

Research Paper

A simple method for determining the limiting distribution of sample central moments for two-point and Binomial distributions

NARGES ABBASI*, NARGES KESHAVARZ, MASOUD YARMOHAMMADI, ABDOLLAH SAADATMAND
DEPARTMENT OF STATISTICS, COLLEGE OF SCIENCE, PAYAME NOOR UNIVERSITY, TEHRAN, IRAN

Received: May 23, 2023/ Revised: September 08, 2023/ Accepted: October 31, 2023

Abstract: Moments play an essential role in the characterization of statistical distributions and criteria such as dispersion, skewness, and kurtosis. This article is a dissection of the central moments of two-point and binomial distributions. First, we consider the Bernoulli distribution of the population and generalize the results. With a simple method, we present the condition that when the sample size is large, the structure of the sample central moment consists of random variables independent of standard normal or chi-square or a combination of both. In the obtained results, the role of points that have a probability of $1/2$ is very influential in the limit distribution.

Keywords: Binomial distribution; Central moment; Convergence in probability.

Mathematics Subject Classification (2010): 60F05, 60E05.

1 Introduction

In statistical inference, when the sample size is large, the distribution of estimators is tied to the normal distribution based on limit theorems, or in higher dimensions to the multivariate normal distribution. Most of the time, these assumptions are acceptable for statistical populations that have finite moments. But if we want to be more precise about the asymptotic distribution of statistics, not all estimators follow the normal distribution. In this paper, we try to clearly obtain their limiting distribution by using a characteristic of Bernoulli random variable. Bernoulli distribution is considered from

*Corresponding author: n_abbasi@pnu.ac.ir

different aspects, based on moments and central moments (Knoblauch, 2008; Nguyen, 2021; Nowakowski, 2021; Ahle, 2022).

Moments and central moments can be seen almost in statistical textbooks. The use of moments plays a significant role in determining skewness and kurtosis and determining statistical distributions. He et al. (2011) have used four central moments to determine the limit of probabilities and estimate the generating function of moments. Fisher (1930) was one of the first to use moments in normality tests.

Determining the limit distribution of moments can also be seen in the topics of U -statistics. The generalization of the theory of U -statistics is referred to Hoeffding (1948). Properties and theorems related to this group of statistics were presented by Denker (1985) and Lee (1990). Heffernan (1997) presented a statistically unbiased estimator for the sample central moment in the family of distributions with finite central moments. These estimators are multiples of the sample central moment was presented by Abbasi (2008). By displaying these relationships, we could determine the distribution of central moments from the distribution of U -statistics. In limit theorems, in a general expression and using operators, depending on the conditions, two types of tendencies in the distribution were mentioned.

What leads to the difference between the results of this article and other previous researches is that this simple method includes polynomial expansion and basic theorems of calculus and differential. First, in Section 2, we obtain the results for the Bernoulli distribution, and we generalize it to any two-point distribution. In Section 3, we obtain results for the binomial distribution. In Section 4, the results obtained in the previous two sections are confirmed by simulation and Kolmogorov-Smirnov test.

2 Bernoulli distribution

Let $X \sim Ber(p)$, then for $k \in N$ and $p \in (0, 1)$, we see

$$\mu_k(p) = E(X - p)^k = pq \left((-1)^k p^{k-1} + q^{k-1} \right), \quad q = 1 - p.$$

Obviously $\mu_k(\frac{1}{2}) = 0$ if k odd and $\mu_k(\frac{1}{2}) = (\frac{1}{2})^k$ if k even.

Now X_1, \dots, X_n are independent identically Bernoulli random variables with parameter p , $p \in (0, 1)$. The mean sample, \bar{X} is the sufficient statistic and maximum likelihood estimator of p , also by weak law of large numbers $\bar{X} \rightarrow p$ in probability and by central limit theorem $Z_n = \frac{\sqrt{n}(\bar{X}-p)}{\sqrt{pq}} \rightarrow Z$ in distribution as $n \rightarrow \infty$, where Z is random variable from standard norm distribution.

