Testing Special Cases of the GB2 Distribution: Comment and an Extension

William Griffiths
University of Melbourne

R. Carter Hill
Louisiana State University

8 December 2024

Abstract

We extend earlier work (Griffiths and Hill, 2025) by pointing out that a score test developed by Prentice (1975) for testing whether a normal distribution is an adequate representation of the log of a GB2 random variable is suitable for testing whether a lognormal distribution is a reasonable special case of the GB2 distribution. Prentice's test involves testing the skewness and kurtosis of the logs of the observations. We use a simulation study to investigate its power and size.

JEL codes: D31, C46, C12

Keywords: income distribution, score test, test power

Corresponding Author:
William Griffiths
Department of Economics
University of Melbourne
Vic 3010
Australia
wegrif@unimelb.edu.au

1. Introduction

In Griffiths and Hill (2025), we reported the results of a thorough investigation into the properties of hypothesis tests for four special cases of the GB2 distribution: the Singh-Maddala, Dagum, beta-2, and Fisk distributions. Prompted by a referee, we briefly mentioned that testing for the special case lognormal distribution was problematic because violating a boundary condition means that the usual properties of the likelihood ratio test no longer hold.[1] In this note, we report the outcome of a further investigation where we discovered a test that overcomes the problem. This test is a score (Lagrange multiplier) test developed by Prentice (1975) who uses a distribution that equates to the log of a GB2 distribution, and a reparameterization that overcomes the boundary condition problem. In what follows, we describe how the distribution considered by Prentice relates to the GB2 distribution, describe the score test statistic, and investigate its size and power.

2. Distributions and Test

The probability density function for a random variable *X* that has a GB2 distribution is

$$f(x) = \frac{ax^{ap-1}}{b^{ap}B(p,q)\left[1+\left(x/b\right)^{a}\right]^{p+q}}$$

where *a*, *b*, *p,* and *q* are positive parameters, and $B(p,q)$ is the beta function. To place this distribution in the context of that considered by Prentice (1975), we note that one way of generating observations on a GB2 random variable is via the relationship (Cummins et al. 1990, p. 259)

$$X = b\left(U_1/U_2\right)^{1/a}$$

where U_1 and U_2 are independent gamma random variables, with scale parameters one, and shape parameters *p* and *q*, respectively. Let $Y = \ln(X)$. It follows that we can represent observations on *Y* as

$$\begin{aligned} Y = \ln(X) &= \ln(b) + \frac{1}{a}\left[\ln\left(U_1\right) - \ln\left(U_2\right)\right] \\ &= \ln(b) + \frac{1}{a}\left[\ln\left(p\right) - \ln\left(q\right)\right] + \frac{1}{a}\left[\ln\left(\frac{U_1}{p}\right) - \ln\left(\frac{U_2}{q}\right)\right] \end{aligned}$$

Prentice considers the distribution of *Y*, parameterized as

---

[1] See McDonald and Xu (1992) for details.

$$Y = \mu + \sigma\left[\ln\left(\frac{U_1}{p}\right) - \ln\left(\frac{U_2}{q}\right)\right]$$

Thus, writing his parameters as functions of those for the GB2 specification, we have

$$\mu = \ln(b) + \frac{1}{a}\left[\ln(p) - \ln(q)\right]$$

and $\sigma = 1/a$.

Prentice notes that the limiting distribution of Y as p and q approach infinity is a normal distribution. He develops another parameterization that avoids the boundary testing problem, enabling him to create a score test for the null hypothesis that Y has a normal distribution. This score test involves two asymptotically independent standard normal random variables and involves testing Y for skewness and excess kurtosis.

Let $\hat{\mu} = N^{-1}\sum_{i=1}^{N} y_i$ and $\hat{\sigma}^2 = N^{-1}\sum_{i=1}^{N}\left(y_i - \hat{\mu}\right)^2$ be the sample mean and variance for Y calculated from a sample of size N, and let $w_i = \left(y_i - \hat{\mu}\right)/\hat{\sigma}$ be the standardized observations. Prentice shows that, under the null hypothesis that Y has a normal distribution,

$$Z_1 = \frac{1}{\sqrt{6N}}\sum_{i=1}^{N} w_i^3 \qquad \text{and} \qquad Z_2 = \frac{1}{\sqrt{24N}}\left(\sum_{i=1}^{N} w_i^4 - 3N\right)$$

have limiting distributions that are independent standard normal. Thus, we can use the following chi-square statistic with 2 degrees of freedom to test for the normality of Y and the lognormality of X.

