KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.1084

“Unit root tests considering initial values and a concise method for computing powers”

Kohtaro Hitomi, Jianwei Jin, Keiji Nagai,
Yoshihiko Nishiyama and Junfan Tao

November 2022

KYOTO UNIVERSITY

KYOTO, JAPAN

Unit root tests considering initial values and a concise method for computing powers*

Kohtaro Hitomi[†], Jianwei Jin[‡], Keiji Nagai[§], Yoshihiko Nishiyama[¶], and Junfan Tao[‖]

November 18, 2022

Abstract

The Dickey-Fuller (DF) unit root tests are widely used in empirical studies on economics. In the local-to-unity asymptotic theory, the effects of initial values vanish as the sample size grows. However, for a small sample size, the initial value will affect the distribution of the test statistics. When ignoring the effect of the initial value, the left-sided unit root test sets the critical value smaller than it should be. Therefore, the size and power of the test become smaller. This paper investigates the effect of the initial value for the DF test (including the t test). Limiting approximations of the DF test statistics are the ratios of two integrals which are represented via a one-dimensional squared Bessel process. We derive the joint density of the squared Bessel process and its integral, enabling us to compute this ratio's distribution. For independent normal errors, the exact distribution of the Dickey-Fuller coefficient test statistic is obtained using the Imhof (1961) method for non-central chi-squared distribution. Numerical results show that when the sample size is small, the limiting distributions of the DF test statistics with initial values fit well with the exact or simulated distributions. We transform the DF test with respect to a local parameter into the test for a shift in the location parameter of normal distributions. As a result, a concise method for computing the powers of DF tests is derived.

Keywords: Dickey-Fuller tests, Squared Bessel process, joint density, powers approximated by normal distribution, exact distribution

JEL Classification: C12, C22, C46

*Acknowledgments: This research was supported by the 2018 Kyoto University Institute of Economic Research Joint Usage and Research Center Project "Asymptotic Theory of Sequential Tests and Estimation of Unit Root Processes" and JSPS KAKENHI Grants Numbers JP19H01473, JP20K01589, JP21K01422, JP22K20133. The authors sincerely appreciate these supports.

[†]Kyoto Institute of Technology. E-mail address: hitomi@kit.ac.jp

[‡]Yokohama National University. E-mail address: jin-jianwei-cd@ynu.jp

[§]Yokohama National University. E-mail address: nagai-keiji-hs@ynu.ac.jp

[¶]Institute of Economic Research, Kyoto University. E-mail address: nishiyama@kier.kyoto-u.ac.jp

[‖]Institute of Economic Research, Kyoto University. E-mail address: tao.junfan.7j@kyoto-u.ac.jp

1 Introduction

Autoregressive processes are basic models in time series analysis. From a practical point of view, the problem of a unit root in a time series is critical in statistical inference. Many procedures for unit root tests have been proposed since the mid-1970s. The most widely used unit root tests are the Dickey-Fuller tests (Dickey and Fuller (1979)). For some essential contributions related to unit root tests, see White (1958), Bobkoski (1983), Cavanagh (1985), Chan (1988), Chan and Wei (1987), Phillips (1987a,b) and Abadir (1995a,b).

In the DF test, the null hypothesis is that the autoregressive process has a unit root, while the alternative hypothesis is that the process is stationary. The limiting distribution of the DF test statistic under the null hypothesis is a non-standard distribution. Researchers have considered the asymptotic properties of the DF test. As the sample size increases, the influence of the initial value decreases, and a state closer to the limit is achieved. However, the effect of the initial value should be addressed for a small sample size. Intuitively, under a fixed sample size, the larger the initial value, the larger the observed Fisher information, which should yield a better estimator of the autoregressive coefficient.

To the best of our knowledge, the primary literature does not consider the effect of initial values in small samples. This study establishes methods for calculating the size and power of the DF test with initial values. We also propose a concise method for calculating power. We derive the limiting joint density of the numerator and denominator in the DF coefficient (or t) statistic and compute the cumulative distribution function (CDF) of the test statistic. We also derive an exact distribution of the DF coefficient test statistic. In the asymptotic theory, the error terms are strict stationary and ergodic martingale differences with mean 0 and finite variance σ^2. On the other hand, the error terms are more strongly assumed to be independent normal random variables in the exact distribution. The exact distribution of the Dickey-Fuller coefficient test statistic is obtained using the Imhof (1961) method for non-central chi-squared distribution by eliminating the linear term containing an initial value from the quadratic polynomial.

