Communications for Statistical Applications and Methods
2014, Vol. 21, No. 6, 539–559

DOI: http://dx.doi.org/10.5351/CSAM.2014.21.6.539
Print ISSN 2287-7843 / Online ISSN 2383-4757

A View on the Validity of Central Limit Theorem: An Empirical Study Using Random Samples from Uniform Distribution

Chanmi Lee[a], Seungah Kim[a], Jaesik Jeong[1,a]

[a]Department of Statistics, Chonnam National University, Korea

Abstract

We derive the exact distribution of summation for random samples from uniform distribution and then compare the exact distribution with the approximated normal distribution obtained by the central limit theorem. To check the similarity between two distributions, we consider five existing normality tests based on the difference between the target normal distribution and empirical distribution: Anderson-Darling test, Kolmogorov-Smirnov test, Cramer-von Mises test, Shapiro-Wilk test and Shaprio-Francia test. For the purpose of comparison, those normality tests are applied to the simulated data.

It can sometimes be difficult to derive an exact distribution. Thus, we try two different transformations to find out which transform is easier to get the exact distribution in terms of calculation complexity. We compare two transformations and comment on the advantages and disadvantages for each transformation.

Keywords: Central limit theorem, uniform distribution, normal distribution, Anderson-Darling test, Kolmogorov-Smirnov test, Cramer-von Mises test, Shapiro-Wilk test and Shapiro-Francia test.

1. Introduction

Many distribution related theories have been developed in the field of statistics. Given data, it is very important to find useful information from data for statistical inference such as parameter estimation. Researchers are often interested in the distributional characteristics of the data like the center of the distribution. To obtain such information, we need to know the distribution of sample mean because the sample mean is a good estimator of the population mean in consideration of many theoretical properties. However, for some distribution, it is little bit complicated to derive the exact distribution of the sample mean if the size of the data is large (say, 8 or 10). To avoid such complicated calculation a central limit theorem(CLT) can be used to approximate the exact distribution. Such approximation is closer to the exact distribution with the rate of $\sqrt{n}$ as the sample size(n) increases.

This paper validates the central limit theorem using random samples from uniform distribution. Our goal is to answer the following question: How big is big enough in terms of sample size so that we can use the approximated normal distribution instead of the exact distribution of summation of random samples without any problem? To answer the question, we consider uniform distribution and derive the exact distribution of the summation of random samples for different sample sizes of up to 8. We then compare the exact distribution with the approximated distribution obtained by CLT. We use five existing normality tests to check the similarity between two distributions. The novelty of this paper is that we provide the derivation of the exact distribution of the summation of random samples.

[1] Corresponding author: Department of Statistics, Chonnam National University, 77 Yongbong-ro Yongbong-Dong, Buk-Gu, Gwangju 500-757, Korea. E-mail: jjs3098@chonnam.ac.kr

Published 30 November 2014 / journal homepage: http://csam.or.kr

2. Method

We provide five normality tests. Also, we provide the definition of CLT: Let $X_1, \ldots, X_n$ be random samples from a distribution with mean μ and variance σ^2. Then

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \to N(0,1) \quad \text{as} \quad n \to \infty.$$

For education purpose, we provide two wonderful references for CLT (Dinov *et al.*, 2008; Micheaux and Liquet, 2009).

2.1. Normality test

To date, about 40 normality test methods have been developed (Dufour *et al.*, 1998) since the seminal paper by Pearson who worked on the skewness and kurtosis coefficients (Althouse *et al.*, 1998). Those tests for normality differ in two ways: (1) the characteristics of the normal distribution the tests focus on (for example, skewness or kurtosis) (2) distance measure the tests use (for example, absolute difference or squared difference between theoretical distribution function and empirical distribution function). Here we focus on the five selected tests: Kolmogorov-Smirnov test, Anderson-Darling test, Cramer-von Mises test, Shapiro-Wilk test and Shapiro-Francia test. To test for normality, those methods use the distance between F (normal distribution) and F_n(empirical distribution) which plays a key role to test for the equality of two distributions. We briefly introduce the five tests and refer the reader to the original papers for more details.

Kolmogorov-Smirnov(KS) test: The empirical distribution function F_n for n iid observations X_i's is defined as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{X_i \leq x}$$

where $I_{X_i \leq x}$ is the indicator function. Suppose that the F is the target normal distribution. Then the Kolmogorov-Smirnov statistic is defined as

$$T_{\mathrm{KS}} = sup_x |F_n(x) - F(x)|$$

where sup_x is the supremum of the set of distances. The null hypothesis that the samples come from the normal distribution is rejected if the T_{KS} is larger than the tabulated values calculated by Nikolai Vasilyevich Smirnov in 1948 (Smirnov, 1948). Note that Glivenko-Cantelli theorem supports the theoretical background of the test.

Anderson-Darling(AD) test: Anderson-Darling test is a modification of the KS test, *i.e.*, more weight is given to the tails compared to KS test. The test statistic is defined as

$$T_{\mathrm{AD}} = -n - S$$

where $S = \sum_{i=1}^{n} \frac{2i-1}{n} [\ln F(Y_i) + \ln(1 - F(Y_{n+1-i}))]$ and F is the cdf of normal distribution, and Y_i is the ordered data. The null hypothesis that the samples come from the normal distribution is rejected if the T_{AD} is larger than the tabulated values calculated by Stephens (Stephens, 1974; Stephens, 1976; Stephens, 1977).

Cramer-von Mises(CvM) test: Let $Y_1, \ldots, Y_n$ be the ordered samples in increasing order. The test statistic is defined as

$$T_{\mathrm{CvM}} = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F(Y_i) \right]^2 .$$

Then the null hypothesis that the data come from the theoretical (normal) distribution (F) is rejected if the value of T_{CvM} is larger than the tabulated values calculated by Anderson, TW in 1962 (Anderson, 1962).

Shapiro-Wilk(SW) test: The Shapiro-Wilk test was developed in 1965 by Samuel Sanford Shapiro and Martin Wilk (Shapiro and Wilk, 1965). The test statistic is defined as

$$T_{\mathrm{SW}} = \frac{\left(\sum_{i=1}^{n} a_i y_i \right)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

where the constants a_i are given by $(a_1, \ldots, a_n) = m^T \Sigma^{-1} / (m^T \Sigma^{-1} \Sigma^{-1} m)^{1/2}$. Here y_i are ordered statistics from normal distribution in increasing order and its expectation $m = (m_1, \ldots, m_n) = (E(Y_1), \ldots, E(Y_n))$. The null hypothesis that the data come from the theoretical normal distribution (F) is then rejected if the value of T_{SW} is larger than the predetermined values. Compared to the Anderson-Darling test, the Shapiro-Wilk test is less affected by ties.

Shapiro-Francia(SF) test: The Shapiro-Francia test was developed in 1972 by Samuel Sanford Shapiro and RS Francia (Shapiro and Francia, 1972). The test statistic is defined as

$$T_{\mathrm{SF}} = \frac{\left(\sum_{i=1}^{n} b_i y_i \right)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

where $b_i = m^T / (m^T m)^{1/2}$ and m is the vector of expected values of standard normal order statistics. Note that T_{SF} presents the squared product moment correlation coefficient between ordered observed data and expected values of standard normal order statistics (Royston, 1983) with large values of T_{SF} indicating normality.

3. Exact Derivation of the Distribution of S_k

Let $U_1, \ldots, U_k$ be random samples from uniform distribution $U(0, 1)$ and $S_k = U_1 + \cdots + U_k$. In this section, we derive the distribution of S_k for four different values of $k = 2, 3, 4, 8$ by using the change-of-variable technique. Note that standard way of getting the distribution of sum of two independent random variables is to use convolution. It is well known that when two random variables X, Y are independent, the distribution of $Z = X + Y$ is represented as

$$h(z) = (f * g)(z) = \int f(x) g(z - x) dx$$

where f, g and h are probability density functions of X, Y and Z, respectively. However, in case of uniform distribution, each density function consists of a couple of different functions on each unit interval. Therefore, we decompose each density function into subfunction which is defined on each unit interval and then calculate the distribution sum of two independent random variables.

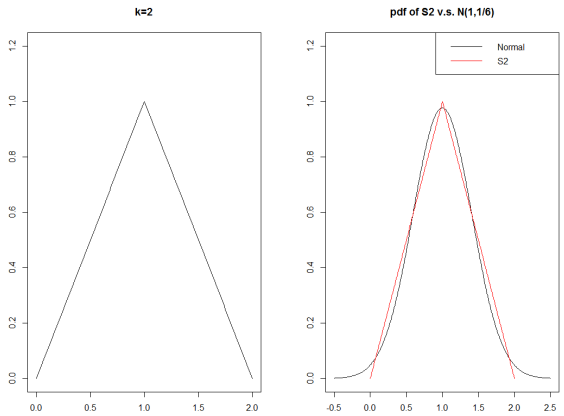

Figure 1: *Pdf of S_2 (left); Normal density function superimposed on pdf of S_2 (right).*

3.1. k=2

Let U_1 and U_2 be random samples from uniform distribution, and $Y_1 = U_1 + U_2$ and $Y_2 = U_1 - U_2$. Then Jacobian is calculated

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Therefore, the joint distribution of Y_1 and Y_2 is

$$g(y_1, y_2) = \frac{1}{2},$$

where the support of the joint distribution is depicted in Figure 1. The marginal distribution of Y_1 (*i.e.*, S_2) is

$$g_2(y) = \begin{cases} y, & 0 \le y \le 1, \\ 2 - y, & 1 \le y \le 2. \end{cases}$$

The mean and variance of S_2 are:

$$E(S_2) = E[U_1 + U_2] = 1,$$
$$\mathrm{Var}(S_2) = \mathrm{Var}(U_1 + U_2) = \frac{1}{6}.$$

Using CLT, we can approximate the distribution of S_2 with normal distribution, $N(1, 1/6)$. Figure 1 includes two density plots. The plot in left panel presents the exact distribution of S_2 and the one in right panel presents approximated normal distribution. Note that there are some difference between two curves, implying that the approximation with small samples is not accurate.