$$\begin{aligned}\chi^2_{(2)} &= \left(\frac{1}{\sqrt{6N}}\sum_{i=1}^{N} w_i^3\right)^2 + \left[\frac{1}{\sqrt{24N}}\left(\sum_{i=1}^{N} w_i^4 - 3N\right)\right]^2 \\ &= N\left(\frac{\hat{\mu}_3^2}{6} + \frac{\left(\hat{\mu}_4 - 3\right)^2}{24}\right)\end{aligned}$$

where $\hat{\mu}_3 = N^{-1}\sum_{i=1}^{N} w_i^3$ and $\hat{\mu}_4 = N^{-1}\sum_{i=1}^{N} w_i^4$. This chi-square statistic is the same as the Lagrange multiplier statistic suggested by Jarque and Bera (1987) for testing normality against an alternative that belongs to the family of Pearson distributions. We can test whether the lognormal distribution is a reasonable special case of the GB2 distribution by simply taking logs of the observations and jointly testing their skewness and kurtosis.

3. A Monte Carlo Study

To investigate the power and size of the test, we generated observations *X* from a GB2 distribution with $a = 2, b = 1$ and values of *p* up to 70 in increments of 0.1, and with $p = q$. We used sample sizes of 400, 1000, and 4000, and 100,000 replications. Using a 5% significance level, we computed the proportion of replications where the $\chi^2_{(2)}$ statistic rejected the null hypothesis that $Y = \ln(X)$ has a normal distribution. We display test power in Figure 1 for values up to $p = q = 40$, a point at which all three sample sizes lead to a rejection probability closely approximating the nominal test size of 0.05[2]. The power functions behave as expected, declining monotonically as $p = q$ increases, and with the larger sample sizes, having greater power. It is concerning that the power can be low. For example, for $N = 400$ and $p = q = 4$, the power is only slightly above 0.2.

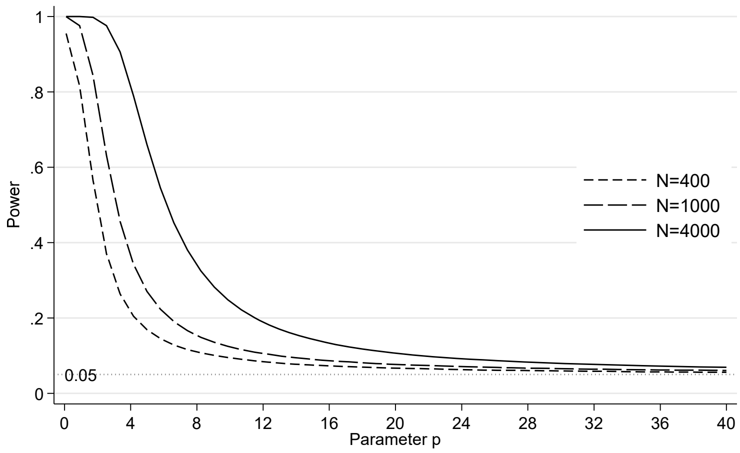

Figure 1. Power of score test for lognormal distribution when $p = q$ for sample sizes 400, 1000 and 4000.

To investigate further, we estimated power with increasing values of *p*, but with *q* kept fixed at two values; first $q = 2$, then $q = 4$, scenarios where the null hypothesis is always false. The results for $q = 2$ and $q = 4$ for sample sizes $N = 400$ and $N = 1000$ are displayed in Figures 2 and 3, respectively, for *p* up to 20. Power is at a minimum when $p = q$, and depends heavily on how much *p* deviates from *q*. This outcome occurs because the log GB2 distribution is symmetric when $p = q$, implying the

[2] We used Gauss Version 24, Aptech Systems, Inc. (2023) www.aptech.com to conduct the simulations. We created the graphs using Stata 18.0 (StataCorp LLC) with command *lpoly* to slightly smooth the curves.

skewness component of the chi-square statistic will be relatively small.[3] Griffiths and Hill (2025) found that the powers of the Wald, likelihood ratio, and Lagrange multiplier tests for the Fisk distribution, where the null hypothesis is $p = q = 1$, are also at a minimum when $p = q$. Comparing the solid (or dashed) curves in Figures 2 and 3 reveals a dramatic difference between the power functions for $q = 2$ and $q = 4$, in terms of the minimum power achieved and the rate of increase in power after its minimum is reached. Having a lower power when $q = 4$ is consistent with the declining power for large (p,q) observed in Figure 1. As expected, a larger sample size leads to higher power.

