KURTOSIS OF THE SAMPLE MEAN AND THE SAMPLE CENTROID AND ITS LIMITING VALUES

Ampalavanar Nanthakumar
Department of Mathematics, State University of New York at Oswego
Email: ampala.nanthakumar@oswego.edu

ABSTRACT

In this paper, we study the limiting value of the Kurtosis of the sample mean and the sample centroid for large samples under some mild conditions in a very general situation. The results are seen to support the Central Limit Theorem in both the univariate and the multivariate situations.

KEY WORDS

Kurtosis, Multivariate, Limiting Value, Central Limit Theorem.

1. INTRODUCTION

The Central Limit Theorem is a well-known result in probability and statistics and several versions of it are in existence since the time the proof was given for the original version which involved identically and independently distributed random variables with finite variance. Polya (1920), Lindberg (1922, 1924), Levy (1923, 1935, 1939), Lyapunov (1900, 1901), Feller (1971), Gnedenko and Kolmogorov (1954), Trotter (1959), Govindarajulu (1973) and many others have discussed the Central Limit Theorem and its different variants during the last century. The work related to this important theorem which plays a central role in probability and statistics may have started as early as 1733 as there are references about approximation theorems in de Moivre (1733) and Laplace (1774, 1812). On the other hand, the results concerning the kurtosis for some univariate distributions are well known and there are many papers in this respect. Mardia (1970, 1974) extended the definition of the kurtosis for the multivariate situation (also see the paper by Miyagawa et al. (2011) for further reading). In this paper, we study the limiting behavior of the kurtosis of the sample mean in a univariate situation where the observations are independent but not identically distributed. Also, we consider the limiting value of the kurtosis of the sample centroid in a bivariate situation where the observations are from a bivariate population but the population is not necessarily bivariate normally distributed. In the univariate case, we show that the limiting value of the kurtosis of the sample mean is 3 and in the bivariate case, the limiting value of the kurtosis of the sample centroid is 8.¹

¹ AMS Subject Classification 62
Kurtosis of the Sample Mean and the Sample Centroid and its Limiting Values

This supports the Central Limit Theorem which relies on the fact that the kurtosis is 3 for the univariate normal distribution, and is 8 for the bivariate normal distribution. This kurtosis based approach is a useful tool in the context of checking for normality and it is fairly easy to calculate the kurtosis even in the multivariate situations to check for the multivariate normality.

First, the paper discusses the univariate situation where the observations are independent but not necessarily identically distributed. Under a very mild condition, we are able to show that the limiting value of the kurtosis of the sample mean is 3 as the sample size gets larger. This finding is in agreement with the Central Limit Theorem. The case where the observations are identically and independently distributed is treated as a special case and again, we are able to verify normality. Next, the paper discusses the bivariate case where the vector components (attributes) follow the non-normal distributions and the observations about the attributes are independent.

2. THEORY AND METHODOLOGY

Univariate Case:

Suppose that you are taking one observation from each of the different layers of a population where the $i^{th}$ layer has a mean $= \mu_i$ and variance $= \sigma_i^2$. We assume that the sample based observations are independent but not identically distributed due to the fact that these observations are about different layers of the population. Also, note that these individual layers can follow different distributions. This means that for observation $X_i$ from the $i^{th}$ layer, $E(X_i) = \mu_i$ and $Var(X_i) = \sigma_i^2$.

Main Results:

This section presents the mathematical results for the large samples. The focus is on the sample mean $\bar{X}$ based on these observations.

Lemma 1:

If $\text{Sup } E(X_i^4) < \infty$, then the kurtosis of $\bar{X}_n$ is 3 as $n \to \infty$.

Proof:

Note that $\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$

This means,

$$\bar{X}^4 = \frac{1}{n^4}. \left\{ \sum_{i=1}^{n} X_i^4 + 6 \sum_{i<j} X_i^2 X_j^2 + 4 \sum_{i \neq j} X_i^3 X_j + 12 \sum_{i<j<k} X_i^2 X_j X_k \right\}$$

$$+ \frac{1}{n^4}. \left\{ \sum_{i<j<k<l} \sum \sum X_i X_j X_k X_l \right\}$$

Let us assume without loss of generality $E(X_i) = 0$. Then,
\[ E(\overline{X}^4) = \frac{1}{n^4} \left\{ \sum_{i=1}^{n} E(X_i^4) + 6 \sum_{i<j} E(X_i^2)E(X_j^2) \right\} \]

\[ = \frac{1}{n^4} \left\{ \sum_{i=1}^{n} E(X_i^4) + 6 \sum_{i<j} \sigma_i^2 \sigma_j^2 \right\} \]