3.2. k=3

Let U_1, U_2 and U_3 be random samples from uniform distribution, and $Y_1 = U_1 + U_2 + U_3$ and $Y_2 = U_1 + U_2 - U_3$. Since we know the distribution of $X_1 = U_1 + U_2$ and $X_2 = U_3$, we can use them

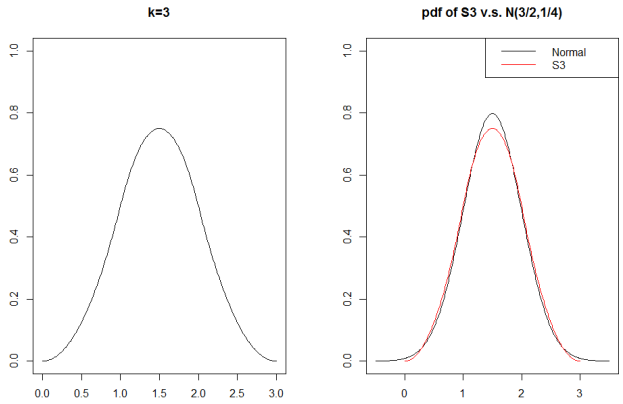

Figure 2: *Pdf of S_3 (left); Normal density function superimposed on pdf of S_3 (right).*

to get the distribution of Y_1. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Therefore, the joint distribution of Y_1 and Y_2 is

$$g(y_1, y_2) = \begin{cases} \frac{1}{4}(y_1 + y_2) & \text{if } 0 \le y_1 + y_2 \le 2, \\ \frac{1}{2}(2 - \frac{y_1 + y_2}{2}) & \text{if } 2 \le y_1 + y_2 \le 4. \end{cases}$$

The marginal distribution of Y_1 (*i.e.*, S_3) is

$$g_2(y) = \begin{cases} \frac{1}{2}y^2, & 0 \le y \le 1, \\ -y^2 + 3y - \frac{3}{2}, & 1 \le y \le 2, \\ \frac{1}{2}(y-3)^2, & 2 \le y \le 3. \end{cases}$$

The mean and variance of S_3 are:

$$E(S_3) = E[X_1 + X_2] = \frac{3}{2},$$
$$\text{Var}(S_3) = \text{Var}(X_1 + X_2) = \frac{1}{4}.$$

Using CLT, we can approximate the distribution of S_3 with normal distribution, $N(3/2, 1/4)$. Figure 2 includes two density plots. The plot in left panel presents the exact distribution of S_3 and the one in right panel presents approximated normal distribution.

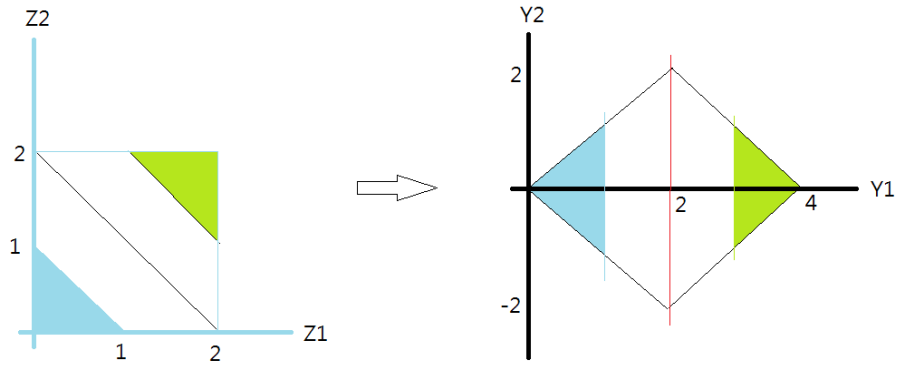

Figure 3: *Transformation of support; support of joint distribution of (Z_1, Z_2) (left) and (Y_1, Y_2) (right), respectively.*

3.3. k=4

Let $Y_1 = \underbrace{(U_1 + U_2)}_{Z_1} + \underbrace{(U_3 + U_4)}_{Z_2}$ and $Y_2 = \underbrace{(U_1 + U_2)}_{Z_1} - \underbrace{(U_3 + U_4)}_{Z_2}$. The distributions of Z_1 and Z_2 are given in the previous section. Then the Jacobian is the same as the previous case

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Figure 3 presents the transformation of the support by such change-of-variable technique: Support of the joint distribution of (Z_1, Z_2) and support of the joint distribution of (Y_1, Y_2). Using given information, we get the joint distribution of Z_1 and Z_2 as follows:

$$f(z_1, z_2) = \begin{cases} z_1 z_2, & 0 \le z_1 \le 1,\ 0 \le z_2 \le 1, \\ z_1(2 - z_2), & 0 \le z_1 \le 1,\ 1 \le z_2 \le 2, \\ (2 - z_1)z_2, & 1 \le z_1 \le 2,\ 0 \le z_2 \le 1, \\ (2 - z_1)(2 - z_2), & 1 \le z_1 \le 2,\ 1 \le z_2 \le 2. \end{cases}$$

In order to derive the marginal distribution of Y_1, we first need to derive the joint distribution of (Y_1, Y_2) and then calculate the marginal distribution of Y_1 by integrating out the joint distribution with respect to Y_2. To get the joint distribution of (Y_1, Y_2), we need to divide the support of Y_1 into 4 sub-intervals and then calculate corresponding joint density functions on each interval separately.

Since the joint density is represented as

$$g(y_1, y_2) = f(z_1, z_2)|J|,$$

the marginal density of Y_1 on the interval $y_2 \in [0, 1)$ is

$$g_1(y_1) = \int_{-y_1}^{y_1} \left(\frac{y_1^2 - y_2^2}{8} \right) dy_2 = \frac{y_1^3}{4} - \frac{y_1^3 + y_1^3}{24} = \frac{1}{6} y_1^3.$$

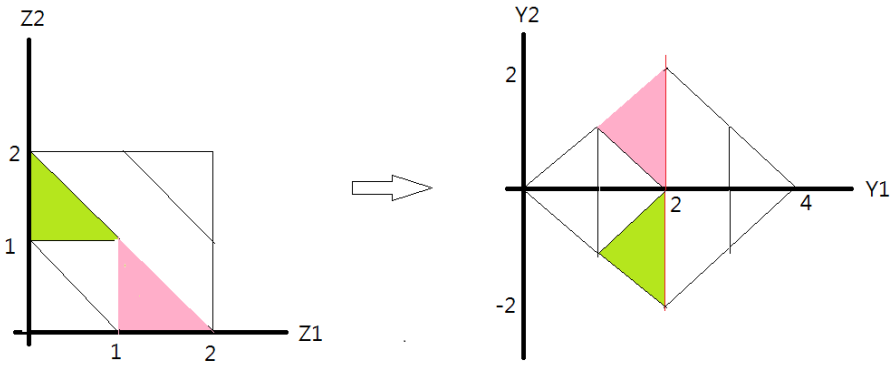

Figure 4: *Transformation of support;* (Z_1, Z_2) *plane to* (Y_1, Y_2) *plane.*

Similarly, marginal density on the interval $[3, 4)$ can be calculated:

$$g_1(y_1) = \int_{y_1-4}^{4-y_1} \left(2 - y_1 + \frac{y_1^2}{8} - \frac{y_2^2}{8}\right) dy_2 = -\frac{y_1^3}{6} + 2y_1^2 - 8y_1 + \frac{32}{3}.$$

Compared to the previous two intervals, the joint density on the interval $[1, 2)$ is more complicated. It can be obtained by combining three different parts, each coming from different supports. Each of three parts on (Z_1, Z_2) plane is transformed to different part on (Y_1, Y_2) plane. Figure 4 represents such transformation. The marginal density of Y_1 on interval $y_2 \in [1, 2)$ consists of the following three separate parts.

$$\begin{aligned} g_1(y_1) &= \int_{-y_1}^{y_1-2} \frac{y_1 + y_2}{2} \left(2 - \frac{y_1 - y_2}{2}\right) \frac{1}{2} dy_2 + \int_{y_1-2}^{2-y_1} \left(\frac{y_1^2}{8} - \frac{y_2^2}{8}\right) dy_2 + \int_{2-y_1}^{y_1} \left(2 - \frac{y_1 + y_2}{2}\right) \frac{y_1 - y_2}{2} \frac{1}{2} dy_2 \\ &= \frac{1}{6}\left(-3y_1^3 + 12y_1^2 - 12y_1 + 4\right). \end{aligned}$$

