*Chilean Journal of Statistics*
Volume 14, Number 1 June 2023, Pages 48–64
DOI: 10.32372/chjs.14-01-04

STATISTICAL INFERENCE
RESEARCH PAPER

# Critical value functions for likelihood-ratio tests under normality

SHAHZAD MUNIR[1,*], SOHAIL RAZA[2], and KHAWAJA ABDUL HAQ[3]

[1]Crop Reporting Service, Agriculture Department, Punjab, Pakistan
[2]Deakin Business School, Deakin University, Australia
[3]Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, China

(Received: 01 October 2022 · Accepted in final form: 06 June 2023)

**Abstract**

The asymptotic distributions of likelihood ratio tests for normality are unknown; only a few critical values have been tabulated using Monte Carlo simulations. This study aims to develop the critical value functions for these likelihood ratio tests using response surface regressions. In these regressions, the simulated critical values depend on sample size; however, practitioners can easily compute the finite-sample critical values for a number of sample sizes using a hand calculator. An extensive Monte Carlo simulation shows that the proposed critical value functions perform very well for both small and large samples.

**Keywords:** Critical values · Monte Carlo · Normality tests · Response surface.

**Mathematics Subject Classification:** Primary 62F03 · Secondary 62E20 · 65C05 · 62J02.

## 1. INTRODUCTION

For many years, the issue of testing the normality of a sample has gained significant attention. It may not be surprising because the assumption of normality is a source of great convenience for researchers in both theoretical and applied settings. The normality assumption is considered a simplifying assumption in a wide variety of statistical procedures used in estimation, inference, and forecasting. A large number of tests have been proposed in the literature to assess or test the normality of a univariate sample.

Among them, traditional Jarque-Bera (Jarque and Bera, 1987), Shapiro-Wilk (Shapiro and Wilk, 1965), Kolmogorov-Smirnov (Massey Jr, 1951), Anderson-Darling (Anderson and Darling, 1952, 1954), D'Agostino (D'Agostino, 1971), Watson (Watson, 1961), and Kuiper (Kuiper, 1960) are some prominent tests that is being used in empirical works, either due to their well known asymptotic distributions or easy access of critical values; Royston (1992) and Stephens (1974) are the classic references.

*Corresponding author. Email: shahzadmunirm@outlook.com

Whereas likelihood ratio tests (Zhang and Wu, 2005) for normality have received less attention in empirical applications than traditional normality tests. Extensive Monte Carlo simulations (Zhang and Wu, 2005; Torabi et al., 2016) show that the likelihood ratio tests generally perform better than the Jarque-Bera, Shapiro-Wilk, Anderson-Darling, D'Augustino, Kolmogorov-Smirnov, Kuiper, and Watson tests. The main problem with these likelihood ratio tests is that they have unknown asymptotic distributions under the null of normality, for which only a few critical values have been tabulated. However, practitioners require to generate the critical values through Monte Carlo simulations for the sample sizes that are not tabulated. For this reason, these likelihood tests have not attained much attention in practice even though they are much more powerful than their competitors.

Monte Carlo simulations and response surface regressions are commonly used to calculate the finite-sample quantiles or critical values for test statistics that do not have well-known asymptotic distributions or differ considerably from their asymptotic null distributions. For example, Lawford (2005), Wuertz and Katzgraber (2005), and Urzua (1996) conducted extensive Monte Carlo studies to compute the precise quantiles of Jarque-Bera type tests to correct their finite-sample properties. Suárez-Espinosa et al. (2018) and Castro-Kuriss (2011) studied the null distribution of their proposed goodness-of-fit tests for the Pareto distribution and the location-scale distribution, respectively, through Monte Carlo simulations. MacKinnon (2010) tabulated the critical values for Dickey-Fuller unit root tests (Dickey and Fuller, 1979; Fuller, 2009) and Engle-Granger cointegration tests (Engle and Granger, 1987) using response surface regressions and Monte Carlo simulations, as these tests do not follow any standard tabulated distribution, either in finite samples or asymptotically. Moreover, Kiefer and Vogelsang (2005) developed a heteroskedasticity and autocorrelation robust Wald statistic for statistical inference in time series settings. In fixed-$b$ theory, their resulting Wald statistic asymptotically follows a nonstandard distribution; however, they developed critical value functions using response surface regressions and simulated critical values for practical convenience.

This study will fill the research gap in the existing line of literature (Lawford, 2005; Wuertz and Katzgraber, 2005; MacKinnon, 2010) by providing the critical value functions for the statistically powerful likelihood ratio tests (Zhang and Wu, 2005) using Monte Carlo experiments. These experiments are summarized by response surface regressions, in which simulated critical values are regressed on various functions of sample size. The coefficients of these response surface regressions are obtained using the feasible generalized least squares (GLS) method and tabulated in such a way that asymptotic critical values can be read off directly. However, practitioners can easily compute the finite-sample critical values for a number of sample sizes with a hand calculator in empirical applications. To the best of our knowledge, these response surface regressions have not been used to develop the critical value functions for these likelihood ratio tests (Zhang and Wu, 2005) for normality. An extensive Monte Carlo simulation shows that our proposed critical value functions give a very accurate size of the test for both small and large samples. Additionally, we provide an empirical exercise for illustration purposes using average yields (kilograms per acre) of wheat, rice, sugarcane, and canola crops from 37, 33, 32, and 26 districts of Punjab (Pakistan), respectively.

The organization of the rest of the study is as follows. Section 2 summarizes the setup of the likelihood ratio tests for normality, Section 3 presents Monte Carlo experiments for estimating the critical value functions and evaluating their finite-sample properties, Section 4 includes an empirical exercise for illustration purposes and Section 5 contains the concluding remarks, limitations of this study, and future work. The `R` codes are provided in Appendix A, while the average yields (kilograms per acre) of sugarcane, rice, wheat, and canola crops are given in Appendix B.

## 2. Setup

Let $X$ be an independent and identically distributed (IID) random variable with continuous cumulative distribution function (CDF) $G(x)$, and $X_1, \ldots, X_t$ be a random sample of size $T$ from $X$ with order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(t)}$. Zhang and Wu (2005) developed the likelihood ratio tests based on empirical distribution function (EDF) to test the null hypothesis that $H_0 : G(x) = G_0(x)$ against the alternative hypothesis that $H_1 : G(x) \neq G_0(x)$, where $G_0(x) = \Phi\{(x-\mu)/\sigma\}$ is a hypothetical distribution function of a normal population with mean $\mu$ and variance $\sigma^2$. It is important to note that the mean $\mu$ and variance $\sigma^2$ are unknown parameters and to be estimated by their sample counterparts i.e. $\overline{X} = (1/T)\sum_{t=1}^{T} X_t$ and $s^2 = (1/[T-1])\sum_{t=1}^{T}(X_t - \overline{X})^2$, respectively. Their statistics $Z_K$, $Z_A$, and $Z_C$ are defined as:

$$Z_K = \max_{1 \leqslant t \leqslant T}\left(\left[t-\frac{1}{2}\right]\log\left(\frac{t-\frac{1}{2}}{TG_0\Big(X_{(t)}\Big)}\right) + \left[T-t+\frac{1}{2}\right]\log\left(\frac{T-t+\frac{1}{2}}{T\Big[1-G_0\Big(X_{(t)}\Big)\Big]}\right)\right),$$

$$Z_{\mathrm{A}} = -\sum_{t=1}^{T}\left[\frac{\log\left(G_0(X_{(t)})\right)}{T-t+\frac{1}{2}} + \frac{\log\left(1-G_0(X_{(t)})\right)}{t-\frac{1}{2}}\right],$$

$$Z_{\mathrm{C}} = \sum_{t=1}^{T}\left[\log\left(\frac{G_0(X_{(i)})^{-1}-1}{[T-\frac{1}{2}]/[t-\frac{3}{4}]-1}\right)\right]^2.$$

We will reject the null of normality for large values of the three test statistics. Unfortunately, the $Z_K$, $Z_A$, and $Z_C$ test statistics do not follow any standard tabulated distribution under the null of normality, either in finite samples or asymptotically. Only a few simulated critical values have been tabulated by Zhang and Wu (2005) to evaluate the finite sample properties of $Z_K$, $Z_A$, and $Z_C$ in relation to other traditional normality tests.