Note that

\[ Var(\overline{X}_n) = \frac{1}{n} \left\{ \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 \right\} \]

So,

\[ \left[ Var(\overline{X}_n) \right]^2 = \frac{1}{n^4} \left\{ \sum_{i=1}^{n} \sigma_i^4 + 2 \sum_{i<j} \sigma_i^2 \sigma_j^2 \right\} \]

\[ \geq \frac{2}{n^4} \sum_{i<j} \sigma_i^2 \sigma_j^2 \]

By definition, kurtosis \( \kappa \) is equal to

\[ \kappa = \frac{E(\overline{X}_n^4)}{\left[ Var(\overline{X}_n) \right]^2} \]

\[ \leq \frac{1}{n^4} \left\{ \sum_{i=1}^{n} E(X_i^4) + 6 \sum_{i<j} \sigma_i^2 \sigma_j^2 \right\} \]

\[ \leq \frac{2}{n^4} \sum_{i<j} \sigma_i^2 \sigma_j^2 \]

\[ \rightarrow 3 \text{ as } n \rightarrow \infty . \]

Also,

\[ \frac{6 \sum_{i<j} \sigma_i^2 \sigma_j^2}{\sum_{i=1}^{n} \sigma_i^4 + 2 \sum_{i<j} \sigma_i^2 \sigma_j^2} \leq 3 \]

\[ \rightarrow 3 \text{ as } n \rightarrow \infty . \]

So, Limit \( \kappa = 3 \) as \( n \rightarrow \infty . \)

**Multivariate Case:**

Suppose that we are taking bivariate observations \( \begin{pmatrix} x_i \\ y_i \end{pmatrix} \) about a population where the marginal distributions are non-normal. Say that without loss of generality that the observations \( x_1, x_2, \ldots, x_n \) are identically and independently distributed
with mean $= 0$ and variance $= \sigma_1^2$. Similarly, the observations $y_1, y_2, \ldots, y_n$ are identically and independently distributed with mean $= 0$ and variance $= \sigma_2^2$. Furthermore, assume that $E(y_i \setminus x_i) = \frac{\sigma_{12}}{\sigma_1^2} x_i$ and $E(y_i^2 \setminus x_i) = \sigma_2^2 \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{12}^2}{\sigma_1^2} x_i^2$, where $\sigma_{12} = \text{Cov}(x_i, y_i)$ and $\rho$ is the correlation coefficient such that $\sigma_{12} = \rho \sigma_1 \sigma_2$.

Note that $\text{Cov}(x_i, y_j) = 0$ when $i \neq j$. In order to proceed, we will use the Multivariate Kurtosis as defined by Mardia. According to the definition given by Mardia (1970, 1974), for a bivariate vector $\begin{pmatrix} x \\ y \end{pmatrix}$, its kurtosis is defined as the expected value for the squared Mahalanobis distance from the centroid of the population. In other words,

Multivariate Kurtosis $= E \left( \left( x - \mu_x, y - \mu_y \right) \cdot \Sigma^{-1} \cdot \left( x - \mu_x, y - \mu_y \right) \right)^2$

where $\Sigma$ is the variance-covariance matrix, and $\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$ is the centroid of the bivariate population. In order to derive the kurtosis, we need the following results. Here, as noted earlier, we assume without loss of generality that $\mu_x = 0$ and $\mu_y = 0$.

Let $A_n = \left( \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \right) \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i \end{pmatrix}$

where $\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$ is the inverse of $\begin{pmatrix} n\sigma_1^2 & n\sigma_{12} \\ n\sigma_{12} & n\sigma_2^2 \end{pmatrix}$.

Note that one can write

$A_n = a_{11} \left( \sum_{i=1}^{n} x_i \right)^2 + 2a_{12} \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{i=1}^{n} y_i \right) + a_{22} \left( \sum_{i=1}^{n} y_i \right)^2$

This implies

$A_n^2 = a_{11}^2 \left( \sum_{i=1}^{n} x_i \right)^4 + a_{22}^2 \left( \sum_{i=1}^{n} y_i \right)^4 + 4a_{12}^2 \left( \sum_{i=1}^{n} x_i \right)^2 \left( \sum_{i=1}^{n} y_i \right)^2 + 4a_{11}a_{12} \left( \sum_{i=1}^{n} x_i \right)^3 \left( \sum_{i=1}^{n} y_i \right) + 4a_{22}a_{12} \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{i=1}^{n} y_i \right)^3$
\[= n^4a_{11}^2(X)^4 + n^4a_{22}^2(Y)^4 + n^4(4a_{12}^2 + 2a_{11}a_{22})(X^2)(Y)^2\]
\[+ 4n^4a_{11}a_{12}(X)^3Y + 4n^4a_{22}a_{12}X(Y)^3\]