Similarly, the marginal density on $[2, 3)$ can be obtained by marginalization with respect to y_2:

$$\begin{aligned} g_1(y_1) &= \int_{y_1-4}^{2-y_1} \frac{y_1 + y_2}{2} \left(2 - \frac{y_1 - y_2}{2}\right) \frac{1}{2} dy_2 + \int_{2-y_1}^{y_1-2} \left(2 - \frac{y_1 + y_2}{2}\right)\left(2 - \frac{y_1 - y_2}{2}\right) \frac{1}{2} dy_2 \\ &\quad + \int_{y_1-2}^{4-y_1} \left(2 - \frac{y_1 + y_2}{2}\right) \frac{y_1 - y_2}{2} \frac{1}{2} dy_2 \\ &= \frac{1}{6}\left(-3y_1^3 - 24y_1^2 + 10y_1 - 44\right). \end{aligned}$$

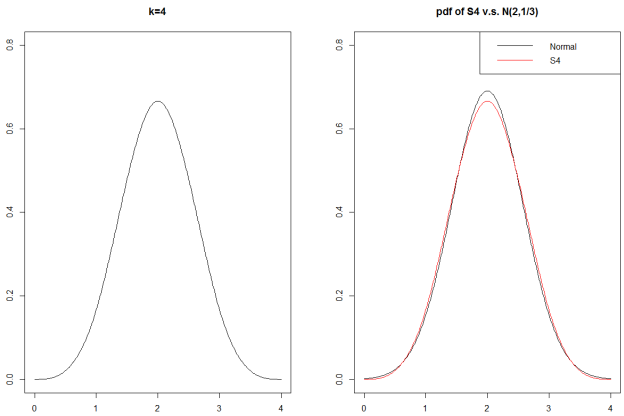

Figure 5: *Pdf of S_4 (left); Normal density function superimposed on pdf of S_4 (right).*

The density function of $S_4 (= U_1 + U_2 + U_3 + U_4)$ is

$$f_4(y) = \begin{cases} \dfrac{y^3}{6}, & 0 \le y < 1, \\ \dfrac{-3y^3 + 12y^2 - 12y + 4}{6}, & 1 \le y < 2, \\ \dfrac{3y^3 - 24y^2 + 60y - 44}{6}, & 2 \le y < 3, \\ \dfrac{-y^3 + 12y^2 - 48y + 64}{6}, & 3 \le y < 4. \end{cases}$$

The mean and variance of S_4 are:

$$E(S_4) = 4E(U_1) = 2,$$
$$\mathrm{Var}(S_4) = 4\mathrm{Var}(U_1) = \frac{1}{3}.$$

Using CLT, we can approximate the distribution of S_4 with normal distribution, $N(2, 1/3)$. Figure 5 includes two density plots. The figure in left panel presents the density function of S_4 while the other presents the corresponding normal distribution.

3.4. k=8

Let Y_1 and Y_2 be defined as:

$$Y_1 = \underbrace{(U_1 + U_2 + U_3 + U_4)}_{Z_1} + \underbrace{(U_5 + U_6 + U_7 + U_8)}_{Z_2},$$
$$Y_2 = \underbrace{(U_1 + U_2 + U_3 + U_4)}_{Z_1} - \underbrace{(U_5 + U_6 + U_7 + U_8)}_{Z_2}.$$

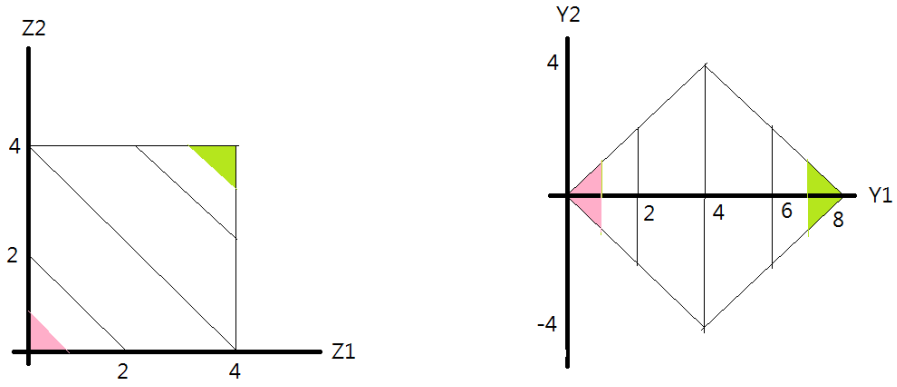

Figure 6: *Transformation of support ($k = 8$); support of joint distribution of (Z_1, Z_2) (left) and (Y_1, Y_2) (right), respectively.*

Note that each distribution of Z_1 and Z_2 is given in the previous section. The Jacobian is calculated

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

The following figure (Figure 6) presents the change of the support of the joint distribution by such transformation. After simple algebra, we can easily get the marginal distribution of Y_1

$$g_1(y_1) = \begin{cases} \frac{y_1^7}{5040}, & 0 \le y_1 \le 1, \\ \frac{1}{4608}\left[-\frac{32}{5}y_1^7 + \frac{256}{5}y_1^6 - \frac{768}{5}y_1^5 + 256y_1^4 - 256y_1^3 + \frac{768}{5}y_1^2 - \frac{256}{5}y_1 + \frac{256}{35}\right], & 1 \le y_1 \le 2, \\ \frac{1}{240}y_1^7 - \frac{1}{15}y_1^6 + \frac{13}{30}y_1^5 - \frac{3}{2}y_1^4 + \frac{55}{18}y_1^3 - \frac{37}{10}y_1^2 + \frac{223}{90}y_1 - \frac{149}{210}, & 2 \le y_1 \le 3, \\ \frac{1}{288}\left[-2y_1^7 + 48y_1^6 - 480y_1^5 + 2592y_1^4 - 8192y_1^3 + 15264y_1^2 - 15616y_1 + \frac{237792}{35}\right], & 3 \le y_1 \le 4, \\ \frac{1}{72}\left[\frac{1}{2}y_1^7 - 16y_1^6 + 216y_1^5 - 1592y_1^4 + 6912y_1^3 - 17688y_1^2 + 24768y_1 - \frac{513992}{35}\right], & 4 \le y_1 \le 5, \\ \frac{1}{36}\left[-\frac{3}{20}y_1^7 + 6y_1^6 - 102y_1^4 + 954y_1^4 - 5294y_1^3 + 17406y_1^2 - 31366y_1 - \frac{836754}{35}\right], & 5 \le y_1 \le 6, \\ \frac{1}{72}\left[\frac{1}{10}y_1^7 - \frac{24}{5}y_1^6 + \frac{492}{5}y_1^4 - 1116y_1^4 + 7556y_1^3 - \frac{152532}{5}y_1^2 + \frac{339524}{5}y_1 - \frac{2245596}{35}\right], & 6 \le y_1 \le 7, \\ \frac{1}{72}\left[-\frac{1}{70}y_1^7 + \frac{4}{5}y_1^6 - \frac{96}{5}y_1^5 + 256y_1^4 - 2048y_1^3 + \frac{49152}{5}y_1^2 - \frac{131072}{5}y_1 + \frac{1048576}{35}\right], & 7 \le y_1 \le 8. \end{cases}$$

More details about the derivation of the distribution of S_8 is given in the Appendix. The mean and variance of S_8 are:

$$E(S_8) = 8E(U_1) = 4,$$
$$\mathrm{Var}(S_8) = 8\mathrm{Var}(U_1) = \frac{2}{3}.$$

Using CLT, we can approximate the distribution of S_8 with normal distribution, $N(4, 2/3)$. Figure 7 includes two density plots. The figure in the left panel presents the density function of S_8 while the other presents the corresponding normal distribution.

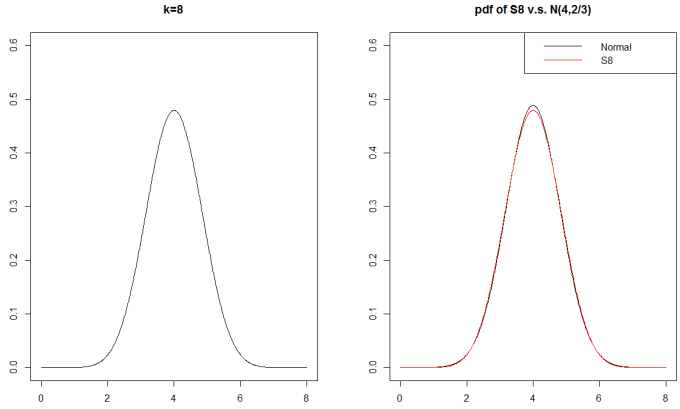

Figure 7: *Pdf of S_8 (left); Normal density function superimposed on pdf of S_8 (right).*