These three likelihood ratio tests can be used to test $H_0$ in situations where $G_0(x)$ is fully specified (meaning $\mu$ and $\sigma^2$ are known). For these scenarios, simulated critical values have been tabulated by Zhang (2001) and Zhang (2002). In this study, we will only estimate the critical value functions for cases where $G_0(x)$ is completely unspecified (that is, $\mu$ and $\sigma^2$ are unknown), as practitioners typically encounter this situation in empirical work or applications. Furthermore, to enhance the power of the tests, it is recommended by Zhang and Wu (2005) to estimate the mean $\mu$ and variance $\sigma^2$ from the data using their sample counterparts, even when they are known.

## 3. Monte Carlo Experiments

It is well-documented that the finite-sample null distributions of the Jarque-Bera test and its modified version (MJarque-Bera) introduced by Urzua (1996) differ considerably from their asymptotic distributions. Using their asymptotic critical values, even for sufficiently large samples, will distort the true size of the test and may lead to erroneous judgments in empirical applications. Therefore, to correct the finite-sample size distortions, critical value functions for the Jarque-Bera and MJarque-Bera tests were estimated by Lawford (2005), Wuertz and Katzgraber (2005), and many others using response surface regressions. Additionally, response surface regressions have been employed by MacKinnon (2010) to approximate the asymptotic distributions of Dickey-Fuller unit root tests (Dickey and Fuller, 1979; Fuller, 2009) and Engle-Granger cointegration tests (Engle and Granger, 1987), as they possess nonstandard asymptotic distributions.

Like Dickey-Fuller unit root tests and Engle-Granger cointegration tests, the likelihood ratio tests $Z_K$, $Z_A$, and $Z_C$ do not follow any standard tabulated distribution under the null of normality. However, in the same spirit as MacKinnon (2010), we estimate the critical value functions for $Z_K$, $Z_A$, and $Z_C$ for the percentage points 99.5%, 99%, 98%, 95%, 90%, and 80%. Using simulation codes and the `MonteCarlo` package (Leschinski, 2019) in the `R` software, we generate 20,000 realizations of $Z_K$, $Z_A$, and $Z_C$ under the null of normality with $\mu = 0$ and $\sigma^2 = 1$ for each sample size $T \in [5, 2500]$. We calculate the 0.5%, 1%, 2%, 5%, 10%, and 20% critical values as the $q(20,000)th$, where $q \in \{99.5\%, 99\%, 98\%, 95\%, 90\%, 80\%\}$, largest values of $Z_K$, $Z_A$, and $Z_C$. Then, we regress these simulated critical values on various functions of a sample size to calculate the critical value functions. It is important to note that choosing the correct functional form for response surface regressions plays a critical role in obtaining good estimates (MacKinnon, 2010). We choose the following form of response surface regression to fit the critical value functions after considerable experimentation, and motivated by the quantile approximations developed by MacKinnon (2010) in the context of the Dickey-Fuller unit root tests and Engle-Granger cointegration tests:

$$cv^q(T)_l = a_0 + a_1 T^{-\frac{1}{2}} + a_2 T^{-1} + a_3 T^{-\frac{3}{2}} + a_4 T^{-2} + a_5 T^{-\frac{5}{2}} + a_6 T^{-3} + e_t, \text{ for } l = \{Z_K, Z_A, Z_C\}, \tag{3.1}$$

where $a_0$ represents the asymptotic critical value of a test as $T^{-1/2}$, $T^{-1}$, $T^{-3/2}$, $T^{-2}$, $T^{-5/2}$, and $T^{-3}$ tend to zero when $T$ approaches infinity. The coefficients of $T^{-1/2}$, $T^{-1}$, $T^{-3/2}$, $T^{-2}$, $T^{-5/2}$, and $T^{-3}$ determine the shape of the response surface for finite samples, while $e_t$ represents the error term. The dependent variable $cv^q(T)_l$ stands for the simulated critical value of the $l^{th}$ test statistic for the $qth$ percentage point with sample size $T$. The residuals $e_t$ in regression given in Equation (3.1) exhibit heteroskedasticity. However, the feasible generalized least squares (GLS) technique is employed to estimate the unknown parameters $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, $a_5$, and $a_6$ of the response surface regression given in Equation (3.1) as follows.

As a first step, the dependent variable $cv^q(T)_l$ is regressed on $T^{-1/2}$, $T^{-1}$, $T^{-3/2}$, $T^{-2}$, $T^{-5/2}$, and $T^{-3}$, where $T \in [5, 2500]$, to calculate the ordinary least squares (OLS) residuals $\widehat{e}_t$. These residuals $\widehat{e}_t$ are used to compute the weights for feasible GLS using the following regression:

$$\log(\widehat{e}_t^2) = \beta_0 + \beta_1 \log(T) + \eta, \tag{3.2}$$

where $\eta$ is the error term. We run the regression given in Equation (3.2) and use the reciprocals of the square roots of exponentials of its fitted values as weights for feasible GLS estimation of regression given in Equation (3.1). Therefore, the final estimated critical value functions for the three test statistics are as follows:

$$\widehat{cv}^q(T)_{Z_K} = \widehat{a}_0 + \widehat{a}_1 T^{-\frac{1}{2}} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-\frac{3}{2}} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-\frac{5}{2}} + \widehat{a}_6 T^{-3}, \tag{3.3}$$

$$\widehat{cv}^q(T)_{Z_A} = \widehat{a}_0 + \widehat{a}_1 T^{-\frac{1}{2}} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-\frac{3}{2}} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-\frac{5}{2}} + \widehat{a}_6 T^{-3}, \tag{3.4}$$

$$\widehat{cv}^q(T)_{Z_C} = \widehat{a}_0 + \widehat{a}_1 T^{-\frac{1}{2}} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-\frac{3}{2}} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-\frac{5}{2}} + \widehat{a}_6 T^{-3}, \tag{3.5}$$

where $\widehat{a}_0$, $\widehat{a}_1$, $\widehat{a}_2$, $\widehat{a}_3$, $\widehat{a}_4$, $\widehat{a}_5$, and $\widehat{a}_6$ are the feasible GLS estimates of $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, $a_5$, and $a_6$, respectively. These feasible GLS estimates are presented in Tables 1, 2, and 3 for $Z_K$, $Z_A$, and $Z_C$, respectively. These are the final findings of this paper. All regression coefficients were found significant at the 1% level. Each test statistic table required approximately 378 minutes to complete. The procedure for calculating the estimated critical values for any sample size $T$ is described in the following examples.