The k^{th} central moment of sample, M_k , is

$$M_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k = \frac{1}{n} \sum_{i=1}^n \sum_{j=0}^k \binom{k}{j} X_i^j (-\bar{X})^{k-j}. \quad (1)$$

Equality $X^j = X$, $j = 1, 2, \dots$, (since $P(X \in \{0, 1\}) = 1$) allows us to write (1) as

$$nM_k = n\bar{X}(1 - \bar{X}) \left((-1)^k \bar{X}^{k-1} + (1 - \bar{X})^{k-1} \right). \quad (2)$$

We obtain the limiting distribution of M_k in Bernoulli population for following three cases.

2-1. Case $p \neq \frac{1}{2}$.

We rewrite formula (2) according to Z_n and μ_k .

$$\begin{aligned} nM_k &= n \frac{1}{\bar{X}} \left(\sqrt{\frac{pq}{n}} \left(\frac{\bar{X} - p}{\sqrt{\frac{pq}{n}}} \right) + p \right) \mu_k(\bar{X}) \\ &= \sqrt{n} \sqrt{pq} \frac{1}{\bar{X}} \mu_k(\bar{X}) Z_n + n \frac{p}{\bar{X}} \mu_k(\bar{X}), \end{aligned}$$

or

$$\sqrt{n} \left(M_k - \frac{p}{\bar{X}} \mu_k(\bar{X}) \right) = \sqrt{pq} \frac{1}{\bar{X}} \mu_k(\bar{X}) Z_n.$$

According to $\bar{X} \xrightarrow{P} p$, $\frac{p}{\bar{X}} \xrightarrow{P} 1$, $\sqrt{pq} \frac{1}{\bar{X}} \xrightarrow{P} \sqrt{\frac{q}{p}}$ and $\mu_k(\bar{X}) \xrightarrow{P} \mu_k(p)$, the following result obtained by Slutsky's theorem (Casetlla and Berger, 1990).

$$\sqrt{n}(M_k - \mu_k(p)) \xrightarrow{D} N \left(0, \frac{q}{p} (\mu_k(p))^2 \right), \quad \text{as } n \rightarrow \infty. \quad (3)$$

2-2. Case $p = \frac{1}{2}$ and k odd.

In this case for k odd the k^{th} central moment Bernoulli is zero and the last term in (2) equals to

$$(-1)^k \bar{X}^{k-1} + (1 - \bar{X})^{k-1} = (1 - \bar{X})^{k-1} - \bar{X}^{k-1}, \quad (4)$$

which (4) divisible by $(\bar{X} - \frac{1}{2})$ and from this, we make random variable Z_n . On the other hand, by L'Hopital's rule (Leithold, 1976), we can see that

$$\lim_{\bar{X} \rightarrow \frac{1}{2}} \frac{(1 - \bar{X})^{k-1} - \bar{X}^{k-1}}{\bar{X} - \frac{1}{2}} = -2(k-1) \left(\frac{1}{2} \right)^{k-2}, \quad n \rightarrow +\infty.$$

Again we will return to (2) and rewrite

$$\begin{aligned} nM_k &= n\bar{X}(1 - \bar{X}) \left(\bar{X} - \frac{1}{2} \right) (-2(k-1)) \left(\frac{1}{2} \right)^{k-2} \\ &= n \frac{1}{4} \sqrt{\frac{1}{4n}} \left(\frac{\bar{X} - \frac{1}{2}}{\sqrt{\frac{1}{4n}}} \right) (-2(k-1)) \left(\frac{1}{2} \right)^{k-2} \\ &= -(k-1) \left(\frac{1}{2} \right)^k \sqrt{n} Z_n, \end{aligned}$$

or

$$\sqrt{n} M_k \xrightarrow{D} N(0, (k-1)^2 \left(\frac{1}{4} \right)^k), \quad \text{as } n \rightarrow \infty. \quad (5)$$

2-3. Case $p = \frac{1}{2}$ and k even.