3.5. Practical view on another transformation

To get the exact distribution of the summation of random sample (S_2) from uniform distribution, we try two different transformations to find out which transformation is easier. In the previous section, we used the change-of-variable below:

$$Y_1 = Z_1 + Z_2, \quad Y_2 = Z_1 - Z_2,$$

because it has symmetry property. Now, we also try another type of transformation:

$$Y_1 = Z_1 + Z_2, \quad Y_2 = Z_1.$$

Here we get the joint distribution of (Y_1, Y_2) using the second transformation. In case of $k = 8$, $Y_1 = U_1 + \cdots + U_8$ and $Y_2 = U_1 + \cdots + U_4$. Then the joint distribution $g(y_1, y_2)$ is given

$$\begin{cases} \frac{1}{36}\left[-y_2^6 + 3y_1y_2^5 - 3y_1^2y_2^4 + y_1^3y_2^3\right], & 0 \le z_1 \le 1,\ 0 \le z_2 \le 1, \\ \frac{1}{36}\left[3y_2^6 + (12 - 9y_1)\,y_2^5 + \left(9y_1^2 - 24y_1 + 12\right)y_2^4 \right. \\ \qquad \left. + \left(-3y_1^3 + 12y_1^2 - 12y_1 + 4\right)y_2^3\right], & 0 \le z_1 \le 1,\ 1 \le z_2 \le 2, \\ \frac{1}{36}\left[-3y_2^6 + (9y_1 - 24)\,y_2^5 + \left(-9y_1^2 + 48y_1 - 60\right)y_2^4 \right. \\ \qquad \left. + \left(3y_1^3 - 24y_1^2 + 60y_1 - 44\right)y_2^3\right], & 0 \le z_1 \le 1,\ 2 \le z_2 \le 3, \\ \frac{1}{36}\left[y_2^6 + (12 - 3y_1)\,y_2^5 + \left(3y_1^2 - 24y_1 + 48\right)y_2^4 \right. \\ \qquad \left. + \left(-y_1^3 + 12y_1^2 - 48y_1 + 64\right)y_2^3\right], & 0 \le z_1 \le 1,\ 3 \le z_2 \le 4, \end{cases}$$

$$
\begin{cases}
\frac{1}{36}\left[3y_2^6-(9y_1+12)\,y_2^5+\left(9y_1^2+36y_1+12\right)y_2^4-\left(3y_1^3+36y_1^2+36y_1+4\right)y_2^3\right.\\
\quad\left.+\left(12y_1^3+36y_1^2+12y_1\right)y_1^2-\left(12y_1^3+12y_1^2\right)y_2+4y_1^3\right], & 1\le z_1\le 2,\ 0\le z_2\le 1,\\
\frac{1}{36}\left[-9y_2^6+27y_1y_2^5+\left(-27y_1^2-36y_1+72\right)y_2^4+\left(9y_1^3+72y_1^2-144y_1\right)y_2^3\right.\\
\quad+\left(-36y_1^3+36y_1^2+108y_1-48\right)y_2^2+\left(36y_1^3-108y_1^2+48y_1\right)y_2\\
\quad\left.+\left(-12y_1^3+48y_1^2-48y_1+16\right)\right], & 1\le z_1\le 2,\ 1\le z_2\le 2,\\
\frac{1}{36}\left[9y_2^6+(-279y_1+36)\,y_2^5+\left(27y_1^2-36y_1-72\right)y_2^4\right.\\
\quad+\left(-9y_1^3-36y_1^2+288y_1-312\right)y_2^3+\left(36y_1^3-180y_1^2+180y_1+96\right)y_2^2\\
\quad\left.+\left(-36y_1^3+252y_1^2-528y_1+288\right)y_2+\left(12y_1^3-96y_1^2+240y_1-176\right)\right], & 1\le z_1\le 2,\ 2\le z_2\le 3,\\
\frac{1}{36}\left[-3y_2^6+(9y_1-24)\,y_2^5+\left(-9y_1^2+36y_1-12\right)y_2^4+\left(3y_1^3-108y_1^2+244\right)y_2^3\right.\\
\quad+\left(-12y_1^3+108y_1^2-300y_1+240\right)y_2^2+\left(12y_1^3-132y_1^2+480y_1-576\right)y_2\\
\quad\left.+\left(-4y_1^3+48y_1^2-192y_1+256\right)\right], & 1\le z_1\le 2,\ 3\le z_2\le 4,
\end{cases}
$$

$$
\begin{cases}
\frac{1}{36}\left[-3y_2^6+(9y_1+24)\,y_2^5+\left(-9y_1^2-72y_1-60\right)y_2^4+\left(3y_1^3+72y_1^2+180y_1+44\right)y_2^3\right.\\
\quad\left.+\left(-24y_1^3-180y_1^2-132y_1\right)y_1^2+\left(60y_1^3+132y_1^2\right)y_2-44y_1^3\right], & 2\le z_1\le 3,\ 0\le z_2\le 1,\\
\frac{1}{36}\left[9y_2^6-(27y_1+36)\,y_2^5+\left(27y_1^2+144y_1-72\right)y_2^4+\left(-9y_1^3-180y_1^2+312\right)y_2^3\right.\\
\quad+\left(72y_1^3+252y_1^2-756y_1+96\right)y_2^2+\left(-180y_1^3+324y_1^2+336y_1-288\right)y_2\\
\quad\left.+\left(132y_1^3-528y_1^2+528y_1-176\right)\right], & 2\le z_1\le 3,\ 1\le z_2\le 2,\\
\frac{1}{36}\left[-9y_2^6+27y_1y_2^5+\left(-27y_1^2-72y_1+216\right)y_2^4+\left(9y_1^3+144y_1^2-432y_1\right)y_2^3\right.\\
\quad+\left(-72y_1^3+36y_1^2+1044y_1-1488\right)y_2^2+\left(180y_1^3-1044y_1^2+1488y_1\right)y_2\\
\quad\left.+\left(-132y_1^3+1056y_1^2-2640y_1+1936\right)\right], & 2\le z_1\le 3,\ 2\le z_2\le 3,\\
\frac{1}{36}\left[3y_2^6+(-9y_1+12)\,y_2^5+\left(9y_1^2-84\right)y_2^4+\left(-3y_1^3-36y_1^2+252y_1-284\right)y_2^3\right.\\
\quad+\left(24y_1^3-108y_1^2-156y_1+816\right)y_2^2+\left(-60y_1^3+588y_1^2-1824y_1+1728\right)y_2\\
\quad\left.+\left(44y_1^3-528y_1^2+2112y_1-2816\right)\right], & 2\le z_1\le 3,\ 3\le z_2\le 4,
\end{cases}
$$

$$
\begin{cases}
\frac{1}{36}\left[y_2^6-(3y_1+12)\,y_2^5+\left(3y_1^2+36y_1+48\right)y_2^4+\left(-y_1^3-36y_1^2-144y_1-64\right)y_2^3\right.\\
\quad\left.+\left(12y_1^3+144y_1^2+192y_1\right)y_1^2+\left(-48y_1^3-192y_1^2\right)y_2+64y_1^3\right], & 3\le z_1\le 4,\ 0\le z_2\le 1,\\
\frac{1}{36}\left[-3y_2^6+(9y_1+24)\,y_2^5+\left(-9y_1^2-84y_1-12\right)y_2^4+\left(3y_1^3+96y_1^2+156y_1-244\right)y_2^3\right.\\
\quad+\left(-36y_1^3-288y_1^2+432y_1+240\right)y_2^2+\left(144y_1^3-960y_1+576\right)y_2\\
\quad\left.+\left(-192y_1^3+768y_1^2-768y_1+256\right)\right], & 3\le z_1\le 4,\ 1\le z_2\le 2,\\
\frac{1}{36}\left[3y_2^6-(9y_1+12)\,y_2^5+\left(9y_1^2+60y_1-84\right)y_2^4+\left(-3y_1^3-84y_1^2+84y_1+284\right)y_2^3\right.\\
\quad+\left(36y_1^3+144y_1^2-1008y_1+816\right)y_2^2+\left(-144y_1^3+576y_1^2+192y_1-1728\right)y_2\\
\quad\left.+\left(192y_1^3-1536y_1^2+3840y_1-2816\right)\right], & 3\le z_1\le 4,\ 2\le z_2\le 3,
\end{cases}
$$

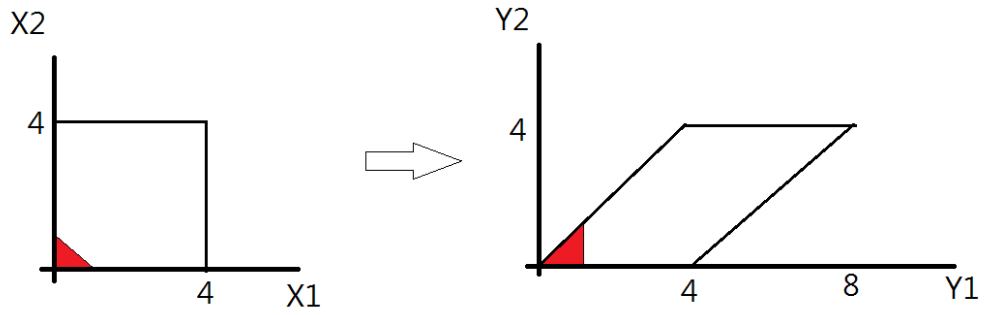

Figure 8: *Support transformation by the second transformation.*

$$\begin{cases} \frac{1}{36}\left[-y_2^6+3y_1y_2^5+\left(-3y_1^2-12y_1+48\right)y_2^4+\left(y_1^3+24y_1^2-96y_1\right)y_2^3\right. \\ \quad +\left(-12y_1^3+384y_1-768\right)y_2^2+\left(48y_1^3-384y_1^2+768y_1\right)y_2 \\ \quad \left.+\left(-64y_1^3+768y_1^2-3072y_1+4096\right)\right], \qquad 3\le z_1\le 4,\ 3\le z_2\le 4. \end{cases}$$

Given the joint distribution of (Y_1, Y_2), we can get the marginal distribution of Y_1 *i.e.*, (S_8) by integrating out the joint distribution with respect to Y_2. Compared to the first type of change-of-variable, the second type has advantages and disadvantages. Advantage is that (1) clearly, it has simple joint distribution function. (2) integration on each subinterval is relatively simple because one of integral section for each integration is constant. The marginal distribution of y_1, $f_1(y_1)$ on [0,1] (shaded in red in Figure 8) is represented as $\int_0^{y_1} f^1 dy_2$. However, the second transformation has a disadvantage in that to get the marginal distribution of Y_1 on [3,4], 7 different integrations (which are calculated on different support) should be done because a symmetry property is not applied to this kind of transformation. In contrast, only the first 4 integration are required since the first type of transformation has a symmetry property. Here, symmetry means that the result of the first integration is the same as the last integration.