**Example 3.1:** When $T = 20$, the estimated critical values for $Z_K$ at significance level 0.05 is $\widehat{cv}^{0.95}(20)_{Z_K} = 4.787 - 58.38(20)^{-1/2} + 591.7(20)^{-1} - 3763(20)^{-3/2} + 13330(20)^{-2} - 23900(20)^{-5/2} + 16820(20)^{-3} = 1.3132$.

**Example 3.2:** When $T = 20$, the estimated critical value for $Z_A$ at significance level 0.05 is $\widehat{cv}^{0.95}(20)_{Z_A} = 3.289 + 0.1309(20)^{-1/2} + 8.559(20)^{-1} - 49.07(20)^{-3/2} + 152.9(20)^{-2} - 272.4(20)^{-5/2} + 194.7(20)^{-3} = 3.4519$.

**Example 3.3:** When $T = 20$, the estimated critical value for $Z_C$ at significance level 0.05 is $\widehat{cv}^{0.95}(20)_{Z_C} = 33.58 - 456.6(20)^{-1/2} + 5237(20)^{-1} - 36210(20)^{-3/2} + 135200(20)^{-2} - 251000(20)^{-5/2} + 181100(20)^{-3} = 8.8153$.

There are two important points to consider. First, each realization was performed by generating independent samples from a standard normal distribution. Consequently, the simulated critical values were independent of each other. Since these critical values served as the unit of observation for the response surface regression; therefore, the correlation across observations is not an issue.

Second, it is common practice to use OLS or GLS (MacKinnon, 2010; Lawford, 2005; Wuertz and Katzgraber, 2005; Kiefer and Vogelsang, 2005) for estimating the response surface regressions in order to develop the critical value functions. However, we can also utilize robust estimation methods such as, median regression (Koenker and Bassett, 1978; Koenker and d'Orey, 1987), least median squares estimates (Rousseeuw, 1984; Siegel, 1982), M-estimates (Huber, 1981), MM-estimates (Yohai, 1987), robust and efficient weighted least squares estimates (Gervini and Yohai, 2002), R-estimates Jaeckel (1972) to obtain more robust regression coefficient estimates for critical value functions. In our specific case, we found that feasible GLS estimates are much more precise than the median regression, M-estimates, and MM-estimates estimates. Based on our analysis, we can conclude that the feasible GLS approach is more appropriate than robust regression methods for estimating Equation (3.1).

Table 1. Asymptotic critical value response function coefficients for $Z_K$

| Quantiles | $\widehat{a}_0$ | $\widehat{a}_1$ | $\widehat{a}_2$ | $\widehat{a}_3$ | $\widehat{a}_4$ | $\widehat{a}_5$ | $\widehat{a}_6$ |
|---|---|---|---|---|---|---|---|
| 99.50% | 6.961 | -74.27 | 648.6 | -3809 | 13090 | -23220 | 16260 |
| 99% | 6.317 | -70.11 | 640 | -3837 | 13230 | -23420 | 16350 |
| 98% | 5.663 | -65.38 | 623.9 | -3832 | 13340 | -23690 | 16560 |
| 95% | 4.787 | -58.38 | 591.7 | -3763 | 13330 | -23900 | 16820 |
| 90% | 4.107 | -52.11 | 550.9 | -3598 | 12940 | -23420 | 16590 |
| 80% | 3.406 | -44.86 | 492.3 | -3297 | 12050 | -22040 | 15740 |

Note: Given the percentage point and the value of $T$ , the critical value of $Z_K$ is $\widehat{cv}^q(T)_{Z_K} = \widehat{a}_0 + \widehat{a}_1 T^{-1/2} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-3/2} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-5/2} + \widehat{a}_6 T^{-3}$.

To evaluate the finite sample performance of our estimated critical value functions, we compute the empirical size of each test statistic $Z_K$, $Z_A$, and $Z_C$. It is the probability of rejecting $H_0$ when it is actually true. For this purpose, we generate 20,000 random samples from a standard normal distribution of size $T$=20, 25, 50, 75, 100, 150, 200, 350, 450, 500, 600, 700, 800, 900, and 1000. Then we compute the values of the three test statistics and calculate their chances of rejecting $H_0$ by comparing them with their corresponding estimated critical $\widehat{cv}^q(T)_l$ values, where $l \in \{Z_K, Z_A, Z_C\}$, at significance level $\alpha = 0.005, 0.01, 0.02, 0.05, 0.10$, and 0.20. These chances are referred to the empirical size or the level of significance.

Table 2. Asymptotic critical value response function coefficients for $Z_A$

| Quantiles | $\widehat{a}_0$ | $\widehat{a}_1$ | $\widehat{a}_2$ | $\widehat{a}_3$ | $\widehat{a}_4$ | $\widehat{a}_5$ | $\widehat{a}_6$ |
|---|---|---|---|---|---|---|---|
| 99.50% | 3.288 | 0.2067 | 14.1 | -83.01 | 299.9 | -599.9 | 453.3 |
| 99% | 3.289 | 0.1771 | 12.54 | -73.45 | 255.9 | -496.9 | 370.3 |
| 98% | 3.289 | 0.152 | 10.95 | -64.37 | 219 | -418.4 | 312 |
| 95% | 3.289 | 0.1309 | 8.559 | -49.07 | 152.9 | -272.4 | 194.7 |
| 90% | 3.289 | 0.108 | 7.037 | -41.57 | 127.7 | -223.3 | 158.8 |
| 80% | 3.289 | 0.09393 | 5.313 | -32.29 | 95.23 | -158.6 | 108.5 |

Note: Given the percentage point and the value of $T$ , the critical value of $Z_A$ is $\widehat{cv}^q(T)_{Z_A} = \widehat{a}_0 + \widehat{a}_1 T^{-1/2} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-3/2} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-5/2} + \widehat{a}_6 T^{-3}$.

Table 3. Asymptotic critical value response function coefficients for $Z_C$

| Quantiles | $\widehat{a}_0$ | $\widehat{a}_1$ | $\widehat{a}_2$ | $\widehat{a}_3$ | $\widehat{a}_4$ | $\widehat{a}_5$ | $\widehat{a}_6$ |
|---|---|---|---|---|---|---|---|
| 99.50% | 50.1 | -516.1 | 5346 | -36600 | 135800 | -249300 | 177700 |
| 99% | 44.75 | -518 | 5559 | -37710 | 138700 | -253600 | 180400 |
| 98% | 39.79 | -500.5 | 5527 | -37590 | 138600 | -254300 | 181600 |
| 95% | 33.58 | -456.6 | 5237 | -36210 | 135200 | -251000 | 181100 |
| 90% | 28.94 | -410.1 | 4810 | -33580 | 126200 | -235200 | 170200 |
| 80% | 24.21 | -355.8 | 4250 | -29860 | 112500 | -210100 | 152200 |

Note: Given the percentage point and the value of $T$ , the critical value for $Z_C$ is $\widehat{cv}^q(T)_{Z_C} = \widehat{a}_0 + \widehat{a}_1 T^{-1/2} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-3/2} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-5/2} + \widehat{a}_6 T^{-3}$.