We discover a concise method for calculating the power of the DF unit root test. Some simulations revealed that at a significance level of 5%, the power of the DF coefficient test for local parameter θ is almost the same as the power of the test of null hypothesis $N(0,1)$ against alternative hypothesis $N(\delta,1)$ by setting $\delta = 0.23\theta$;

$$\Psi_\theta(\Psi_0^{-1}(\alpha)) \approx \Phi(\Phi^{-1}(\alpha) - \delta).$$

where Ψ_θ is the CDF of the DF coefficient test statistic for local parameter θ and Φ is the CDF of $N(0,1)$. By equating the above approximation and defining an implicit function $\delta = \delta(\theta)$, the problem of the unit root test could be transformed into the problem of testing the shift δ in the location parameter from $N(0,1)$. Finding the third-order Taylor approximation to the function $\delta(\theta)$ and substituting it into the right-hand in the above equation shows that the approximation is perfect.

2 Limiting Approximation of Dickey-Fuller Test Statistic

For the AR(1) process:

$$x_n = \beta x_{n-1} + \varepsilon_n \quad (n = 1, 2, ..., N),$$

one can localize the regression coefficient by $\beta = 1 + \theta/N$. We consider a unit root test with respect to the null and alternative hypotheses;

$$H_0 : \theta = 0 \text{ vs } H_1 : \theta < 0.$$

The estimators of β, σ^2, θ and the standard error of $\hat{\beta}_N$ (s.e. $\left(\hat{\beta}_N\right)$) are

$$\hat{\beta}_N = \frac{\sum_{n=1}^{N} x_n x_{n-1}}{\sum_{n=1}^{N} x_{n-1}^2}, \quad \hat{\sigma}_N^2 = \frac{1}{N} \sum_{n=1}^{N} \left(x_n - \hat{\beta}_N x_{n-1}\right)^2$$

$$\hat{\theta}_N = N(\hat{\beta}_N - 1), \ \text{s.e.}\left(\hat{\beta}_N\right) = \frac{\hat{\sigma}_N}{\sqrt{\sum_{n=1}^{N} x_{n-1}^2}}.$$

The Dickey–Fuller test employs $\hat{\theta}_N$ as the coefficient test statistic or $(\hat{\beta}_N - 1)/s.e.\left(\hat{\beta}_N\right)$ as t test statistic . We obtain their limiting approximations and an exact distribution of the coefficient test statistic θ. In the asymptotic theory, $\varepsilon_1, \varepsilon_2, \cdots$ are strict stationary and ergodic martingale differences with mean 0 and finite variance σ^2. On the other hand, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N$ is a sequence of independent normal random variables with mean 0 and variance σ^2 in the exact distribution theory.

We make the following asymptotic assumption to investigate the effect of the initial value x_0 in the AR(1) process (1). Letting X_0 be an L_2 random variable, we assume that as $Nc \to \infty$,

$$x_0/\sqrt{N} \xrightarrow{p} X_0$$

where $\xrightarrow{p}$ represents convergence in probability. Of course, when considering small-sample theory and simulations, we simply set $X_0 = x_0/\sqrt{N}$, since c is fixed at a constant value.

As N tends to ∞, $x_{\lfloor Nt \rfloor}/\sqrt{N}$ converges in distribution to an Ornstein–Uhlenbeck process X_t;

$$x_{\lfloor Nt \rfloor}/\sqrt{N} \Rightarrow X_t = X_0 + \theta \int_0^t X_s ds + \sigma W_t \tag{1}$$

where W_s is a standard Brownian motion and $x_0/\sqrt{N}$ converges in probability to X_0.

The stochastic process satisfying the following stochastic integral is called a δ-dimensional squared Bessel process with drift 2θ started at x, denoted by $q_t \sim \text{BESQ}_{x,2\theta}^{\delta}$.

$$q_t = x + 2\int_0^t \sqrt{q_s} d\mathcal{W}_s + 2\theta \int_0^t q_s ds + \delta t \tag{2}$$

By Ito's lemma, we get

$$X_t^2 = X_0^2 + 2\sigma \int_0^t X_s dW_s + 2\theta \int_0^t X_s^2 ds + \sigma^2 t.$$

From Levy's theorem, $\mathcal{W}_t \equiv \int_0^t I\left\{X_s \neq 0\right\} X_s/\left|X_s\right| dW_s$ is a Brownian motion due to $\langle \mathcal{W} \rangle_t = t$. Hence, $q_t = X_t^2/\sigma^2$ becomes $\text{BESQ}_{x,2\theta}^1$ with $x = X_0^2/\sigma^2$. .