4. Simulation

In this section, we check how close the distribution of S_k is to its approximated normal distribution obtained by CLT.

4.1. Simulation setup

Data generation consists of three steps:

Step1 Generate k random samples from $U(0,1)$: $u_1, \ldots, u_k$

Step2 Sum k samples: $y = \sum_{i=1}^{k} u_i$

Step3 Repeat the above two steps many times (say n): $y_1, \ldots, y_n$

Table 1: The number of rejection of null hypothesis for 3 different $n = 100, 500, 1000$ when $\alpha = 0.05$.

		AD	CvM	KS	SW	SF
n=100	S_2	70	51	52	94	32
	S_4	42	47	53	28	9
	S_8	40	40	32	29	19
	Norm	46	47	43	37	38
n=500	S_2	555	287	148	972	886
	S_4	93	73	60	118	58
	S_8	40	34	45	45	29
	Norm	52	49	57	55	53
n=1000	S_2	964	671	380	1000	1000
	S_4	187	148	96	324	182
	S_8	58	54	46	62	36
	Norm	51	50	49	43	49

Table 2: The number of rejection of null hypothesis for 3 different $n = 100, 500, 1000$ when $\alpha = 0.01$.

		AD	CvM	KS	SW	SF
n=100	S_2	12	7	5	17	3
	S_4	5	9	8	3	1
	S_8	9	8	5	5	5
	Norm	11	14	7	10	8
n=500	S_2	238	65	26	818	524
	S_4	21	18	14	24	7
	S_8	10	9	10	6	1
	Norm	9	11	10	10	12
n=1000	S_2	749	305	117	1000	999
	S_4	54	39	22	105	41
	S_8	11	12	11	15	7
	Norm	7	7	7	7	7

Given the simulated data, we can check the normality of the data by using five normality tests mentioned in previous section. We repeat the process many times (say B) and count the number of times that the normality test does not reject the null hypothesis that the data come from a normal distribution.

4.2. Simulation results

Normal approximation improves as sample size increases. Clearly, the distribution of S_k is closer to the corresponding approximated normal distribution as k increases. Here we consider $k = 2, 4, 8$, $n = 100, 500, 1000$, and $B = 1000$. As a reference, we also generate normal samples and check how many times the normality tests reject the null hypothesis that the data come from in a normal distribution. We consider two levels of significance $\alpha = 0.05$ and 0.01. Normality test results for all combination of parameters are given in the following tables. Table 1 includes the results of a normality test when $\alpha = 0.05$ while Table 2 includes the results for $\alpha = 0.01$.

In the results of Anderson Darling test (the first column in Table 1) especially for $n = 1000$, there are four numbers (such as 964, 187, 58 and 51). For example, 58 means the number of times that AD test correctly rejects the null hypothesis that data come from normal distribution because data were simulated from the distribution of S_8. The AD test made a wrong decision most of the time (942 times out of 1000). The reason is that the distribution of S_8 is very close to normal distribution such that normality test cannot tell the samples of S_8 from normal samples. Similar to AD test, other tests also made a wrong decision for the same reason. Generally speaking, the reason of poor performance of the normality tests is because the two distributions are very similar and not because the tests are

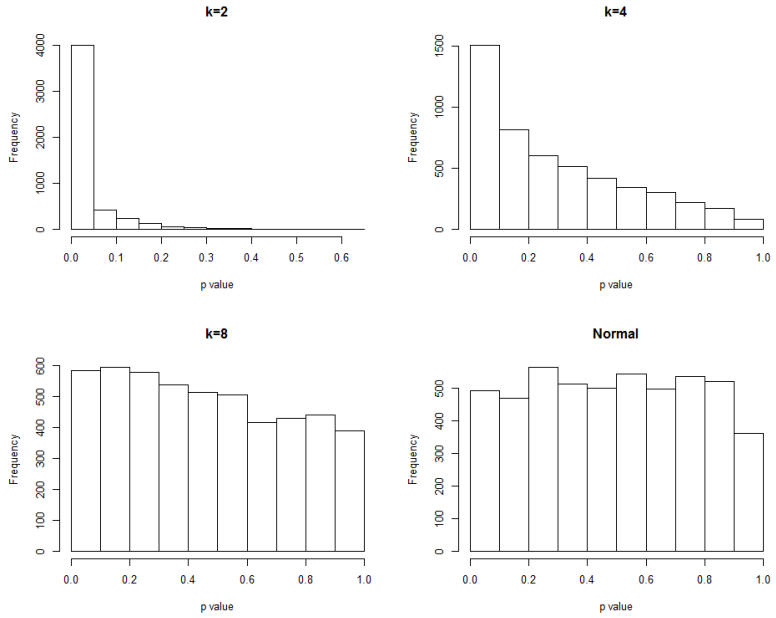

Figure 9: *Histogram of p values when $n = 1000$; $k = 2$ (top left); $k = 4$ (top right); $k = 8$ (bottom left); Normal data (bottom right).*

improper.

We also provide the histogram of p values obtained by combining all normality test results together, *i.e.*, 5000 values. Figure 9 shows that the histogram of p values from S_k samples is closer to that of the p values from normal samples as k increases. Especially, the histogram of p values from the S_8 (bottom left histogram in the figure) is similar to the p values obtained from normal samples (bottom right histogram in the figure). The two histograms provide some hints of why a normality test cannot tell normal samples from the samples of S_8. Such phenomenon is not so obvious in S_2 and S_4 compared to S_8. Note that the four histograms tell us that the distribution of S_k is similar to the normal distribution with sample size.

From another angle, we provide a plot that represents the number of correct rejection (NCR or empirical power) for the combination of all parameters such as k, n and α. Note that NCR and empirical power are equivalent measure when we use non-normal samples. Based on the results in Figure 10, each test show a different performance for the case $k = 2$; however, performance is getting similar as k increases. We also noticed that the effect of sample size on performance (NCR or empirical power) is tremendous when k is small. However, such effect is insignificant when k is big. That is, for small k, empirical power rapidly increased as sample size increases; however, empirical power increases relatively slowly with sample size for large k.

5. Conclusion

In practice, we sometimes want to know the distribution of the summation of random samples for statistical inference purposes. However, for many distribution, it is not too simple to derive and exact distribution. The central limit theorem(CLT) helps provide an approximated normal distribution we can use instead of and exact distribution. CLT enables a statistical inference without the derivation

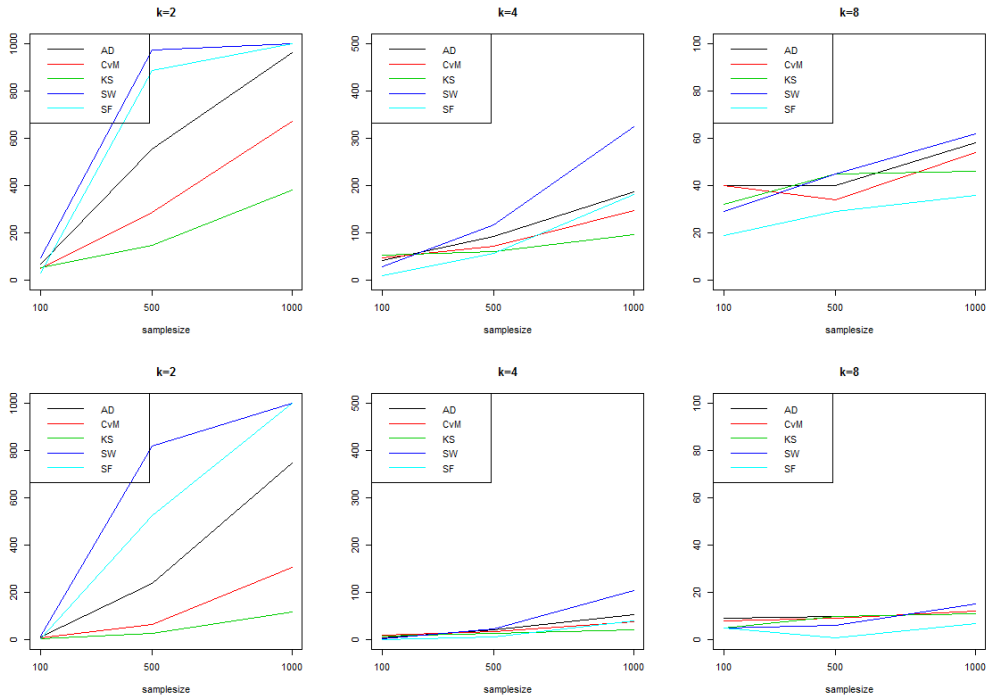

Figure 10: *The number of correct rejection(NCR); Each row corresponds to two* $\alpha = 0.05$ *and 0.01, respectively. Each column corresponds to* k*, i.e.* $k = 2$ *(left),* $k = 4$ *(center),* $k = 8$ *(right).*

of the exact distribution if the sample size is large. However, we still face the following practical question. How big is big enough for us to use the approximated distribution without any problem? We used random samples from uniform distribution to answer the question from a practical perspective.