The empirical rejection rates of $Z_K$, $Z_A$, and $Z_C$, are given in Tables 4, 5, and 6 respectively. Each test statistic table required approximately 20 minutes to complete using simulation codes and the MonteCarlo package (Leschinski, 2019) in the *R* software. The Monte Carlo results show that the estimated critical value functions give a very accurate size of the tests for each percentage point and sample size. However, we can conclude that our proposed critical value functions work very well for both small and large samples.

All simulations and computations were performed on a laptop with an AMD Ryzen 7 2700U 2.20 GHz processor and 8GB of RAM, running `R` version 3.6.2 (R Core Team , 2023) under Microsoft Windows 10 Home version 22H2. Additionally, we used an `R` package called **`writexl`** (Ooms, 2021) to export data frames from `R` to Microsoft Excel 2010, while the package **`readxl`** (Wickham and Bryan, 2019) was used to import Excel files in the `.xls` format into `R`. The simulation codes are provided in Appendix A.

## 4. Empirical Illustration

For illustration purposes, we consider the average yields (kilograms per acre) of sugarcane, rice, wheat, and canola crops for the year 2021-22 of 37, 33, 32, and 26 districts of Punjab (Pakistan), respectively. These average yields are obtained from Crop Reporting Service, Agriculture Department, Punjab (Pakistan); and provided in Table 8 (see, Appendix B). We utilize $Z_K$, $Z_A$, and $Z_C$ to test the normality of the data at a significance level $\alpha = 0.05$. For this purpose, we calculate these three test statistics using the DistributionTest (Ning Cui, 2020) package in the *R* software. The critical values are computed from Equations (3.3), (3.4), and (3.5) for $T = 37, 33, 32,$ and 26 at $\alpha = 0.05$.

Table 4. Empirical size of the $Z_K$ test statistic using $\widehat{cv}^q(T)_{Z_K}$

| | $\alpha$ | | | | | |
|---|---|---|---|---|---|---|
| $T$ | 0.005 | 0.01 | 0.02 | 0.05 | 0.1 | 0.2 |
| 20 | 0.00523 | 0.01144 | 0.02247 | 0.05523 | 0.11167 | 0.21249 |
| 25 | 0.00520 | 0.01104 | 0.02164 | 0.05460 | 0.10884 | 0.21260 |
| 50 | 0.00483 | 0.00991 | 0.01936 | 0.04883 | 0.09864 | 0.19643 |
| 75 | 0.00464 | 0.00974 | 0.01940 | 0.04859 | 0.09718 | 0.19351 |
| 100 | 0.00470 | 0.00996 | 0.01982 | 0.04904 | 0.09834 | 0.19525 |
| 150 | 0.00547 | 0.01012 | 0.02035 | 0.04961 | 0.09973 | 0.19976 |
| 200 | 0.00505 | 0.01029 | 0.02042 | 0.05064 | 0.10097 | 0.20174 |
| 350 | 0.00517 | 0.01014 | 0.02023 | 0.05095 | 0.10117 | 0.20187 |
| 450 | 0.00515 | 0.01002 | 0.01983 | 0.05089 | 0.10119 | 0.20106 |
| 500 | 0.00482 | 0.01028 | 0.01958 | 0.05042 | 0.10120 | 0.20013 |
| 600 | 0.00522 | 0.01000 | 0.02008 | 0.05080 | 0.10007 | 0.20000 |
| 700 | 0.00470 | 0.01013 | 0.01954 | 0.05061 | 0.10054 | 0.19964 |
| 800 | 0.00500 | 0.00966 | 0.01970 | 0.05074 | 0.09925 | 0.19992 |
| 900 | 0.00477 | 0.01030 | 0.01963 | 0.04974 | 0.09854 | 0.19943 |
| 1000 | 0.00481 | 0.01001 | 0.01986 | 0.04951 | 0.10004 | 0.19816 |

Note: $\alpha$ is the significance level and $T$ is the sample size.

Table 5. Empirical size of the $Z_A$ test statistic using $\widehat{cv}^q(T)_{Z_A}$

| | $\alpha$ | | | | | |
|---|---|---|---|---|---|---|
| $T$ | 0.005 | 0.01 | 0.02 | 0.05 | 0.1 | 0.2 |
| 20 | 0.00514 | 0.00970 | 0.01975 | 0.04999 | 0.10092 | 0.20161 |
| 25 | 0.00490 | 0.01043 | 0.02024 | 0.05003 | 0.09931 | 0.20203 |
| 50 | 0.00535 | 0.01000 | 0.01975 | 0.05006 | 0.09977 | 0.20194 |
| 75 | 0.00514 | 0.00969 | 0.01962 | 0.04944 | 0.10016 | 0.20228 |
| 100 | 0.00490 | 0.00930 | 0.01980 | 0.04968 | 0.10021 | 0.20216 |
| 150 | 0.00523 | 0.00940 | 0.01950 | 0.04987 | 0.10098 | 0.20420 |
| 200 | 0.00497 | 0.00953 | 0.01917 | 0.05016 | 0.10094 | 0.20467 |
| 350 | 0.00578 | 0.00917 | 0.01945 | 0.05031 | 0.10309 | 0.20706 |
| 450 | 0.00537 | 0.00970 | 0.02000 | 0.04914 | 0.10299 | 0.20910 |
| 500 | 0.00523 | 0.00980 | 0.01918 | 0.04974 | 0.10313 | 0.21406 |
| 600 | 0.00535 | 0.00919 | 0.01960 | 0.04964 | 0.10510 | 0.21363 |
| 700 | 0.00541 | 0.00898 | 0.01902 | 0.05021 | 0.10415 | 0.21535 |
| 800 | 0.00564 | 0.00856 | 0.01944 | 0.04848 | 0.10535 | 0.21592 |
| 900 | 0.00541 | 0.00904 | 0.01861 | 0.04875 | 0.10622 | 0.21863 |
| 1000 | 0.00556 | 0.00883 | 0.01850 | 0.04880 | 0.10678 | 0.21984 |

Note: $\alpha$ is the significance level and $T$ is the sample size.

The results of applying these statistics are given in Table 7. The first three columns contain the values of the test statistics $Z_K$, $Z_A$, and $Z_C$, respectively, while the last three columns report the values of $\widehat{cv}^{0.95}(T)_{Z_K}$, $\widehat{cv}^{0.95}(T)_{Z_A}$, and $\widehat{cv}^{0.95}(T)_{Z_C}$ respectively. We reject the null hypothesis of normality if the calculated statistic exceeds its critical value. The results indicate that the $Z_K$ test, $Z_A$ test, and $Z_C$ test reject the null of normality for the average yield of the wheat crop. In contrast, these tests do not provide sufficient evidence to reject the null of normality for the average yields of rice, sugarcane, and canola crops.