We need the following lemmas in order to show that in the bivariate case, 
\[E\left(A_n^2\right) \rightarrow 8 \text{ as } n \rightarrow \infty.\]

**Lemma 2:**
\[E\left(\sum_{i=1}^{n} y_i \setminus \sum_{i=1}^{n} x_i\right) = \frac{\sigma_{12}}{\sigma_1^2}\left(\sum_{i=1}^{n} x_i\right)\]

**Proof:**
\[E\left(\sum_{i=1}^{n} y_i \setminus \sum_{i=1}^{n} x_i\right) = E\left(y_1 \setminus \sum_{i=1}^{n} x_i\right) + E\left(y_2 \setminus \sum_{i=1}^{n} x_i\right) + \ldots + E\left(y_n \setminus \sum_{i=1}^{n} x_i\right)\]
\[= E\left(y_1 \setminus x_1\right) + E\left(y_2 \setminus x_2\right) + \ldots + E\left(y_n \setminus x_n\right)\]
\[= \frac{\sigma_{12}}{\sigma_1^2}x_1 + \frac{\sigma_{12}}{\sigma_1^2}x_2 + \ldots + \frac{\sigma_{12}}{\sigma_1^2}x_n\]
\[= \frac{\sigma_{12}}{\sigma_1^2}\left(\sum_{i=1}^{n} x_i\right)\]

**Corollary 1:**
\[E\left(\bar{Y} \setminus \bar{X}\right) = \frac{\sigma_{12}}{\sigma_1^2}\bar{X}\]

**Lemma 3:**
\[Cov(\bar{X}, \bar{X}) = \frac{\sigma_{12}}{n}\]

**Proof:**
From the above corollary, it follows that

\[E\left(\bar{X} \cdot (\bar{Y} \setminus \bar{X})\right) = \frac{\sigma_{12}}{\sigma_1^2}\left(\bar{X}\right)^2\]

\[E\left(\bar{X} \cdot (\bar{Y})\right) = E\left(E\left(\bar{X} \cdot (\bar{Y} \setminus \bar{X})\right)\right)\]

\[= E\left(\frac{\sigma_{12}^2}{\sigma_1^2}\left(\bar{X}\right)^2\right) = \frac{\sigma_{12}^2}{\sigma_1^2}E\left(\left(\bar{X}\right)^2\right) = \frac{\sigma_{12}^2}{\sigma_1^2} \cdot \frac{\sigma_1^2}{n} = \frac{\sigma_{12}}{n}\]
Lemma 4:

\[ E\left( (\bar{Y})^2 \setminus \bar{X} \right) = \frac{1}{n^2} \cdot \left\{ n \cdot \sigma_Y^2 \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_Y^2}{\sigma_{X}^2} \cdot \sum \limits_{i=1}^{n} x_i^2 + \frac{2}{n^2} \cdot \sum \sum \frac{\sigma_{x_i}^2}{\sigma_{X}^2} \cdot x_i \cdot x_j \right\} + \frac{2}{n^2} \cdot \sum \sum \frac{\sigma_{x_i}^2}{\sigma_{X}^2} \cdot x_i \cdot x_j \]

Proof:

\[ E\left( (\bar{Y})^2 \setminus \bar{X} \right) = \frac{1}{n^2} \cdot E\left( (Y_1 + Y_2 + \ldots + Y_n)^2 \setminus \bar{X} \right) \]

\[ = \frac{1}{n^2} \cdot E\left( Y_1^2 + Y_2^2 + \ldots + Y_n^2 + 2 \cdot \sum \sum Y_i \cdot Y_j \right) \setminus \bar{X} \]

\[ = \frac{1}{n^2} \cdot \left\{ E(Y_1^2 \setminus x_1) + E(Y_2^2 \setminus x_2) + \ldots + E(Y_n^2 \setminus x_n) \right\} \]

\[ + \frac{2}{n^2} \cdot \sum \sum E(Y_i \cdot Y_j \setminus x_i \cdot x_j) \]

\[ = \frac{1}{n} \cdot \left\{ \sigma_{Y_1}^2 \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot \sum \sum x_i^2 + \frac{2}{n} \cdot \sum \sum \frac{\sigma_{x_i}^2}{\sigma_{X}^2} \cdot x_i \cdot x_j \right\} \]