For $k = 2$, we can easily see $n(M_2 - \frac{1}{4}) \xrightarrow{D} \frac{-1}{4} \chi_1^2$. and by expanding polynomials $n(M_4 - \frac{1}{16}) \xrightarrow{D} \frac{1}{8} \chi_1^2$. By changing the power of $\frac{1}{2}$, the simulation studies of the sequence $n[M_k - \mu_k(\frac{1}{2})]$ shows that the chi-square coefficient is different with these two

cases (Appendix A1). Therefore, the M_k should be analyzed in another way. For $k \geq 6$ even, the formula of (2) equals to

$$nM_k = n\bar{X}(1 - \bar{X})(\bar{X} - \frac{1}{2})^2 \left(\frac{\bar{X}^{k-1} + (1 - \bar{X})^{k-1} - (\frac{1}{2})^{k-2}}{(\bar{X} - \frac{1}{2})^2} \right) + n\bar{X}(1 - \bar{X})(\frac{1}{2})^{k-2}.$$

Since $\lim_{\bar{X} \rightarrow \frac{1}{2}} \frac{\bar{X}^{k-1} + (1 - \bar{X})^{k-1} - (\frac{1}{2})^{k-2}}{(\bar{X} - \frac{1}{2})^2} = (k-1)(k-2)(\frac{1}{2})^{k-3}$, we have

$$n[M_k - \mu_k(\frac{1}{2})] \xrightarrow{D} (k-1)(k-2)(\frac{1}{2})^{k+1} \chi_1^2, \quad (6)$$

where χ_1^2 is the random variable with chi-square distribution with one degree freedom.

2-4. Two-points distributions.

Let discrete random variable Y have probability model as $P(Y = y_1) = p, P(Y = y_2) = q$. With the linear transformation of random variable Bernoulli X on the random variable Y , we have

$$Y = y_1X + y_2(1 - X) = (y_1 - y_2)X + y_2.$$

Then $\bar{Y} = (y_1 - y_2)\bar{X} + y_2$ and $Y_i - \bar{Y} = (y_1 - y_2)(X_i - \bar{X})$, therefore

$$nM_k^Y = \sum_{i=1}^n ((y_1 - y_2)(X_i - \bar{X}))^k = (y_1 - y_2)^k nM_k,$$

where M_k^Y is k^{th} sample central moment of Y . It is obvious that the rest of the results are obtained by applying the $(y_1 - y_2)^k$ coefficient.

3 Binomial distribution

One of the special cases of discrete variables is the binomial random variable, which is equivalent to the sum of independent Bernoulli random variables.

$$Y_i = X_{1i} + \cdots + X_{mi}, \quad i = 1, \dots, n,$$

where $X_{ji} \sim Ber(p), i = 1, \dots, n, j = 1, \dots, m$ are independent. also

$$\bar{Y} = \bar{X}_1 + \cdots + \bar{X}_m.$$

This advantage will lead to the decomposition of sample central moments into the sum of Bernoulli independent sample central moments and the product moments of those independent variables.

$$\begin{aligned} nM_k^{(m)} &= \sum_{i=1}^n (Y_i - \bar{Y})^k \\ &= \sum_{i=1}^n \left(\sum_{j=1}^m (X_{ji} - \bar{X}_j) \right)^k \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^m \sum_{i=1}^n (X_{ji} - \bar{X}_j)^k + \sum_{i=1}^n \sum_{\substack{0 \leq l_1, \dots, l_m < k \\ \sum_{i=1}^k l_i = k}} \binom{k}{l_1, \dots, l_m} \prod_{j=1}^m (X_{ji} - \bar{X}_j)^{l_j} \\
&= n(M_k^{X_1} + \dots + M_k^{X_m}) + \sum_{i=1}^n \sum_{\substack{0 \leq l_1, \dots, l_m < k \\ \sum_{i=1}^k l_i = k}} \binom{k}{l_1, \dots, l_m} \prod_{j=1}^m (X_{ji} - \bar{X}_j)^{l_j}, \quad (7)
\end{aligned}$$

where $M_k^{(m)}$ is k^{th} sample central moment for Binomial distribution. As $n \rightarrow \infty$ the second part of (7) becomes the product of central moments.