Table 6. Empirical size of the $Z_C$ test statistic using $\widehat{cv}^q(T)_{Z_C}$

| | $\alpha$ | | | | | |
|---|---|---|---|---|---|---|
| $T$ | 0.005 | 0.01 | 0.02 | 0.05 | 0.1 | 0.2 |
| 20 | 0.00512 | 0.01122 | 0.02198 | 0.05925 | 0.10811 | 0.22302 |
| 25 | 0.00526 | 0.01131 | 0.02247 | 0.05794 | 0.11045 | 0.22134 |
| 50 | 0.00480 | 0.00956 | 0.01913 | 0.04749 | 0.09579 | 0.19305 |
| 75 | 0.00470 | 0.00960 | 0.01892 | 0.04723 | 0.09513 | 0.19143 |
| 100 | 0.00488 | 0.00933 | 0.01950 | 0.04881 | 0.09510 | 0.19437 |
| 150 | 0.00515 | 0.01058 | 0.01942 | 0.04930 | 0.09862 | 0.19815 |
| 200 | 0.00522 | 0.01010 | 0.01996 | 0.05039 | 0.10146 | 0.20054 |
| 350 | 0.00476 | 0.00954 | 0.01988 | 0.05095 | 0.10040 | 0.20241 |
| 450 | 0.00492 | 0.01015 | 0.02072 | 0.04944 | 0.09963 | 0.20169 |
| 500 | 0.00505 | 0.00987 | 0.02012 | 0.05046 | 0.10073 | 0.20286 |
| 600 | 0.00499 | 0.00999 | 0.01982 | 0.05102 | 0.10010 | 0.19930 |
| 700 | 0.00507 | 0.00973 | 0.01982 | 0.05083 | 0.09955 | 0.19897 |
| 800 | 0.00499 | 0.00994 | 0.01971 | 0.04958 | 0.09999 | 0.20010 |
| 900 | 0.00490 | 0.00993 | 0.01959 | 0.05002 | 0.09957 | 0.19851 |
| 1000 | 0.00478 | 0.01001 | 0.02033 | 0.04976 | 0.09986 | 0.19981 |

Note: $\alpha$ is the significance level and $T$ is the sample size.

Table 7. Application of $Z_K$, $Z_C$, and $Z_C$ to average yield of wheat, rice, and sugarcane crops

| Variables | $Z_K$ | $Z_A$ | $Z_C$ | $\widehat{cv}^{0.95}(T)_{Z_K}$ | $\widehat{cv}^{0.95}(T)_{Z_A}$ | $\widehat{cv}^{0.95}(T)_{Z_C}$ |
|---|---|---|---|---|---|---|
| Wheat | 2.5631 | 3.5310 | 18.0533 | 1.6605 | 3.4066 | 11.3588 |
| Rice | 1.0931 | 3.3474 | 4.6087 | 1.5927 | 3.4146 | 10.8498 |
| Sugarcane | 0.7225 | 3.3414 | 4.9444 | 1.5745 | 3.4168 | 10.7131 |
| Canola | 0.6615 | 3.3067 | 4.6229 | 1.4536 | 3.4004 | 9.8126 |

Note: The critical values for $Z_K$, $Z_A$ and $Z_C$ are calculated using the critical value function $\widehat{cv}^q(T)_l = \widehat{a}_0 + \widehat{a}_1 T^{-1/2} + \widehat{a}_2 T^{-1} + \widehat{a}_3 T^{-3/2} + \widehat{a}_4 T^{-2} + \widehat{a}_5 T^{-5/2} + \widehat{a}_6 T^{-3}$ for $l = \{Z_K, Z_A, Z_C\}$, where $T$ takes values 37, 33, 32, and 26 for wheat, rice, sugarcane, and canola respectively, at the 5% significance level.

## 5. Conclusions, Limitations, and Future Research

For practical convenience, we have developed the critical value functions for likelihood ratio tests (Zhang and Wu, 2005) for normality using response surface regressions and Monte Carlo experiments. These functions tabulate the simulated critical values as a function of sample size, allowing practitioners to easily calculate the finite-sample critical values for any sample size in their empirical applications. Our extensive Monte Carlo simulations have demonstrated that the proposed critical value functions yield reasonably accurate test sizes for both small and large samples. Therefore, the proposed critical value functions offer practitioners the opportunity to enhance the accuracy of their statistical analyses and make meaningful conclusions from data by utilizing statistically powerful tests $Z_K$, $Z_A$, and $Z_C$. However, it is important to note that our proposed critical value functions are specifically designed for the test statistics $Z_K$, $Z_A$, and $Z_C$ under the null of normality. When dealing with other tests or hypotheses, it becomes necessary to develop new critical value functions under their specific settings. One can effectively utilize response surface regressions to estimate critical value functions for cases where test statistics have nonstandard asymptotic distributions (MacKinnon, 2010).

As part of our future work, we intend to explore the estimation of critical value functions for the tests of normality proposed by Torabi et al. (2016) and other researchers. Additionally, we will investigate the application of robust regressions to approximate the finite-sample distributions of test statistics in cases where the asymptotic distributions are unknown.

## Appendix A

This appendix provides the R codes used in the paper. All simulations and computations were performed on a laptop with an AMD Ryzen 7 2700U 2.20 GHz processor and 8GB of RAM, running R version 3.6.2 (R Core Team , 2023) under Microsoft Windows 10 Home version 22H2.

### Codes for Generating Simulated Critical Values

**Code for the $Z_A$ test statistics.**

```
rm(list=ls())
library(MonteCarlo)
library(writexl)
############################################################
shahzad<-function(n,loc,scale){
  x=rnorm(n,loc,scale)
  n1=length(x)
  Data=sort(x)
  MU=mean(Data)
  SDV=sd(Data)
  F=pnorm(Data,MU,SDV)
  i=1:n1
  part1=(log(F))/(n1-i+0.5)
  part2=(log(1-F))/(i-0.5)
  Comb=part1+part2
  Z_A_Stat=-sum(Comb)
  return(list("Z_A_Stat"=Z_A_Stat))
}
ns=seq(5,2500,1)
loc_grid=0
scale_grid=1
param_list=list("n"=ns,"loc"=loc_grid, "scale"=scale_grid)
s=MonteCarlo(func=shahzad,nrep=20000,param_list=param_list,ncpus=1,
             time_n_test = TRUE)
f<-MakeFrame(s)
qaunt=function(x){quantile(x,probs=c(0.80,0.90,0.95,0.98,0.99,0.995))}
qaunt.1=function(x){quantile(x,probs=c(0.2,0.1,0.05,0.02,0.01,0.005))}
uperlimit=aggregate(abs(f$Z_A_Stat),list(f$n),qaunt)
lowerlimit=aggregate(abs(f$Z_A_Stat),list(f$n),qaunt.1)
uper=data.frame(n=uperlimit$Group.1,uperlimit$x)
lower=data.frame(n=lowerlimit$Group.1,lowerlimit$x)
```

**Code for the $Z_C$ test statistics.**

```
rm(list=ls())
library(MonteCarlo)
library(writexl)
#########################################################################
shahzad<-function(n,loc,scale){
  x=rnorm(n,loc,scale)
```

```
  n1=length(x)
  Data=sort(x)
  MU=mean(Data)
  SDV=sd(Data)
  F=pnorm(Data,MU,SDV)
  i=1:n1
  F_Inv=1/F
  Uper=F_Inv-1
  Lower=((n1-0.5)/(i-0.75))-1
  lg=(log(Uper/Lower))^2
  Z_C_Stat=sum(lg)
  return(list("Z_C_Stat"=Z_C_Stat))
}
ns=seq(5,2500,1)
loc_grid=0
scale_grid=1
param_list=list("n"=ns,"loc"=loc_grid, "scale"=scale_grid)
s=MonteCarlo(func=shahzad,nrep=20000,param_list=param_list,ncpus=1,
             time_n_test = TRUE)
f<-MakeFrame(s)
qaunt=function(x){quantile(x,probs=c(0.80,0.90,0.95,0.98,0.99,0.995))}
qaunt.1=function(x){quantile(x,probs=c(0.2,0.1,0.05,0.02,0.01,0.005))}
uperlimit=aggregate(abs(f$Z_C_Stat),list(f$n),qaunt)
lowerlimit=aggregate(abs(f$Z_C_Stat),list(f$n),qaunt.1)
uper=data.frame(n=uperlimit$Group.1,uperlimit$x)
lower=data.frame(n=lowerlimit$Group.1,lowerlimit$x)
```