\[ = \frac{\sigma_{Y_1}^2}{n} \cdot \left\{ (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot \sum \sum x_i^2 + \frac{2}{n} \cdot \sum \sum \frac{\sigma_{x_i}^2}{\sigma_{X}^2} \cdot x_i \cdot x_j \right\} \]

\[ = \frac{\sigma_{Y_1}^2}{n} \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot \sum \sum x_i \cdot x_j \]

\[ = \frac{\sigma_{Y_1}^2}{n} \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot \sum \sum x_i \cdot x_j \]

\[ = \frac{\sigma_{Y_1}^2}{n} \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot \left( \sum x_i \right)^2 \]

\[ = \frac{\sigma_{Y_1}^2}{n} \cdot (1 - \rho^2) + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot \left( \bar{X} \right)^2 \]

Lemma 5:

\[ E\left( (\bar{X})^2 \cdot (\bar{Y})^2 \setminus \bar{X} \right) = \frac{1}{n^2} \cdot \left\{ n \cdot \sigma_{Y_1}^2 \cdot (1 - \rho^2) \cdot (\bar{X})^2 \right\} + \rho^2 \cdot \frac{\sigma_{Y_1}^2}{\sigma_{X}^2} \cdot (\bar{X})^4 \]

Proof:

Follows from Lemma 4.

Lemma 6:

\[ E\left( (\bar{X})^4 \right) = \frac{h}{n^3} + \frac{3 \cdot (n - 1) \cdot \sigma_1^4}{n^3} \quad \text{where} \quad h = E\left( X^4 \right) \]
Lemma 7:
\[ E\left( \left( \bar{X} \right)^2 \cdot \left( \bar{Y} \right)^2 \right) = \frac{\sigma_2^2 \cdot \sigma_1^2 \cdot \left( 1 - \rho^2 \right)}{n^2} + \frac{\rho^2 \cdot \sigma_1^2}{n^3} \cdot \left\{ h + \frac{3 \cdot \left( n - 1 \right) \cdot \sigma_1^4}{n^3} \right\}. \]

Proof:
The result follows easily from Lemmas 5 and 6.

Remark 1:
From the definition of Kurtosis, it follows that
\[ h = (\text{kurtosis}) \cdot (\text{variance}^2) = (3 + \theta) \cdot \sigma_1^4 \]

Note: \( h = E\left( X^4 \right) \) as defined in Lemma 6 and \( \theta \) is the excess kurtosis of \( X \).

Remark 2:
From Lemma 7 and the above remark, it follows that
\[ E\left( \left( \bar{X} \right)^2 \cdot \left( \bar{Y} \right)^2 \right) = \frac{\left( 1 + 2 \cdot \rho^2 \right) \cdot \sigma_1^2 \cdot \sigma_2^2}{n^2} + \frac{\theta \rho \cdot \sigma_1^2 \cdot \sigma_2^2}{n^3}. \]

Lemma 8:
\[ E\left( \left( \bar{X} \right)^3 \cdot \bar{Y} \right) = \frac{\sigma_{12}}{\sigma_1^2} \cdot E\left( \left( \bar{X} \right)^4 \right) = \frac{h \cdot \sigma_{12}}{n^3 \cdot \sigma_1^2} + \frac{3 \cdot \left( n - 1 \right) \cdot \sigma_{12} \cdot \sigma_1^2}{n^3} \]

Proof:
Follows from the Corollary 1 and Lemma 6.

Remark 3:
One can easily show that
\[ E\left( \left( \bar{X} \right)^3 \cdot \bar{Y} \right) = \frac{3 \cdot \rho \cdot \sigma_1^3 \cdot \sigma_2}{n^2} + \frac{\theta \rho \cdot \sigma_1^3 \cdot \sigma_2}{n^3}. \]

Lemma 9:
\[ E\left( \left( \bar{Y} \right)^3 \cdot \bar{X} \right) = \frac{\sigma_{12}}{\sigma_1^2} \cdot E\left( \left( \bar{Y} \right)^4 \right) = \frac{\tilde{h} \cdot \sigma_{12}}{n^3 \cdot \sigma_2^2} + \frac{3 \cdot \left( n - 1 \right) \cdot \sigma_{12} \cdot \sigma_2^2}{n^3} \]
where \( \tilde{h} = E\left( Y^4 \right) = \frac{3 \cdot \rho \cdot \sigma_1 \cdot \sigma_2^3}{n^2} + \frac{\tilde{\theta} \rho \cdot \sigma_1 \cdot \sigma_2^3}{n^3}. \)

Proof:
Follows from Lemma 8 and symmetry.
Note: \( \bar{h} = E \left( Y^4 \right) \) as defined in Lemma 9 and \( \bar{\theta} \) is the excess kurtosis of \( Y \).