$$\hat{L}(p; k, m) = \frac{1}{n} \sum_{i=1}^n \sum_{\substack{0 \leq l_1, \dots, l_m < k \\ \sum_{i=1}^k l_i = k}} \binom{k}{l_1, \dots, l_m} \prod_{j=1}^m (X_{ji} - \bar{X}_j)^{l_j}$$

So,

$$L(p, k, m) = \sum_{\substack{0 \leq l_1, \dots, l_m < k \\ \sum_{i=1}^k l_i = k}} \binom{k}{l_1, \dots, l_m} \prod_{j=1}^m \mu_{l_j}(p), \quad (8)$$

where from the independence of variables, the limit result equals to $L(p; k, m)$ with the moment estimator $\hat{L}(p; k, m)$.

Notice that: $\mu_k^m(p) = E(Y - mp)^k = m\mu_k(p) + L(p; k, m)$. The most applications of central moments are considered in the second, third and fourth orders: $\mu_2^m(p) = m\mu_2(p)$, $\mu_3^m(p) = m\mu_3(p)$ and $\mu_4^m(p) = m\mu_4(p) + 3m(m-1)(\mu_2(p))^2$, also, $\mu_2^m(\frac{1}{2}) = m\frac{1}{4}$, $\mu_4^m(\frac{1}{2}) = m(3m-2)(\frac{1}{2})^4$. And the number of non-zero terms of the right side of (8), $L(p)$ is equal to $\binom{m-1+\frac{k}{2}}{m-1} - m$. It is obtained by solving the equation $l_1 + l_2 + \dots + l_m = k$ where l_1, l_2, \dots, l_m even and does not count the terms of $l_j = k; j = 1, 2, \dots, m$. Despite the appearance of $L(p; k, m)$, it is possible to obtain bounds for it. If $p = \max(p, q)$, then $2q^{l_j+1} \leq \mu_{l_j}(p) = pq^{l_j-1} + q^{l_j-1} \leq 2p^{l_j+1}$. As a result

$$2^m q^{k+m} m(m^{k-1} - 1) \leq L(p; k, m) \leq 2^m p^{k+m} m(m^{k-1} - 1).$$

At first glance, it seems that in (7), in the limiting case, term $\hat{L}(p, k, m)$ can be omitted and the limiting distribution of the central moment can be assigned to $\sum_{i=1}^m M_k^{X_i}$. But the simulation study shows us that this assumption is not correct (Appendix A2). In the following, we try to get results for $k = 2, 3$.

3-1. Case $k = 2$.

For $p \in (0, 1)$,

$$\begin{aligned}
E(\hat{L}(p; 2, m)) &= E\left(\frac{1}{n} \sum_{1 \leq k \leq j < m} \sum_{i=1}^n (X_{ki} - \bar{X}_k)(X_{ji} - \bar{X}_j)\right) = 0, \\
E((X_{ki} - \bar{X}_k)^2) &= \frac{n-1}{n} \mu_2(p); \quad i = 1, 2, \dots, n, \quad k = 1, 2, \dots, m,
\end{aligned}$$

$$C = \sum_{\substack{1 \leq k_1 \leq j_1 < m \\ 1 \leq k_2 \leq j_2 < m \\ (k_1, j_1) \neq (k_2, j_2)}} \text{cov} \left[\sum_{i=1}^n (X_{k_2 i} - \bar{X}_{k_2})(X_{j_1 i} - \bar{X}_{j_1}), \sum_{i=1}^n (X_{k_2 i} - \bar{X}_{k_2})(X_{j_2 i} - \bar{X}_{j_2}) \right] = 0,$$