**Code for the $Z_K$ test statistics.**

```
rm(list=ls())
library(MonteCarlo)
library(writexl)
########################################################################
shahzad<-function(n,loc,scale){
  x=rnorm(n,loc,scale)
  n1=length(x)
  Data=sort(x)
  MU=mean(Data)
  SDV=sd(Data)
  F=pnorm(Data,MU,SDV)
  i=1:n1
  part=(i-0.5)*log((i-0.5)/(n1*(F)))+(n1-i+0.5)*log((n1-i+0.5)/(n1*(1-F)))
  Z_K_Stat=max(part)
  return(list("Z_K_Stat"=Z_K_Stat))
}
ns=seq(5,2500,1)
loc_grid=0
scale_grid=1
param_list=list("n"=ns,"loc"=loc_grid, "scale"=scale_grid)
s=MonteCarlo(func=shahzad,nrep=20000,param_list=param_list,ncpus=1,
             time_n_test = TRUE)
f<-MakeFrame(s)
qaunt=function(x){quantile(x,probs=c(0.80,0.90,0.95,0.98,0.99,0.995))}
qaunt.1=function(x){quantile(x,probs=c(0.2,0.1,0.05,0.02,0.01,0.005))}
uperlimit=aggregate(abs(f$Z_K_Stat),list(f$n),qaunt)
lowerlimit=aggregate(abs(f$Z_K_Stat),list(f$n),qaunt.1)
uper=data.frame(n=uperlimit$Group.1,uperlimit$x)
lower=data.frame(n=lowerlimit$Group.1,lowerlimit$x)
```

## Codes for Size of the Test

**Code for the $Z_A$ test statistics.**

```
rm(list=ls())
library(MonteCarlo)
library(writexl)
####################################################################
ZA_Crit=function(n1,alpha){
  if      (alpha==0.005){3.288+0.2067*(1/sqrt(n1))+14.1*(1/n1)-83.01
                         *(1/n1^1.5)+299.9*(1/n1^2)-599.9*(1/n1^2.5)
                         +453.3*(1/n1^3)}
  else if (alpha==0.01){3.289+0.1771*(1/sqrt(n1))+12.54*(1/n1)-73.45
                         *(1/n1^1.5)+255.9*(1/n1^2)-496.9*(1/n1^2.5)+
                          370.3*(1/n1^3)}
  else if (alpha==0.02){3.289+0.1520*(1/sqrt(n1))+10.95*(1/n1)-64.37
                         *(1/n1^1.5)+219*(1/n1^2)-418.4*(1/n1^2.5)+
                           312*(1/n1^3)}
  else if (alpha==0.05){3.289+0.1309*(1/sqrt(n1))+8.559*(1/n1)-49.07
                         *(1/n1^1.5)+152.9*(1/n1^2)-272.4*(1/n1^2.5)+
                           194.70*(1/n1^3)}
  else if (alpha==0.1){3.289+0.1080*(1/sqrt(n1))+7.037*(1/n1)-41.57
                       *(1/n1^1.5)+127.7*(1/n1^2)-223.3*(1/n1^2.5)+
                        158.8*(1/n1^3)}
  else if (alpha==0.2){3.289+0.09393*(1/sqrt(n1))+5.313*(1/n1)-32.29
                        *(1/n1^1.5)+95.23*(1/n1^2)-158.6*(1/n1^2.5)+
                         108.5*(1/n1^3)}
}
#####################################################################
shahzad<-function(n,alpha,loc,scale){
  x=(rnorm(n,loc,scale))
  n1=length(x)
  Data=sort(x)
  MU=mean(Data)
  SDV=sd(Data)
  F=pnorm(Data,MU,SDV)
  i=1:n1
  part1=(log(F))/(n1-i+0.5)
  part2=(log(1-F))/(i-0.5)
  Comb=part1+part2
  Z_A_Stat=-sum(Comb)
  Z_A_Stat=Z_A_Stat>ZA_Crit(n1,alpha)
  return(list("Z_A_Stat"=Z_A_Stat))
}
ns=c(20,25,50,75,100,150,200,350,450,500,600,700,800,900,1000)
loc_grid=0
scale_grid=1
alpha= c(0.005,0.01,0.02,0.05,0.1,0.2)
param_list=list("n"=ns,"alpha"=alpha, "loc"=loc_grid, "scale"=scale_grid)
s=MonteCarlo(func=shahzad,nrep=20000,param_list=param_list,ncpus=1,
             time_n_test = TRUE)
f<-MakeFrame(s)
Size_ZA=aggregate(f$Z_A_Stat,list(f$n,f$alpha),mean)
Size_ZA
```

**Code for the $Z_C$ test statistics.**

```
rm(list=ls())
library(writexl)
library(MonteCarlo)
#################################################################
ZC_Crit=function(n1,alpha){
```

```
  if      (alpha==0.005){50.1-516.1*(1/sqrt(n1))+5346*(1/n1)-
                         36600*(1/n1^1.5)+135800*(1/n1^2)-249300*
                         (1/n1^2.5)+177700*(1/n1^3)}
  else if (alpha==0.01) {44.75-518*(1/sqrt(n1))+5559*(1/n1)-
                         37710*(1/n1^1.5)+138700*(1/n1^2)-253600*
                         (1/n1^2.5)+180400*(1/n1^3)}
  else if (alpha==0.02) {39.79-500.5*(1/sqrt(n1))+5527*(1/n1)-
                         37590*(1/n1^1.5)+138600*(1/n1^2)-254300*
                         (1/n1^2.5)+181600*(1/n1^3)}
  else if (alpha==0.05) {33.58-456.6*(1/sqrt(n1))+5237*(1/n1)-
                          36210*(1/n1^1.5)+135200*(1/n1^2)-251000*
                          (1/n1^2.5)+181100*(1/n1^3)}
  else if (alpha==0.1)  {28.94-410.1*(1/sqrt(n1))+4810*(1/n1)-
                         33580*(1/n1^1.5)+126200*(1/n1^2)-235200*
                         (1/n1^2.5)+170200*(1/n1^3)}
  else if (alpha==0.2)  {24.21-355.8*(1/sqrt(n1))+4250*(1/n1)
                          -29860*(1/n1^1.5)+112500*(1/n1^2)-210100*
                          (1/n1^2.5)+152200*(1/n1^3)}
}
################################################################
shahzad<-function(n,alpha,loc,scale){
  x=(rnorm(n,loc,scale))
  n1=length(x)
  Data=sort(x)
  MU=mean(Data)
  SDV=sd(Data)
  F=pnorm(Data,MU,SDV)
  i=1:n1
  F_Inv=1/F
  Uper=F_Inv-1
  Lower=((n1-0.5)/(i-0.75))-1
  lg=(log(Uper/Lower))^2
  Z_C_Stat=sum(lg)
  Z_C_Stat=Z_C_Stat>ZC_Crit(n1,alpha)
  return(list("Z_C_Stat"=Z_C_Stat))
}
ns=c(20,25,50,75,100,150,200,350,450,500,600,700,800,900,1000)
loc_grid=0
scale_grid=1
alpha= c(0.005,0.01,0.02,0.05,0.1,0.2)
param_list=list("n"=ns,"alpha"=alpha, "loc"=loc_grid, "scale"=scale_grid)
s=MonteCarlo(func=shahzad,nrep=20000,param_list=param_list,ncpus=1,
             time_n_test = TRUE)
f<-MakeFrame(s)
Size_ZC=aggregate(f$Z_C_Stat,list(f$n,f$alpha),mean)
Size_ZC
```