In this paper, we first derived the exact distribution of summation of random samples from uniform distribution. We then compared the exact distribution with the approximated normal distribution. From this comparison, we noticed that the sample size of 8 is big enough to support the validity of the central limit theorem. Two curves (pdf of S_8 and approximated normal pdf in Figure 5) were overlapped over the whole range of the support of S_8. Furthermore, the biggest difference between two curves happen at the center and at 0.0089. For comparison purpose, in case of S_4, the biggest difference between two curves is about 0.0242. It is clear that such biggest difference decreases as k increases. However, an adequate sample size should be considered depending on situation.

Appendix: Derivation of the Distribution of the S_8

Let Y_1 and Y_2 be denoted by:

$$Y_1 = \underbrace{(U_1 + U_2 + U_3 + U_4)}_{Z_1} + \underbrace{(U_5 + U_6 + U_7 + U_8)}_{Z_2},$$
$$Y_2 = \underbrace{(U_1 + U_2 + U_3 + U_4)}_{Z_1} - \underbrace{(U_5 + U_6 + U_7 + U_8)}_{Z_2}.$$

Note that each distribution of Z_1 and Z_2 is given in Section 3.3. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Using given information, we get the joint distribution of Z_1 and Z_2

$$f(z_1, z_2) = \begin{cases} \frac{z_1^3 z_2^3}{36}, & 0 \le z_1 \le 1,\ 0 \le z_2 \le 1, \\ \frac{z_1^3}{6} \cdot \frac{-3z_2^3 + 12z_2^2 - 12z_2 + 4}{6}, & 0 \le z_1 \le 1,\ 1 \le z_2 \le 2, \\ \frac{z_1^3}{6} \cdot \frac{3z_2^3 - 24z_2^2 + 60z_2 - 44}{6}, & 0 \le z_1 \le 1,\ 2 \le z_2 \le 3, \\ \frac{z_1^3}{6} \cdot \frac{-z_2^3 + 12z_2^2 - 48z_2 + 64}{6}, & 0 \le z_1 \le 1,\ 3 \le z_2 \le 4, \\ \frac{-3z_1^3 + 12z_1^2 - 12z_1 + 4}{6} \cdot \frac{z_2^3}{6}, & 1 \le z_1 \le 2,\ 0 \le z_2 \le 1, \\ \frac{-3z_1^3 + 12z_1^2 - 12z_1 + 4}{6} \cdot \frac{-3z_2^3 + 12z_2^2 - 12z_2 + 4}{6}, & 1 \le z_1 \le 2,\ 1 \le z_2 \le 2, \\ \frac{-3z_1^3 + 12z_1^2 - 12z_1 + 4}{6} \cdot \frac{3z_2^3 - 24z_2^2 + 60z_2 - 44}{6}, & 1 \le z_1 \le 2,\ 2 \le z_2 \le 3, \\ \frac{-3z_1^3 + 12z_1^2 - 12z_1 + 4}{6} \cdot \frac{-z_2^3 + 12z_2^2 - 48z_2 + 64}{6}, & 1 \le z_1 \le 2,\ 3 \le z_2 \le 4, \\ \frac{3z_1^3 - 24z_1^2 + 60z_1 - 44}{6} \cdot \frac{z_2^3}{6}, & 2 \le z_1 \le 3,\ 0 \le z_2 \le 1, \\ \frac{3z_1^3 - 24z_1^2 + 60z_1 - 44}{6} \cdot \frac{-3z_2^3 + 12z_2^2 - 12z_2 + 4}{6}, & 2 \le z_1 \le 3,\ 1 \le z_2 \le 2, \\ \frac{3z_1^3 - 24z_1^2 + 60z_1 - 44}{6} \cdot \frac{3z_2^3 - 24z_2^2 + 60z_2 - 44}{6}, & 2 \le z_1 \le 3,\ 2 \le z_2 \le 3, \\ \frac{3z_1^3 - 24z_1^2 + 60z_1 - 44}{6} \cdot \frac{-z_2^3 + 12z_2^2 - 48z_2 + 64}{6}, & 2 \le z_1 \le 3,\ 3 \le z_2 \le 4, \\ \frac{-z_1^3 + 12z_1^2 - 48z_1 + 64}{6} \cdot \frac{z_2^3}{6}, & 3 \le z_1 \le 4,\ 0 \le z_2 \le 1, \\ \frac{-z_1^3 + 12z_1^2 - 48z_1 + 64}{6} \cdot \frac{-3z_2^3 + 12z_2^2 - 12z_2 + 4}{6}, & 3 \le z_1 \le 4,\ 1 \le z_2 \le 2, \\ \frac{-z_1^3 + 12z_1^2 - 48z_1 + 64}{6} \cdot \frac{3z_2^3 - 24z_2^2 + 60z_2 - 44}{6}, & 3 \le z_1 \le 4,\ 2 \le z_2 \le 3, \\ \frac{-z_1^3 + 12z_1^2 - 48z_1 + 64}{6} \cdot \frac{-z_2^3 + 12z_2^2 - 48z_2 + 64}{6}, & 3 \le z_1 \le 4,\ 3 \le z_2 \le 4. \end{cases}$$

To re-express the joint distribution of (Z_1, Z_2) into the function of (Y_1, Y_2), we have to replace (Z_1, Z_2) in the 16 formula above with (Y_1, Y_2). Thus the re-expression of the joint distribution of Z_1 and Z_2 is:

$$
\begin{cases}
f^1 = \dfrac{1}{2304}\left(-y_2^6 + 3y_1^2y_2^4 - 3y_1^4y_2^2 + y_1^6\right), & 0 \le z_1 \le 1,\ 0 \le z_2 \le 1,\\
f^2 = \dfrac{1}{288}\left[\dfrac{3}{8}y_2^6 + 3y_2^5 + \left(-\dfrac{9}{8}y_1^2 + 3y_1 + 6\right)y_2^4 + \left(-6y_1^2 + 12y_1 + 4\right)y_2^3\right.\\
\quad + \left(\dfrac{18}{8}y_1^4 - \dfrac{57}{8}y_1^3 + 12y_1\right)y_2^2 + \left(3y_1 - 12y_1^3 + 12y_1^2\right)y_2\\
\quad \left. + y_1^3\left(-\dfrac{3}{8}y_1^3 + 3y_1^2 - 6y_1 + 4\right)\right], & 0 \le z_1 \le 1,\ 1 \le z_2 \le 2,\\
f^3 = \dfrac{1}{288}\left[-\dfrac{3}{8}y_2^6 - 6y_2^5 + \left(\dfrac{9}{8}y_1^2 - 6y_1 - 30\right)y_2^4 + \left(12y_1^2 - 60y_1 - 44\right)y_2^3\right.\\
\quad + \left(-\dfrac{9}{8}y_1^4 + 12y_1^3 - 132y_1\right)y_2^2 + \left(-6y_1^4 + 60y_1^3 - 132y_1^2\right)y_2\\
\quad \left. + \left(\dfrac{3}{8}y_1^6 - 6y_1^5 + 30y_1^4 - 44y_1^3\right)\right], & 0 \le z_1 \le 1,\ 2 \le z_2 \le 3,\\
f^4 = \dfrac{1}{288}\left[\dfrac{1}{8}y_2^6 + 3y_2^5 + \left(-\dfrac{3}{8}y_1^2 + 3y_1 + 24\right)y_2^4 + \left(-6y_1^2 + 48y_1 + 64\right)y_2^3\right.\\
\quad + \left(\dfrac{3}{8}y_1^4 - 6y_1^3 + 192y_1\right)y_2^2 + \left(3y_1^4 - 48y_1^3 + 192y_1^2\right)y_2\\
\quad \left. + y_1^3\left(-\dfrac{1}{8}y_1^3 + 3y_1^2 - 24y_1 + 64\right)\right], & 0 \le z_1 \le 1,\ 3 \le z_2 \le 4,
\end{cases}
$$