According to Tanaka (1996), as N tends to ∞, Dickey-Fuller's coefficient test statistic $N\left(\hat{\beta}_N - 1\right)$ has the following nonstandard approximation.

$$N\left(\hat{\beta}_N - 1\right) \Rightarrow \frac{\int_0^1 X_s dX_s}{\int_0^1 X_s^2 ds}$$

Further, Dickey-Fuller's t test statistic has the following nonstandard limit.

$$t_N = \frac{\hat{\beta}_N - 1}{\hat{\sigma}/\sqrt{\sum_{n=1}^{N} x_{n-1}^2}} \Rightarrow \frac{\int_0^1 X_s dX_s}{\sigma\sqrt{\int_0^1 X_s^2 ds}}$$

Using Ito's lemma again, the numerator of this ratio is $\int_0^1 X_s dX_s = \left(X_1^2 - X_0^2 - \sigma^2\right)/2$. Thus, the limiting CDFs of the two test statistics can be computed through the joint probability density function (PDF) $f_{q_1, \int_0^1 q_s ds}(y, v)$ of q_1 and $\int_0^1 q_s ds$ as follows. The limiting CDF of Dickey-Fuller's coefficient test statistic is

$$\Psi_\theta(z) = P\left(\frac{\int_0^1 X_s dX_s}{\int_0^1 X_s^2 ds} \le z\right) = P\left(\frac{\frac{1}{2}\left(\frac{X_1^2}{\sigma^2} - \frac{X_0^2}{\sigma^2} - 1\right)}{\int_0^1 \frac{X_s^2}{\sigma^2} ds} \le z\right)$$

$$= P\left(q_1 - x - 1 \le 2z \int_0^1 q_s ds\right)$$

$$= \begin{cases} \int_0^{x+1} \int_0^{\frac{y-x-1}{2z}} f_{q_1, \int_0^1 q_s ds}(y,v)\, dvdy & z < 0 \\ \int_0^{\infty} \int_0^{2zv+x+1} f_{q_1, \int_0^1 q_s ds}(y,v)\, dydv & z \ge 0 \end{cases}$$

The limiting CDF of Dickey-Fuller's t test statistic is

$$\Psi_\theta^t(z) = P\left(\frac{\int_0^1 X_s dX_s}{\sigma\sqrt{\int_0^1 X_s^2 ds}} \le z\right) = P\left(\frac{\frac{1}{2}\left(\frac{X_1^2}{\sigma^2} - \frac{X_0^2}{\sigma^2} - 1\right)}{\sqrt{\int_0^1 \frac{X_s^2}{\sigma^2} ds}} \le z\right)$$

$$= P\left(q_1 - x - 1 \le 2z\sqrt{\int_0^1 q_s ds}\right)$$

$$= \begin{cases} \int_0^{x+1} \int_0^{\left(\frac{y-x-1}{2z}\right)^2} f_{q_1, \int_0^1 q_s ds}(y,v)\, dvdy & z < 0 \\ \int_0^{\infty} \int_0^{2z\sqrt{v}+x+1} f_{q_1, \int_0^1 q_s ds}(y,v)\, dydv & z \ge 0 \end{cases}$$

Abadir (1995a,b) considered the same expression with an initial value of 0 for the null hypothesis $H_0 : \theta = 0$. However, here we also adopt the above expressions with a nonzero initial value for alternative hypothesis $H_1 : \theta < 0$. The joint PDF $f_{q_1, \int_0^1 q_s ds}(y,v)$ is computed in the following subsections in several cases. In the following subsections, under each of the null and alternative hypotheses, the joint PDFs are derived for each case where the initial value is zero and nonzero.

2.1 Joint Density With Nonzero Initial Value Under H_0

As the first step in obtaining the joint density of q_t and $\int_0^t q_s ds$, we investigate the null hypothesis, i.e. $\theta = 0$. See Borodin & Selminen (2002) for functions is_v,es_v, s_v and $D_\mu(x)$. Let P_x^0 and E_x^0 be the probability and the expectation with $q_0 = x$ under H_0.

Lemma 1. *For a δ-dimensional squared Bessel process q_t with initial value $x \ne 0$, the joint PDF of q_t and its integral $\int_0^t q_s ds$ is*

$$f_{q_t, \int_0^t q_s ds}(y,v) = \mathrm{is}_v\left(\nu, t, 0, \frac{x+y}{2}, \frac{\sqrt{xy}}{2}\right) \frac{1}{2}\left(\frac{y}{x}\right)^{\frac{\nu}{2}} \tag{3}$$

where $\nu = \delta/2 - 1$ is the index of q_t, and for $\nu \ge -1$, $t + \nu t + r + z > 0$, $t > 0$

