E(e^{is\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j-E(tXj)]}) = e^{\frac{-1}{2}s^2 t\Sigma t'}.$$

Like in [2, p. 383] or in [1, p. 81], we take $s=1$ in order to obtain

$$\lim_{n\to\infty}\phi(t,1) = \lim_{n\to\infty}E(e^{i\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j-E(tXj)]}) = e^{\frac{-1}{2}t\Sigma t'}.$$

Note that $\phi(t,1) = E(e^{i\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j - E(tXj)]}) = E(e^{it\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j-\mu)})$ is the characteristic function of $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j-\mu)$. On the other hand $\lim_{n\to\infty}\phi(t,1) = e^{\frac{-1}{2}t\Sigma t'}$ is continuous at $t = (t_1,...,t_p) = (0,...,0) = \underset{\sim}{0}$, with $\lim_{n\to\infty}\phi(\underset{\sim}{0},1) = 1$, and so by means of the continuity theorem for characteristic functions that concern the multivariate case (see Theorem 2.6.4 of [1, pp. 48-49]) we have that $\lim_{n\to\infty}\phi(t,1)$ is identical with the characteristic function $e^{\frac{-1}{2}t\Sigma t'}$ of a normally distributed vector random variable N($\underset{\sim}{0}$,Σ) (see [5, p. 187]), Example (6), or Theorem 2.6.1 of [1, p. 45], for documentation on the characteristic function $e^{\frac{-1}{2}t\Sigma t'}$). We thus conclude that $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j-\mu)$ converges in distribution towards a normally distributed N($\underset{\sim}{0}$,Σ) vector random variable, and hence Theorem 1 is proved.

3. Conclusion

The above arguments underline the obvious implication (and thus the crucial role) of univariate statistical techniques for proving multivariate results. These techniques arise from the use of the random variable $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j - E(tX_j)]$ and its characteristic function $\phi(t,s)$ in the proof of Theorem 1. Remark that they are also involved in the proof of Lindeberg-Lévy's theorem. Finally it is noteworthy to stress that the idea of using the variable $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[tX_j - E(tX_j)]$ and its characteristic function in the preceding section occurs naturally in view of Theorem 29.4 of [2, p. 383].

References

[1] Anderson, T.W., 1984: *An introduction to multivariate statistical analysis*. 2nd Ed., Wiley, New-York. 667.

[2] Billingsley, P., 1995: *Probability and measure*. 3rd Ed., Wiley, New-York. 593.

[3] Cramer, H., 1946: *Mathematical methods of statistics*. Princeton University Press. 575.

[4] Grimmett, G., and Stirzaker, D., 2001: *Probability and random processes*. 3rd Ed., Oxford University Press. 596.

[5] Grimmett, G., and Stirzaker, D., 2001: *One thousand exercises in probability*. Oxford University Press. 438.