$$\begin{aligned} \text{Var}(\hat{L}(p; 2, m)) &= \frac{4}{n^2} \text{Var} \left(\sum_{1 \leq k \leq j < m} \sum_{i=1}^n (X_{ki} - \bar{X}_k)(X_{ji} - \bar{X}_j) \right) \\ &= \frac{4}{n^2} \left\{ \sum_{1 \leq k \leq j < m} \text{Var} \left[\sum_{i=1}^n (X_{ki} - \bar{X}_k)(X_{ji} - \bar{X}_j) \right] + C \right\} \\ &= \frac{4}{n^2} \sum_{1 \leq k \leq j < m} \sum_{i=1}^n \text{Var} \left((X_{ki} - \bar{X}_k)(X_{ji} - \bar{X}_j) \right) \\ &= \frac{4}{n^2} \sum_{1 \leq k \leq j < m} \sum_{i=1}^n E \left((X_{ki} - \bar{X}_k)^2 (X_{ji} - \bar{X}_j)^2 \right) \\ &= \frac{2}{n} m(m-1) \left(\frac{n-1}{n} \mu_2(p) \right)^2, \end{aligned}$$

also with a little simple calculation, it can be shown that the correlation coefficient between $\hat{L}(p; 2, m)$ and $M_2^{X_i}, i = 1, 2, \dots, m$ is zero. Then

$$E_n := \hat{L}(p; k, m) / \left(\mu_2(p) \frac{n-1}{n} \sqrt{\frac{2}{n}} \sqrt{m(m-1)} \right) \xrightarrow{D} Z, \quad \text{as } n \rightarrow \infty.$$

So

$$\begin{aligned} nM_2^{(m)} &= n \sum_{i=1}^m M_2^{(X_i)} + \mu_2(p) \left(\frac{n-1}{n} \right) \sqrt{2n} \sqrt{m(m-1)} E_n, \\ n(M_2^{(m)} - \mu_2^m) &= n \sum_{i=1}^m (M_2^{(X_i)} - \mu_2(p)) + \mu_2(p) \left(\frac{n-1}{n} \right) \sqrt{2n} \sqrt{m(m-1)} E_n, \end{aligned}$$

then by (3), $p \neq \frac{1}{2}$, as $n \rightarrow \infty$

$$\sqrt{n}(M_2^{(m)} - \mu_2^m(p)) = \sqrt{\frac{q}{p}} \mu_2(p) (Z_1 + \dots + Z_m) + \left(\frac{n-1}{n} \right) \sqrt{2m(m-1)} \mu_2(p) Z,$$

and $p = \frac{1}{2}$, as $n \rightarrow \infty$

$$n(M_2^{(m)} - \mu_2^m(p)) = -\frac{1}{4} (\chi_{11}^2 + \dots + \chi_{1m}^2) + \sqrt{2n} \left(\frac{n-1}{n} \right) \sqrt{m(m-1)} \mu_2(p) Z. \quad (9)$$

3-2. Case k odd.

For any k odd, By comparing with (5), the limiting distribution of $\sqrt{n}(M_3^{(m)} - \mu_3^m(p))$ will consist of two normal parts and with coefficients that take time to calculate. At this stage of the work, we are satisfied with the fact that the limiting distribution is normal. It does not give us a new reality.

4 Simulation study

In the simulation calculations, we divided the chi-square coefficient on the left side and worked with a variable whose asymptotic distribution is chi-square with one degree of freedom. Since we wanted to find the conditions that lead to a limit distribution according to the chi-square distribution, therefore, we consider the simulation study for the values of n to be even and p to be $1/2$. Table 1 shows the p -values of the Kolmogorov-Smirnov test according to the chi-square distribution with one degree of freedom.