**Lemma 10:**

\[ E \left( A_n^2 \right) \rightarrow 8 \text{ as } n \rightarrow \infty \]

**Proof:**

Note that

\[
A_n^2 = n^4.a_{11}^2 \cdot (\bar{X})^4 + n^4.a_{22}^2 \cdot (\bar{Y})^4 + n^4.(4.a_{12}^2 + 2.a_{11}.a_{22}) \cdot (\bar{X})^2 \cdot (\bar{Y})^2 + 4.n^4.a_{11}.a_{12} \cdot (\bar{X})^3 \cdot \bar{Y} + 4.n^4.a_{22}.a_{12} \cdot \bar{X} \cdot (\bar{Y})^3
\]

By taking the expected value on both sides, we have

\[
E \left( A_n^2 \right) = n^4.a_{11}^2.E \left( (\bar{X})^4 \right) + n^4.a_{22}^2.E \left( (\bar{Y})^4 \right) + n^4.(4.a_{12}^2 + 2.a_{11}.a_{22}) \cdot E \left( (\bar{X})^2 \cdot (\bar{Y})^2 \right) + 4.n^4.a_{11}.a_{12} \cdot E \left( (\bar{X})^3 \cdot \bar{Y} \right) + 4.n^4.a_{22}.a_{12} \cdot E \left( \bar{X} \cdot (\bar{Y})^3 \right)
\]

Noting that

\[
a_{11} = \frac{1}{n.\sigma_1^2 \cdot (1 - \rho^2)}
\]

\[
a_{22} = \frac{1}{n.\sigma_2^2 \cdot (1 - \rho^2)}
\]

\[
a_{12} = \frac{-\rho}{n.\sigma_1.\sigma_2 \cdot (1 - \rho^2)}
\]

So, we have

\[
E \left( A_n^2 \right) = n^4 \cdot \frac{1}{n^2.\sigma_1^4 \cdot (1 - \rho^2)^2} \cdot \left\{ \frac{3.\sigma_1^4}{n^2} + \frac{\theta.\sigma_1^4}{n^3} \right\} + n^4 \cdot \frac{1}{n^2.\sigma_2^4 \cdot (1 - \rho^2)^2} \cdot \left\{ \frac{3.\sigma_2^4}{n^2} + \frac{\theta.\sigma_2^4}{n^3} \right\}
\]
Further simplification yields
\[ E(A_n^2) = \left(1 - \rho^2\right)^2 \left\{ 3 + 3 + 2 \left(1 + 2 \rho^2\right)^2 - 24 \rho^2 \right\} \]
\[ + \frac{2 \theta \rho^2 \left(1 + 2 \rho^2\right)}{n} - \frac{4 \theta \rho^2}{n \left(1 - \rho^2\right)^2} - \frac{4 \tilde{\theta} \rho^2}{n \left(1 - \rho^2\right)^2} \]
\[ + \frac{\theta}{n \left(1 - \rho^2\right)^2} + \frac{\tilde{\theta}}{n \left(1 - \rho^2\right)^2} \]
\[ = 8 \left(1 - \rho^2\right)^2 \left(1 - \rho^2\right)^2 + \frac{2 \theta \rho^2 \left(1 + 2 \rho^2\right)}{n} - \frac{4 \theta \rho^2}{n \left(1 - \rho^2\right)^2} - \frac{4 \tilde{\theta} \rho^2}{n \left(1 - \rho^2\right)^2} \]
\[ + \frac{\theta}{n \left(1 - \rho^2\right)^2} + \frac{\tilde{\theta}}{n \left(1 - \rho^2\right)^2} \]
So, \( E(A_n^2) \to 8 \) as \( n \to \infty \)

So, the limiting value of the Multivariate Kurtosis is 8 as \( n \to \infty \).

3. CONCLUSION

In this paper, we considered the general cases to check the behavior of the kurtosis of the sample mean and the kurtosis of the sample centroid as the sample sizes get larger. In both the univariate and the multivariate situations, we were dealing with random samples which were drawn from non-normal populations. As we can see throughout this paper, the limiting value of the kurtosis for the sample mean and the sample centroid (for the
large samples) is supporting the Central Limit Theorem which is an important result in Probability and Statistics. We are able to arrive at this conclusion by imposing a very mild condition that the fourth moment exists in the univariate situation. As for the bivariate situation, in addition to the fourth moment’s existence for each of the two variables, we needed two other conditions concerning the conditional first moment and the conditional second moment which are in the main body of this paper.

REFERENCES