$$
\begin{cases}
f^5 = \dfrac{1}{288}\left[\dfrac{3}{8}y_2^3 - 3y_2^5 + \left(-\dfrac{9}{8}y_1^2 + 3y_1 + 6\right)y_2^4 + \left(6y_1^2 - 12y_1 - 4\right)y_2^3\right.\\
\quad + \left(-\dfrac{9}{8}y_1^4 - 6y_1^3 + 12y_1\right)y_2^2 + \left(-3y_1^4 + 12y_1^3 - 12y_1^2\right)y_2\\
\quad \left. + y_1^3\left(-\dfrac{3}{8}y_1^3 + 3y_1^2 - 6y_1 + 4\right)\right], & 1 \le z_1 \le 2,\ 0 \le z_2 \le 1,\\
f^6 = \dfrac{1}{36}\left[-\dfrac{9}{64}y_2^6 + \left(\dfrac{27}{64}y_1^2 - \dfrac{9}{4}y_1 + \dfrac{9}{2}\right)y_2^4 + \left(-\dfrac{27}{64}y_1^4 + \dfrac{9}{2}y_1^3 - 18y_1^2 + 27y_1 - 12\right)y_2^2\right.\\
\quad \left. + \left(-\dfrac{9}{64}y_1^6 - \dfrac{9}{4}y_1^5 + \dfrac{27}{2}y_1^4 - 39y_1^3 + 60y_1^2 - 48y_1 + 16\right)\right], & 1 \le z_1 \le 2,\ 1 \le z_2 \le 2,\\
f^7 = \dfrac{1}{36}\left[\dfrac{9}{64}y_2^6 + \dfrac{9}{8}y_2^5 + \left(-\dfrac{27}{64}y_1^2 + \dfrac{27}{8}y_1 - \dfrac{9}{2}\right)y_2^4 + \left(-\dfrac{9}{4}y_1^2 + 18y_1 - 39\right)y_2^3\right.\\
\quad + \left(\dfrac{27}{64}y_1^4 - \dfrac{27}{4}y_1^3 + 36y_1^2 - 72y_1 + 24\right)y_2^2\\
\quad + \left(\dfrac{9}{8}y_1^4 - 18y_1^3 + 99y_1^2 - 216y_1 + 144\right)y_2\\
\quad \left. + \left(-\dfrac{9}{64}y_1^6 + \dfrac{27}{8}y_1^5 - \dfrac{63}{2}y_1^4 + 144y1^3 - 336y_1^2 + 384y_1 - 176\right)\right], & 1 \le z_1 \le 2,\ 2 \le z_2 \le 3,
\end{cases}
$$

$$
\begin{cases}
f^8 = \frac{1}{36}\left[-\frac{3}{64}y_2^6 - \frac{3}{4}y_2^5 + \left(\frac{9}{64}y_1^2 - \frac{3}{2}y_1 - \frac{3}{4}\right)y_2^4 + \left(\frac{3}{2}y_1^2 - \frac{33}{2}y_1 + \frac{61}{2}\right)y_2^3\right. \\
\quad + \left(-\frac{9}{64}y_1^4 + 3y_1^3 - 18y_1^2 + \frac{33}{2}y_1 + 60\right)y_2^2 \\
\quad + \left(-\frac{3}{4}y_1^4 + \frac{33}{2}y_1^3 - \frac{249}{2}y_1^2 + 360y_1 - 288\right)y_2 \\
\quad \left. + \left(\frac{3}{64}y_1^6 - \frac{3}{2}y_1^5 + \frac{75}{4}y_1^4 - \frac{229}{2}y1^3 + 348y_1^2 - 480y_1 + 256\right)\right], \qquad 1 \le z_1 \le 2,\ 3 \le z_2 \le 4, \\
f^9 = \frac{1}{288}\left[-\frac{3}{8}y_2^6 + 6y_2^5 + \left(\frac{9}{8}y_1^2 - 6y_1 + 30\right)y_2^4 + \left(-12y_1^2 + 60y_1 + 44\right)y_2^3\right. \\
\quad + \left(-\frac{9}{8}y_1^4 + 12y_1^3 - 132y_1\right)y_2^2 + \left(6y_1^4 - 60y_1^3 + 132y_1^2\right)y_2 \\
\quad \left. + \left(\frac{3}{8}y_1^6 - 6y_1^5 + 30y_1^4 - 44y_1^3\right)\right], \qquad 2 \le z_1 \le 3,\ 0 \le z_2 \le 1, \\
f^{10} = \frac{1}{36}\left[\frac{9}{64}y_2^6 - \frac{9}{8}y_2^5 + \left(-\frac{27}{64}y_1^2 + \frac{27}{8}y_1 - \frac{9}{2}\right)y_2^4 + \left(\frac{9}{4}y_1^2 - 18y_1 + 39\right)y_2^3\right. \\
\quad + \left(\frac{27}{64}y_1^4 - \frac{27}{4}y_1^3 + 36y_1^2 - 72y_1 + 24\right)y_2^2 \\
\quad + \left(-\frac{9}{8}y_1^4 + 18y_1^3 - 99y_1^2 + 216y_1 - 144\right)y_2 \\
\quad \left. + \left(-\frac{9}{64}y_1^6 + \frac{27}{8}y_1^5 - \frac{63}{2}y_1^4 + 144y1^3 - 336y_1^2 + 384y_1 - 176\right)\right], \qquad 2 \le z_1 \le 3,\ 1 \le z_2 \le 2, \\
f^{11} = \frac{1}{36}\left[-\frac{9}{64}y_2^6 + \left(\frac{27}{64}y_1^2 - \frac{9}{2}y_1 + \frac{27}{2}\right)y_2^4 + \right. \\
\quad + \left(-\frac{27}{64}y_1^4 + 9y_1^3 - 72y_1^2 + 261y_1 - 372\right)y_2^2 \\
\quad \left. + \left(\frac{9}{64}y_1^6 - \frac{9}{2}y_1^5 + \frac{117}{2}y_1^4 - 393y1^3 + 1428y_1^2 - 2640y_1 + 1936\right)\right], \qquad 2 \le z_1 \le 3,\ 2 \le z_2 \le 3, \\
f^{12} = \frac{1}{36}\left[\frac{3}{64}y_2^6 + \frac{3}{8}y_2^5 + \left(-\frac{9}{64}y_1^2 + \frac{15}{8}y_1 - \frac{21}{4}\right)y_2^4 + \left(-\frac{3}{4}y_1^2 + \frac{21}{2}y_1 - \frac{71}{2}\right)y_2^3\right. \\
\quad + \left(\frac{9}{64}y_1^4 - \frac{15}{4}y_1^3 + 36y_1^2 - \frac{291}{2}y_1 + 204\right)y_2^2 \\
\quad + \left(\frac{3}{8}y_1^4 - \frac{21}{2}y_1^3 + \frac{219}{2}y_1^2 - 504y_1 + 864\right)y_2 \\
\quad \left. + \left(-\frac{3}{64}y_1^6 + \frac{15}{8}y_1^5 - \frac{123}{4}y_1^4 + \frac{527}{2}y1^3 - 1236y_1^2 + 2976y_1 - 2816\right)\right], \quad 2 \le z_1 \le 3,\ 3 \le z_2 \le 4, \\
f^{13} = \frac{1}{288}\left[\frac{1}{8}y_2^6 - 3y_2^5 + \left(-\frac{3}{8}y_1^2 + 3y_1 + 24\right)y_2^4 + \left(6y_1^2 - 48y_1 - 64\right)y_2^3\right. \\
\quad + \left(\frac{3}{8}y_1^4 - 6y_1^3 + 192y_1\right)y_2^2 + \left(-3y_1^4 + 48y_1^3 - 192y_1^2\right)y_2 \\
\quad \left. + \left(-\frac{1}{8}y_1^6 + 3y_1^5 - 24y_1^4 + 64y_1^3\right)\right], \qquad 3 \le z_1 \le 4,\ 0 \le z_2 \le 1,
\end{cases}
$$

$$
\begin{cases}
f^{14} = \dfrac{1}{36}\left[-\dfrac{3}{64}y_2^6 + \dfrac{3}{4}y_2^5 + \left(\dfrac{9}{64}y_1^2 - \dfrac{3}{2}y_1 - \dfrac{3}{4}\right)y_2^4 + \left(-\dfrac{3}{2}y_1^2 + \dfrac{33}{2}y_1 - \dfrac{61}{2}\right)y_2^3\right. \\
\quad + \left(-\dfrac{9}{64}y_1^4 + 3y_1^3 - 18y_1^2 + \dfrac{33}{2}y_1 + 60\right)y_2^2 \\
\quad + \left(\dfrac{3}{4}y_1^4 - \dfrac{33}{2}y_1^3 + \dfrac{249}{2}y_1^2 - 360y_1 + 288\right)y_2 \\
\quad \left. + \left(\dfrac{3}{64}y_1^6 - \dfrac{3}{2}y_1^5 + \dfrac{75}{4}y_1^4 - \dfrac{229}{2}y1^3 + 348y_1^2 - 480y_1 + 256\right)\right], \qquad 3 \le z_1 \le 4,\ 1 \le z_2 \le 2, \\
f^{15} = \dfrac{1}{36}\left[\dfrac{3}{64}y_2^6 - \dfrac{3}{8}y_2^5 + \left(-\dfrac{9}{64}y_1^2 + \dfrac{15}{8}y_1 - \dfrac{21}{4}\right)y_2^4 + \left(\dfrac{3}{4}y_1^2 - \dfrac{21}{2}y_1 + \dfrac{71}{2}\right)y_2^3\right. \\
\quad + \left(\dfrac{9}{64}y_1^4 - \dfrac{15}{4}y_1^3 + 36y_1^2 - \dfrac{291}{2}y_1 + 204\right)y_2^2 \\
\quad + \left(-\dfrac{3}{8}y_1^4 + \dfrac{21}{2}y_1^3 - \dfrac{219}{2}y_1^2 + 504y_1 - 864\right)y_2 \\
\quad \left. + \left(-\dfrac{3}{64}y_1^6 + \dfrac{15}{8}y_1^5 - \dfrac{123}{4}y_1^4 + \dfrac{527}{2}y1^3 - 1236y_1^2 + 2976y_1 - 2816\right)\right], \quad 3 \le z_1 \le 4,\ 2 \le z_2 \le 3, \\
f^{16} = \dfrac{1}{36}\left[-\dfrac{1}{64}y_2^6 + \left(\dfrac{3}{64}y_1^2 - \dfrac{3}{4}y_1 + 3\right)y_2^4 + \left(-\dfrac{3}{64}y_1^4 + \dfrac{3}{2}y_1^3 - 18y_1^2 + 96y_1 - 192\right)y_2^2\right. \\
\quad \left. + \left(\dfrac{1}{64}y_1^6 - \dfrac{3}{4}y_1^5 + 15y_1^4 - 160y_1^3 + 960y_1^2 - 3072y_1 + 4096\right)\right], \qquad 3 \le z_1 \le 4,\ 3 \le z_2 \le 4.
\end{cases}
$$