Table 1: The p -values of the K-S test for different values of n and k .

n	$k = 4$	$k = 6$	$k = 8$	$k = 10$	$k = 12$	$k = 14$
30	0.0483	0.0610	0.0612	0.0582	0.0535	0.0432
40	0.1385	0.1047	0.1008	0.0969	0.0891	0.0748
50	0.1035	0.1377	0.1366	0.1316	0.1193	0.1065
60	0.1658	0.1637	0.1657	0.1501	0.1566	0.1444
70	0.1361	0.1846	0.1955	0.1846	0.1728	0.1677
80	0.3045	0.1992	0.2102	0.2078	0.2018	0.1863
90	0.3095	0.2095	0.2300	0.2280	0.2200	0.2136
100	0.2530	0.2393	0.2376	0.2508	0.2412	0.2277
200	0.4737	0.3249	0.3487	0.3266	0.3360	0.3446

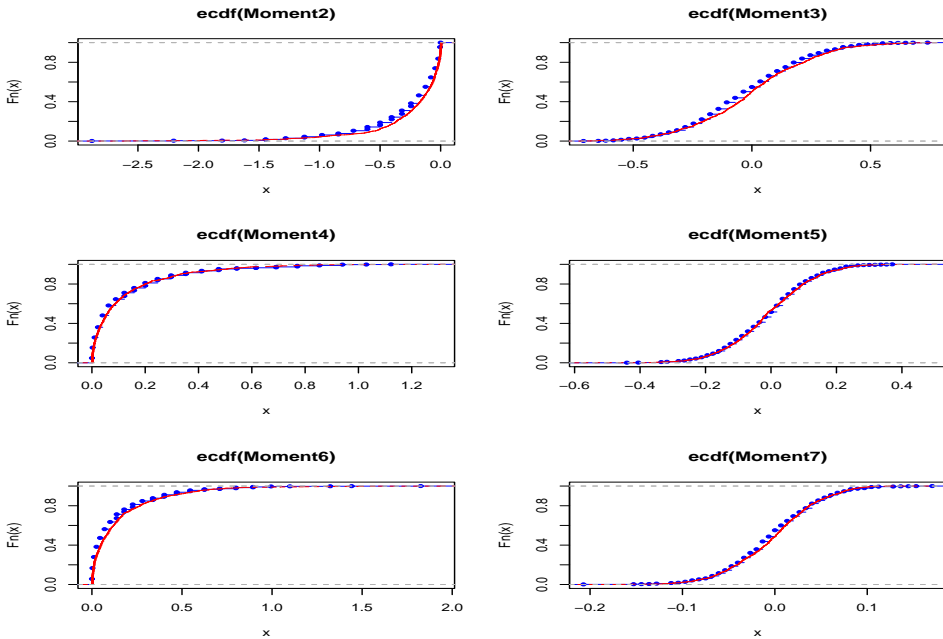


Figure 1: The blue color is the limit distribution and the red color is the result of the simulated empirical distribution function of the central moments in different orders from (5) and (6) with $p = \frac{1}{2}$.

The simulation was done in the R software under the library “dgof”. The results show that too for small random samples, the asymptotic distributions in (3), (5) and (6) are valid. Figure 1 shows the distribution matching well with the chi-square distribution with $n = 100$, and k even. Figure 2, with $m = 2, 3, 4, 5$, $n = 100$, $k = 2$, shows the result (9).