$$\mathrm{is}_v(\nu,t,r,z,w) := \mathcal{L}_\gamma^{-1}\left(\frac{\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma})} \exp\left(-r\sqrt{2\gamma} - z\sqrt{2\gamma}\coth(t\sqrt{2\gamma})\right) I_\nu\left(\frac{2w\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma})}\right)\right)$$
$$= \sum_{l=0}^{\infty} \frac{w^{\nu+2l}}{\Gamma(\nu+l+1)l!} \mathrm{es}_v(1+\nu+2l, 1+\nu+2l, t, r, z)$$

$$\mathrm{es}_v(\mu,\nu,t,r,z) := \mathcal{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\frac{\mu}{2}}}{\sinh^\nu(t\sqrt{2\gamma})} \exp\left(-r\sqrt{2\gamma} - z\sqrt{2\gamma}\coth(t\sqrt{2\gamma})\right)\right)$$
$$= \sum_{k=0}^{\infty} \frac{(-z)^k}{k!} s_v(\mu+k, \nu+k, t, r+z+kt)$$

$$\mathrm{s}_v\left(\mu,\nu,t,z\right) := \mathcal{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\sinh^\nu(t\sqrt{2\gamma})}e^{-z\sqrt{2\gamma}}\right)$$
$$= 2^\nu\sum_{k=0}^{\infty}\frac{\Gamma(\nu+k)e^{-(\nu t+z+2kt)^2/4y}}{\sqrt{2\pi}v^{1+\mu/2}\Gamma(\nu)k!}D_{\mu+1}\left(\frac{\nu t+z+2kt}{\sqrt{v}}\right),\quad \nu\geq 0,\quad \nu t+z>0.$$

$D_\mu\left(x\right)$ is the Parabolic cylinder function (See Borodin & Selminen(2002)).

Proof. According to Revuz and Yor (2013, p441), for $q_t \sim \mathrm{BESQ}_{x,0}^\delta$ with no drift, the density of q_t is

$$f_{q_t}(y) = q_t^\delta(x,y) = \frac{1}{2t}\left(\frac{y}{x}\right)^{\frac{\nu}{2}}\exp\left(-\frac{x+y}{2t}\right)I_\nu\left(\frac{\sqrt{xy}}{t}\right) \tag{4}$$

where $\nu=\delta/2-1$ is its index and I_ν is the modified Bessel function defined by

$$I_\nu\left(x\right) = \sum_{k=0}^{\infty}\frac{\left(x/2\right)^{2k+\nu}}{k!\Gamma\left(\nu+k+1\right)}\quad \nu\geq -1, x>0.$$

Pitman and Yor (1982) gave the expression of a Laplace transform for the conditional PDF of $\int_0^t q_s ds$, called Bessel bridge, which is

$$E_x^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right]$$
$$=\frac{\sqrt{2\gamma}t}{\sinh(\sqrt{2\gamma}t)}\exp\left\{\frac{x+y}{2t}\left(1-\sqrt{2\gamma}t\coth\sqrt{2\gamma}t\right)\right\}\frac{I_\nu\left(\frac{\sqrt{xy}\sqrt{2\gamma}}{\sinh\sqrt{2\gamma}t}\right)}{I_\nu\left(\frac{\sqrt{xy}}{t}\right)} \tag{5}$$

Multiplying (4) by (5), one sees that

$$E_x^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right]\cdot f_{q_t}(y)$$
$$=\frac{\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma})}\exp\left(-\frac{x+y}{2}\sqrt{2\gamma}\coth(t\sqrt{2\gamma})\right)I_\nu\left(\frac{\sqrt{xy}\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma})}\right)\frac{1}{2}\left(\frac{y}{x}\right)^{\frac{\nu}{2}} \tag{6}$$

On the other hand, this is also a Laplace transform of $f_{q_t,\int_0^t q_s ds}\left(y,v\right)$ with respect to v, because

$$E_x^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right]\cdot f_{q_t}(y) = \int_0^\infty e^{-\gamma v}f_{\int_0^t q_s ds|q_t}\left(v|y\right)dv\cdot f_{q_t}(y)$$
$$= \int_0^\infty e^{-\gamma v}f_{q_t,\int_0^t q_s ds}\left(y,v\right)dv$$
$$= \mathcal{L}_\gamma\left\{f_{q_t,\int_0^t q_s ds}\left(y,v\right)\right\}$$

By taking the inverse Laplace transform of equation (6), $f_{q_t,\int_0^t q_s ds}\left(y,v\right)$ is therefore derived. □

2.2 Joint Density with Zero Initial Value under H_0

Lemma 2. *For a δ-dimensional squared Bessel process q_t with initial value $x\neq 0$, the joint PDF of q_t and its integral $\int_0^t q_s ds$ is*

$$f_{q_t,\int_0^t q_s ds}\left(y,v\right) = \frac{y^\nu 2^{-(\nu+1)}}{\Gamma(\nu+1)}\mathrm{es}_v\left(\nu+1,\nu+1,t,0,\frac{y}{2}\right)$$

where $\nu=\delta/2-1$ is the index of q_t, and $\mathrm{es}_v\left(\mu,\nu,t,x,z\right)$ is defined in Lemma 1.