**Code for the $Z_C$ test statistics.**

```
library(writexl)
library(MonteCarlo)
###############################################################
ZK_Crit=function(n1,alpha){
  if      (alpha==0.005){6.961-74.27*(1/sqrt(n1))+648.6*(1/n1)
                         -3809*(1/n1^1.5)+13090*(1/n1^2)-23220*
                         (1/n1^2.5)+16260*(1/n1^3)}
  else if (alpha==0.01) {6.317-70.11*(1/sqrt(n1))+640*(1/n1)
                         -3837*(1/n1^1.5)+13230*(1/n1^2)-23240*
                         (1/n1^2.5)+16350*(1/n1^3)}
  else if (alpha==0.02) {5.663-65.38*(1/sqrt(n1))+623.9*(1/n1)
                         -3832*(1/n1^1.5)+13340*(1/n1^2)-23690*
                         (1/n1^2.5)+16560*(1/n1^3)}
```

```
  else if (alpha==0.05) {4.787-58.38*(1/sqrt(n1))+591.7*(1/n1)
                        -3763*(1/n1^1.5)+13330*(1/n1^2)-23900*
                        (1/n1^2.5)+16820*(1/n1^3)}
  else if (alpha==0.1)  {4.107-52.11*(1/sqrt(n1))+550.9*(1/n1)
                        -3598*(1/n1^1.5)+12940*(1/n1^2)-23420*
                        (1/n1^2.5)+16590*(1/n1^3)}
  else if (alpha==0.2)  {3.406-44.86*(1/sqrt(n1))+492.3*(1/n1)
                        -3297*(1/n1^1.5)+12050*(1/n1^2)-22040*
                        (1/n1^2.5)+15740*(1/n1^3)}
}
##############################################################
shahzad<-function(n,alpha,loc,scale){
  x=(rnorm(n,loc,scale))
  n1=length(x)
  Data=sort(x)
  MU=mean(Data)
  SDV=sd(Data)
  F=pnorm(Data,MU,SDV)
  i=1:n1
  part=(i-0.5)*log((i-0.5)/(n1*(F)))+(n1-i+0.5)*log((n1-i+0.5)/(n1*(1-F)))
  Z_K_Stat=max(part)
  Z_K_Stat=Z_K_Stat>ZK_Crit(n1,alpha)
  return(list("Z_K_Stat"=Z_K_Stat))
}
ns=c(20,25,50,75,100,150,200,350,450,500,600,700,800,900,1000)
loc_grid=0
scale_grid=1
alpha= c(0.005,0.01,0.02,0.05,0.1,0.2)
param_list=list("n"=ns,"alpha"=alpha, "loc"=loc_grid, "scale"=scale_grid)
s=MonteCarlo(func=shahzad,nrep=20000,param_list=param_list,ncpus=1,
             time_n_test = TRUE)
f<-MakeFrame(s)
Size_Z_K=aggregate(f$Z_K_Stat,list(f$n,f$alpha),mean)
Size_Z_K
```

## Code for Numerical Illustration

```
rm(list=ls())
library(readxl)
library(DistributionTest) #Jin Zhang(2005) Likelihood-ratio
###########################Data###################################
Crops_Data_For_Normality <- read_excel("C:/Users/hp/Desktop/
CropsDataForNormality.xlsx")
wheat=c(Crops_Data_For_Normality$`Wheat In Kg`)
sugarcan=c(na.omit(Crops_Data_For_Normality$`Sugarcane In Kg`))
rice=c(na.omit(Crops_Data_For_Normality$`Rice In Kg`))
Canola=c(na.omit(Crops_Data_For_Normality$`Canola in Kg`))
##########################Test Statistics##########################
Z_K_test_W = zk.test(wheat,"norm",N=0)
Z_K_test_R = zk.test(rice,"norm",N=0)
Z_K_test_S = zk.test(sugarcan,"norm",N=0)
Z_K_test_C = zk.test(Canola,"norm",N=0)

Z_A_test_W = za.test(wheat,"norm",N=0)
Z_A_test_R = za.test(rice,"norm",N=0)
Z_A_test_S = za.test(sugarcan,"norm",N=0)
Z_A_test_C = za.test(Canola,"norm",N=0)
```

```
Z_C_test_W = zc.test(wheat,"norm",N=0)
Z_C_test_R = zc.test(rice,"norm",N=0)
Z_C_test_S = zc.test(sugarcan,"norm",N=0)
Z_C_test_C = zc.test(Canola,"norm",N=0)

Z_K_test_W$statistic
Z_K_test_R$statistic
Z_K_test_S$statistic
Z_K_test_C$statistic

Z_A_test_W$statistic
Z_A_test_R$statistic
Z_A_test_S$statistic
Z_A_test_C$statistic

Z_C_test_W$statistic
Z_C_test_R$statistic
Z_C_test_S$statistic
Z_C_test_C$statistic
```

## Appendix B

This Appendix provides the data on average yields (kilograms per acre) of sugarcane, rice, wheat, and canola crops used in Section 4. These average yields are obtained from Crop Reporting Service, Agriculture Department, Punjab (Pakistan). We can download these average yields from their website https://crs.agripunjab.gov.pk/reports.

Table 8. Average yield (kilograms per acre) of sugarcane, rice, wheat, and canola