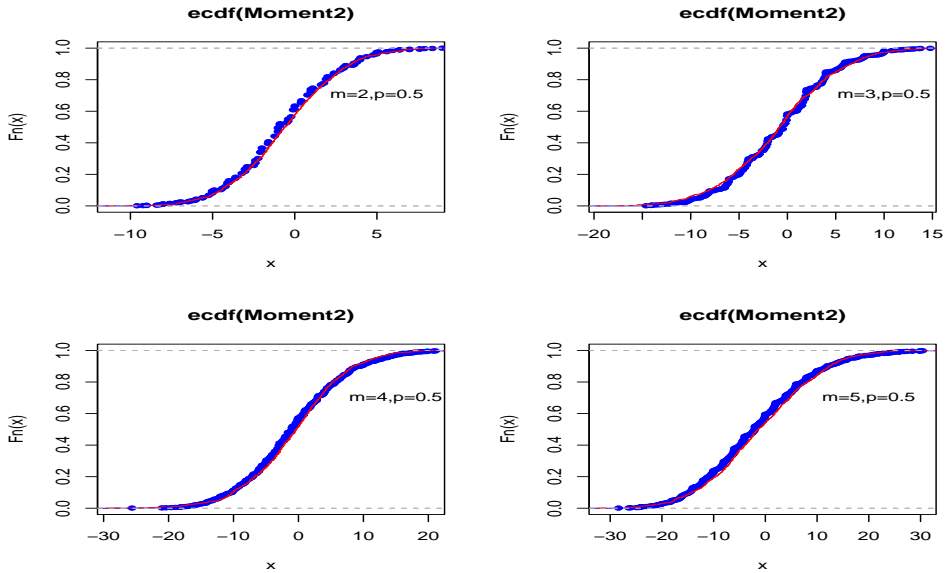


Figure 2: The blue color is the limit distribution and the red color is the result of the simulated empirical distribution function of the central moments in $k = 2$ from (9) with $p = \frac{1}{2}$.

5 Discussion and conclusion

We have seen that in the case of two-point distribution, the point which has a probability of half and the order of the central moment is even, creates a random variable of chi-square factor in the limit distribution of the sample central moment. But in the binomial case, the chi-square factor in the limit distribution was not directly related to the probability of the occurrence of a point in the support, but the chi-square factor appeared depending on the decomposition of the binomial variable into the sum of independent Bernoulli variables. What was important for us is that we were able to show under what conditions the independent combination of chi-square variables appears in the limit distribution. There are problems in generalizing to any discrete random variable using its representation in terms of Bernoulli random variables, and we cannot easily obtain results from the above arguments.

Acknowledgments

The authors of this work are grateful to the reviewers of this article for their accuracy and the suggestions.

Appendix

We have checked four hypothetical models, all of which did not fit well (Figure 3). The figure 4 shows that removing $\hat{L}(p; k, m)$ in the limiting case does not give an accurate

distribution, and the effectiveness of this factor can be seen well.

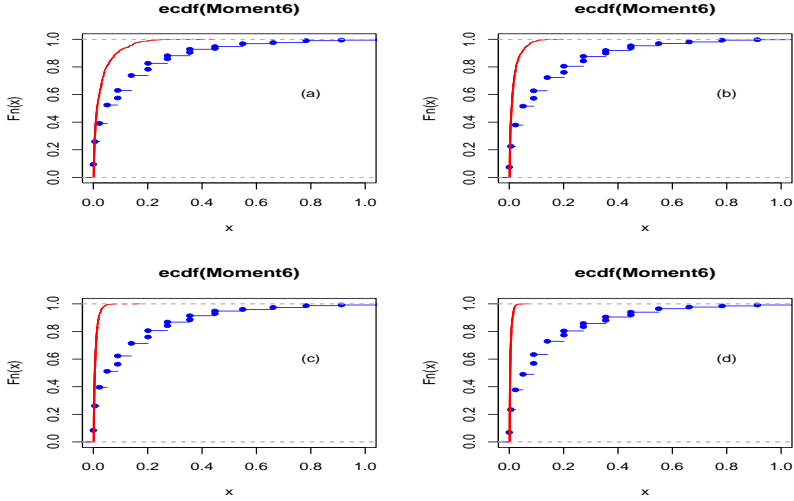


Figure 3: From 4, we have checked four hypothetical models for $n(M_6 - (\frac{1}{2})^6) \xrightarrow{D} Q_i\chi_1^2$: (a) $Q_a = (\frac{1}{2})^5$, (b) $Q_b = (\frac{1}{2})^6$, (c) $Q_c = (\frac{1}{2})^7$, (d) $Q_d = (\frac{1}{2})^8$.