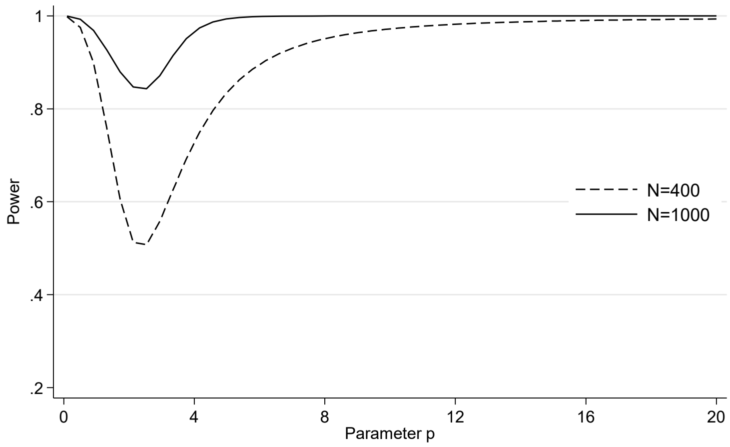

Figure 2. Power of score test for lognormal distribution for $q = 2$ and sample sizes 400 and 1000.

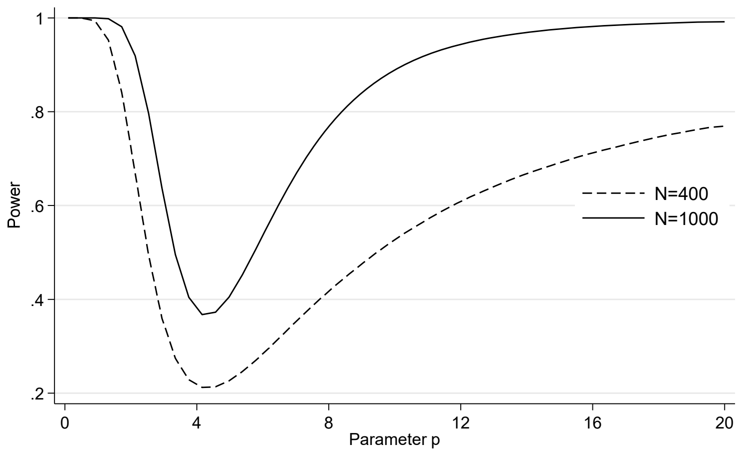

Figure 3. Power of score test for lognormal distribution for $q = 4$ and sample sizes 400 and 1000.

[3] We are using the term log GB2 distribution to refer to the distribution of $Y = \ln(X)$, not the distribution of e^X.

4. Concluding remarks

Jointly testing the significance of skewness and kurtosis of the logs of the observations provides a simple and convenient way of assessing whether the more parsimonious lognormal distribution offers a suitable alternative to the GB2 distribution. If p or q is small, the test will have large power. The power will decline if both p and q are large. However, given that we often estimate income distributions with tens of thousands of observations, the chance of incorrectly concluding that a lognormal distribution is adequate will be small.

References

Cummins, J. D., Dionne, G., McDonald, J. B. and Pritchett, B. M. (1990), "Applications of the GB2 family of distributions in modelling insurance loss processes," *Insurance: Mathematics and Economics,* 9, 257-272.

Griffiths, W. and R. C. Hill (2025), "Testing special cases of the GB2 distribution," *Journal of Income Distribution*, in press.

Jarque, C. M. and A. Bera (1987), "A test for normality of observations and regression residuals," *International Statistical Review*, 55(2), 163-172.

McDonald, J. B. and Y. Xu (1992), "An empirical investigation of the likelihood ratio test when the boundary condition is violated", *Communications in Statistics: Simulation and Computation*, 21(3), 879-892.

Prentice, R.L (1975), "Discrimination among some parametric models", *Biometrika* 62(3), 607-614.