| Sr. No. | Sugarcane | Rice | Wheat | Canola |
|---|---|---|---|---|
| 1 | 29760 | 970 | 686.8 | 784 |
| 2 | 23680 | 822.4 | 787.6 | 571.2 |
| 3 | 27160 | 807.2 | 736.8 | 644.4 |
| 4 | 29360 | 1010.8 | 797.2 | 594.4 |
| 5 | 27640 | 835.2 | 677.2 | 674.8 |
| 6 | 27400 | 873.6 | 1132 | 924.4 |
| 7 | 30760 | 1004.8 | 837.2 | 932.4 |
| 8 | 30000 | 968.8 | 1090.8 | 710 |
| 9 | 16960 | 1029.6 | 1068.4 | 494.4 |
| 10 | 22080 | 742 | 1325.6 | 635.2 |
| 11 | 21560 | 884 | 1351.2 | 500 |
| 12 | 14720 | 606.8 | 1282.8 | 790.4 |
| 13 | 23440 | 649.2 | 1388.8 | 688.4 |
| 14 | 24280 | 877.6 | 829.6 | 728 |
| 15 | 24480 | 970.8 | 1132.8 | 722.8 |
| 16 | 25400 | 803.6 | 1009.6 | 589.2 |
| 17 | 24680 | 973.2 | 917.2 | 804.8 |
| 18 | 22720 | 852.8 | 1202 | 951.2 |
| 19 | 31200 | 912.8 | 1332.8 | 799.6 |
| 20 | 23800 | 1192 | 1225.6 | 848.4 |
| 21 | 22320 | 973.2 | 1378 | 996.8 |
| 22 | 27000 | 987.2 | 1390 | 794 |
| 23 | 30400 | 866.8 | 1265.6 | 673.2 |
| 24 | 28000 | 791.2 | 1527.6 | 940.4 |
| 25 | 27960 | 958.8 | 1379.6 | 793.2 |
| 26 | 29320 | 1007.2 | 1394.4 | 819.2 |
| 27 | 25800 | 833.6 | 1469.2 | |
| 28 | 30800 | 901.6 | 1522.8 | |
| 29 | 32280 | 1259.6 | 1493.2 | |
| 30 | 25960 | 1154 | 1410.4 | |
| 31 | 35440 | 862.4 | 1228.8 | |
| 32 | 26920 | 800 | 1211.6 | |
| 33 | | 974.8 | 1284.8 | |
| 34 | | | 1403.6 | |
| 35 | | | 1439.2 | |
| 36 | | | 1500.8 | |
| 37 | | | 1441.6 | |

Source: https://crs.agripunjab.gov.pk/reports

Author contributions Conceptualization, S.M.; Methodology, S.M.; Software,S.M.; Validation, S.M.; Formal analysis, S.M.; Investigation, S.M.; Data Curation,S.M., S.R.; Visualization, S.M.; Supervision, S.M., S.R.; Writing-original draft, S.M., S.R., K.A.H.; Writing-review and editing, S.M., S.R., K.A.H. All authors have read and agreed to the published version of the manuscript.

Acknowledgements The authors thank the editors, associate editor, anonymous referees, and Andrew Pua for their valuable comments to improve this manuscript.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Conflicts of interest The authors declare no conflict of interest.

## References

Anderson, T.W. and Darling, D.A., 1952. Asymptotic theory of certain goodness of fit criteria based on stochastic processes. The Annals of Mathematical Statistics. 23, 193-212.

Anderson, T.W. and Darling, D.A., 1954. A test of goodness of fit. Journal of The American Statistical Association, 49, 765-769.

Castro-Kuriss, C., 2011. On a goodness-of-fit test for censored data from a location-scale distribution with applications. Chilean Journal of Statistics, 2, 115-136.

Cizek, P. and Visek, J.A., 2000. Least trimmed squares. XPLORE: Application Guide, 49-64.

D'Agostino, R., 1971. An omnibus test of normality for moderate and large sample sizes. Biometrika, 58, 1-348.

Dickey, D.A. and Fuller, W.A, 1978. Distribution of the estimators for autoregressive time series with a unit root. Journal of The American Statistical Association, 74, 427-431.

Engle, R.F. and Granger, C.W., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica, 251-276.

Fuller, W.A., 2009. Introduction to statistical time series. Wiley, New York.

Gervini, D. and Yohai, V.J., 2002. A class of robust and fully efficient regression estimators. The Annals of Statistics. 30, 583-616.

Huber, P.J., 1981. Robust statistics. Wiley, New York.

Jaeckel, L.A., 1972. Estimating regression coefficients by minimizing the dispersion of the residuals. The Annals Of Mathematical Statistics. 1449-1458.

Jarque, C. and Bera, A., 1987. A test for normality of observations and regression residuals. International Statistical Review. 55, 163-172.

Kuiper, N.H., 1960. Tests concerning random points on a circle. Proceedings of the Royal Netherlands Academy of Arts and Sciences A, 63, 38-47.

Koenker, R. and Bassett Jr, G., 1978. Regression quantiles. Econometrica, 46, 33-50.

Koenker, R. and D'Orey, V., 1987 Algorithm AS 229: Computing regression quantiles. Applied Statistics, 383-393.

Kiefer, N.M. and Vogelsang, T.J., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory. 21(6), 1130-1164.

Lawford, S., 2005. Finite-sample quantiles of the Jarque–Bera test. Applied Economics Letters, 12, 351-354.

Leschinski, C.H., 2019 MonteCarlo: automatic parallelized Monte Carlo simulations. Downloadable from: https://CRAN.R-project.org/package=MonteCarlo.

Massey Jr., F., 1951 The Kolmogorov-Smirnov test for goodness of fit. Journal of The American Statistical Association, 46, 68-78.

MacKinnon, J.G., 2010. Critical values for cointegration tests. Working Paper, Department of Economics, Queen.

Ning, C. and Maoyuan, Z., 2020. Distribution test: powerful goodness-of-fit tests based on the likelihood ratio. Downloadable from: https://CRAN.R-project.org/package=DistributionTest.

Ooms, J., 2021. writexl: Export Data Frames to Excel 'xlsx' Format. R package version 1.4.0. Downloadable from: https://CRAN.R-project.org/package=writexl.

Rousseeuw, P.J., 1984. Least median of squares regression. Journal of The American Statistical Association, 79, 871-880.

Royston, P., 1992. Approximating the Shapiro-Wilk W-test for non-normality. Statistics and Computing, 2, 117-119.

R Core Team, 2023. R: A language and environment for statistical computing. Statistics and Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Shapiro, S. and Wilks, M., 1965. An analysis of variance test for normality (complete samples). Biometrika, 52, 591-611.

Stephens, M.A., 1974. EDF statistics for goodness of fit and some comparisons. Journal of The American Statistical Association, 69, 730-737.

Siegel, A.F., 1982. Robust regression using repeated medians. Biometrika, 69, 242-244.

Suárez-Espinosa, J., Villaseñor-Alva, J.A., Hurtado-Jaramillo, A., and Pérez-Rodríguez, P., 2018. A goodness of fit test for the Pareto distribution. Chilean Journal of Statistics, 9(1), 33-46.

Torabi, H., Montazeri, N.H., and Grané, A., 2016. A test for normality based on the empirical distribution function. Statistics and Operations Research Transactions, 40, 55-88.

Urzúa, C.M., 1996. On the correct use of omnibus tests for normality. Economics Letters, 53(3), 247-251.

Watson, G.S., 1961. Goodness-of-fit tests on a circle. Biometrika. 48, 109-114.

Wuertz, D. and Katzgraber, H.G., 2005. Precise finite-sample quantiles of the Jarque-Bera adjusted Lagrange multiplier test. ArXiv Preprint Math/0509423.

Wickham, H. and Bryan, J., 2019. readxl: Read Excel Files. R package version 1.3.1. Download from: https://CRAN.R-project.org/package=readxl.

Yohai, V.J., 1987. High breakdown-point and high efficiency robust estimates for regression. The Annals of Statistics, 642-656.

Zhang, J., 2001. Powerful goodness-of-fit and multi-sample tests. York University, Toronto.

Zhang, J., 2002. Powerful goodness-of-fit tests based on the likelihood ratio. Journal of The Royal Statistical Society B, 64, 281-294.

Zhang, J. and Wu, Y., 2005. Likelihood-ratio tests for normality. Computational Statistics and Data Analysis, 49, 709-721.