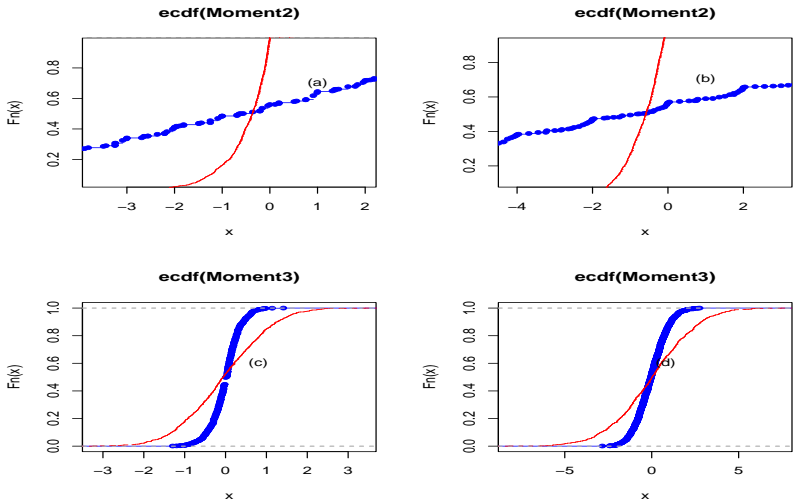


Figure 4: Investigating relation (7) with the idea of how much the $\hat{L}(p; k, m)$ affects the approximation distribution of $\sqrt{n}M_k^{(m)}$ and $n(M_k^{(m)} - \mu_k^{(m)})$: (a) $m = 2, p = \frac{1}{2}, k = 2$, (b) $m = 3, p = \frac{1}{2}, k = 2$, (c) $m = 2, p = \frac{1}{2}, k = 3$, (d) $m = 3, p = \frac{1}{2}, k = 2$.

References

- Abbasi, N. (2008). Asymptotic distribution of UMVUE of r th central moments in Bernoulli distribution. *Australian Journal of Basic and Applied Sciences*, **2**(4):1373–1377.
- Ahle, D.T. (2022). Sharp and simple bounds for the raw moments of the Binomial and Poisson distributions. *Statistics & Probability Letters*, **182**:109–306.
- Casetlla, G. and Berger, R.L. (1990). *Statistical Inference*. Wadsworth & Brooks, Cole, Pacific Grove, CA.
- Denker, M. (1985). *Asymptotic Distribution Theory in Nonparametric Statistics*. Braunschweig: Vieweg.
- Fisher, R.A. (1930). The moments of the distribution for normal samples of measures of departure from normality. *Proceedings of the Royal Society A*, **130**(812):16–28.
- He, S., Zhang, J. and Zhang, S. (2011). Bounding probability of small deviation: A fourth moment approach. *Mathematics of Operations Research*, **35**(1):208–232.
- Heffernan, P.M. (1997). Unbiased estimation of central moments by using U-statistics. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **59**(4):861–863.
- Hoffding, W. (1948). A class of statistics with asymptotically normal distributions. *The Annals of Mathematical Statistics*, **19**(3):293–325.
- Knoblauch, A. (2008). Closed-form expressions for the moments of the Binomial probability distribution. *SIAM Journal on Applied Mathematics*, **69**(1):197–204.
- Lee, A.J. (1990). *U-Statistics: Theory and Practice*. New York: Dekker.
- Leithold, L. (1976). *The Calculus with Analytic Geometry*. Third edition, Harper & Row.
- Nguyen, D. (2021). A probabilistic approach to the moments of binomial random variables and application. *The American Statistician*, **75**(1):101–103.
- Nowakowski, S. (2021). Uniqueness of a median of a Binomial distribution with rational probability. *Advances in Mathematics: Scientific Journal*, **10**(4):1951–1958.