Support also should be represented using y_1 and y_2 instead of z_1 and z_2 even though we retain z_1 and z_2 for notational simplicity. If we denote the 16 functions above by $f^1, \ldots, f^{16}$ in order, then the marginal distribution of Y_1 is represented as:

$$
g_1(y_1) = \begin{cases}
\dfrac{1}{2}\displaystyle\int_{-y_1}^{y_1} f^1 dy_2, & 0 \le y_1 \le 1, \\
\dfrac{1}{2}\left[\displaystyle\int_{-y_1}^{y_1-2} f^2 dy_2 + \int_{y_1-2}^{2-y_1} f^1 dy_2 + \int_{2-y_1}^{y_1} f^5 dy_2\right], & 1 \le y_1 \le 2, \\
\dfrac{1}{2}\left[\displaystyle\int_{-y_1}^{y_1-4} f^3 dy_2 + \int_{y_1-4}^{2-y_1} f^2 dy_2 + \int_{2-y_1}^{y_1-2} f^6 dy_2 + \int_{y_1-2}^{4-y_1} f^5 dy_2 + \int_{4-y_1}^{y_1} f^9 dy_2\right], & 2 \le y_1 \le 3, \\
\dfrac{1}{2}\left[\displaystyle\int_{-y_1}^{y_1-6} f^4 dy_2 + \int_{y_1-6}^{2-y_1} f^3 dy_2 + \int_{2-y_1}^{y_1-4} f^7 dy_2 + \int_{y_1-4}^{4-y_1} f^6 dy_2 + \int_{4-y_1}^{y_1-2} f^{10} dy_2\right. \\
\quad \left. + \displaystyle\int_{y_1-2}^{6-y_1} f^9 dy_2 + \int_{6-y_1}^{y_1} f^{13} dy_2\right], & 3 \le y_1 \le 4, \\
\dfrac{1}{2}\left[\displaystyle\int_{y_1-8}^{2-y_1} f^4 dy_2 + \int_{2-y_1}^{y_1-6} f^8 dy_2 + \int_{y_1-6}^{4-y_1} f^7 dy_2 + \int_{4-y_1}^{y_1-4} f^{11} dy_2 + \int_{y_1-4}^{6-y_1} f^{10} dy_2\right. \\
\quad \left. + \displaystyle\int_{6-y_1}^{y_1-2} f^{14} dy_2 + \int_{y_1-2}^{8-y_1} f^{13} dy_2\right], & 4 \le y_1 \le 5, \\
\dfrac{1}{2}\left[\displaystyle\int_{y_1-8}^{4-y_1} f^8 dy_2 + \int_{4-y_1}^{y_1-6} f^{12} dy_2 + \int_{y_1-6}^{6-y_1} f^{11} dy_2 + \int_{6-y_1}^{y_1-4} f^{15} dy_2 + \int_{y_1-4}^{8-y_1} f^{14} dy_2\right], & 5 \le y_1 \le 6,
\end{cases}
$$

$$\begin{cases} \frac{1}{2}\left[\int_{y_1-8}^{6-y_1} f^{12}dy_2 + \int_{6-y_1}^{y_1-6} f^{16}dy_2 + \int_{y_1-6}^{8-y_1} f^{15}dy_2\right], & 6 \le y_1 \le 7, \\ \frac{1}{2}\int_{y_1-8}^{8-y_1} f^{16}dy_2, & 7 \le y_1 \le 8. \end{cases}$$

As can be seen in the formula above, the marginal distribution of Y_1 on [3,4] consists of seven integration. However, it can be obtained by finishing the first four integration because the first integration is the same as the last integration, *i.e.*, $\int_{-y_1}^{y_1-6} f^4 dy_2 = \int_{6-y_1}^{y_1} f^{13}dy_2$. Let us call this property symmetry. Due to this kind of symmetry, we can easily get the marginal distribution of Y_1 after simple algebra,

$$g_1(y_1) =$$
$$\begin{cases} \frac{y_1^7}{5040}, & 0 \le y_1 \le 1, \\ \frac{1}{4608}\left[-\frac{32}{5}y_1^7 + \frac{256}{5}y_1^6 - \frac{768}{5}y_1^5 + 256y_1^4 - 256y_1^3 + \frac{768}{5}y_1^2 - \frac{256}{5}y_1 + \frac{256}{35}\right], & 1 \le y_1 \le 2, \\ \frac{1}{240}y_1^7 - \frac{1}{15}y_1^6 + \frac{13}{30}y_1^5 - \frac{3}{2}y_1^4 + \frac{55}{18}y_1^3 - \frac{37}{10}y_1^2 + \frac{223}{90}y_1 - \frac{149}{210}, & 2 \le y_1 \le 3, \\ \frac{1}{288}\left[-2y_1^7 + 48y_1^6 - 480y_1^5 + 2592y_1^4 - 8192y_1^3 + 15264y_1^2 - 15616y_1 + \frac{237792}{35}\right], & 3 \le y_1 \le 4, \\ \frac{1}{72}\left[\frac{1}{2}y_1^7 - 16y_1^6 + 216y_1^5 - 1592y_1^4 + 6912y_1^3 - 17688y_1^2 + 24768y_1 - \frac{513992}{35}\right], & 4 \le y_1 \le 5, \\ \frac{1}{36}\left[-\frac{3}{20}y_1^7 + 6y_1^6 - 102y_1^4 + 954y_1^4 - 5294y_1^3 + 17406y_1^2 - 31366y_1 - \frac{836754}{35}\right], & 5 \le y_1 \le 6, \\ \frac{1}{72}\left[\frac{1}{10}y_1^7 - \frac{24}{5}y_1^6 + \frac{492}{5}y_1^4 - 1116y_1^4 + 7556y_1^3 - \frac{152532}{5}y_1^2 + \frac{339524}{5}y_1 - \frac{2245596}{35}\right], & 6 \le y_1 \le 7, \\ \frac{1}{72}\left[-\frac{1}{70}y_1^7 + \frac{4}{5}y_1^6 - \frac{96}{5}y_1^5 + 256y_1^4 - 2048y_1^3 + \frac{49152}{5}y_1^2 - \frac{131072}{5}y_1 + \frac{1048576}{35}\right], & 7 \le y_1 \le 8. \end{cases}$$

References

Anderson, T. W. (1962). On the distribution of the two-sample Cramer-von Mises Criterion, *The Annals of Mathematical Statistics*, **33**, 1148–1159.

Althouse, L. A., Ware, W. B. and Ferron, J. M. (1998). Detecting departures from normality: A Monte Carlo simulation of a new omnibus test based on moments, *Annual meeting of the American Educational Research Association*.

Dufour, J. M., Farhat, A., Gardiol, L. and Khalaf, L. (1998). Simulation-based finite sample normality tests in linear regression, *Econometrics Journal*, **1**, 154–173.

Dinov, I. D., Christou, N. and Sanchez, J. (2008). Central limit theorem: New SOCR applet and demonstration activity, *Journal of Statistics Education*, **16**, 2.

Micheaux, P. L. and Liquet, B. (2009). Understanding convergence concepts: A visual-minded and graphical simulation-based approach, *American Statistician*, **63**, 173–178.

Royston, J. P. (1983). A simple method for evaluating the Shapiro-Francia W' test of non-normality, *The Statistician*, **32**, 297–300.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions, *The Annals of Mathematical Statistics*, **19**, 279–281.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality, *Biometrika*, **52**, 591–611.

Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality, *Journal of the American Statistical Association*, **67**, 215–216.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons, *Journal of the American Statistical Association*, **69**, 730–737.

Stephens, M. A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters, *Annals of Statistics*, **4**, 357–369.

Stephens, M. A. (1977). Goodness of fit for the extreme value distribution, *Biometrika*, **64**, 583–588.

Received September 30, 2014; Revised October 29, 2014; Accepted November 1, 2014