Proof. If the initial value x_0 is 0, then the squared Bessel process q_t also starts at 0. Revus and Yor (2013, p441) gave the PDF of q_t, by substituting $\delta/2$ with $\nu+1$ then we have

$$\begin{aligned} f_{q_t}(y) = q_t^{\delta}(0,y) &= (2t)^{-\frac{\delta}{2}}\Gamma(\frac{\delta}{2})^{-1}y^{\frac{\delta}{2}-1}\exp\left(-\frac{y}{2t}\right) \\ &= \frac{(2t)^{-(\nu+1)}}{\Gamma(\nu+1)}y^{\nu}\exp\left(-\frac{y}{2t}\right) \end{aligned} \tag{7}$$

Let x tend to 0 in the Bessel bridge (5), we get

$$\begin{aligned} E_0^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right] &= \lim_{x\to 0}E_x^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right] \\ &= \left(\frac{t\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma})}\right)^{\nu+1}\exp\left\{\frac{y}{2t}\left[1-t\sqrt{2\gamma}\coth\left(t\sqrt{2\gamma}\right)\right]\right\} \end{aligned} \tag{8}$$

By multiplying (7) and (8), we obtain

$$E_0^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right]\cdot f_{q_t}(y) = \frac{y^{\nu}2^{-(\nu+1)}}{\Gamma(\nu+1)}\left(\frac{\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma})}\right)^{\nu+1}\exp\left\{-\frac{y}{2}\sqrt{2\gamma}\coth\left(t\sqrt{2\gamma}\right)\right\}$$

This is also a Laplace transform of $f_{q_t,\int_0^t q_s ds}(y,w)$ since

$$\begin{aligned} E_0^0\left[\exp\left(-\gamma\int_0^t q_s ds\right)|q_t=y\right]\cdot f_{q_t}(y) &= \int_0^{\infty}e^{-\gamma w}f_{\int_0^t q_s ds|q_t}(v|y)\,dv\cdot f_{q_t}(y) \\ &= \int_0^{\infty}e^{-\gamma w}f_{q_t,\int_0^t q_s ds}(y,v)\,dv \\ &= \mathcal{L}_{\gamma}\left\{f_{q_t,\int_0^t q_s ds}(y,v)\right\} \end{aligned}$$

Thus, the joint density can be derived by taking a inverse Laplace transform.

$$\begin{aligned} f_{q_t,\int_0^t q_s ds}(y,v) &= \mathcal{L}_{\gamma}^{-1}\left\{\frac{y^{\nu}2^{-(\nu+1)}}{\Gamma(\nu+1)}\left(\frac{\sqrt{2\gamma}}{\sinh(t\sqrt{2\gamma}}\right)^{\nu+1}\exp\left\{-\frac{y}{2}\sqrt{2\gamma}\coth\left(t\sqrt{2\gamma}\right)\right\}\right\} \\ &= \frac{y^{\nu}2^{-(\nu+1)}}{\Gamma(\nu+1)}\mathrm{es}_v\left(\nu+1,\nu+1,t,0,\frac{y}{2}\right) \end{aligned}$$

□

For the AR(1) process, $q_t \equiv W_t^2$ has dimension $\delta=1$ and index $\nu=-1/2$. Since $\Gamma(\nu+1)=\sqrt{\pi}$, we have

$$f_{W_t^2,\int_0^t W_s^2 ds}(y,v) = \frac{y^{-\frac{1}{2}}}{\sqrt{2\pi}}\mathrm{es}_v\left(\frac{1}{2},\frac{1}{2},t,0,\frac{y}{2}\right)$$

2.3 Joint Density under Alternative via Girsanov Transformation

In the null and alternative hypotheses, the squared Bessel processes are different in a drift θ.

$$P^0: q_t = x + 2\int_0^t \sqrt{q_s}d\mathcal{W}_s + t$$

$$P^{\theta}: q_t = x + 2\int_0^t \sqrt{q_s}d\mathcal{W}_s + 2\theta\int_0^t q_s ds + t$$

The Girsanov transformation can be applied to remove the drift. The Radon–Nikodym derivative of P^θ with respect to P^0 is

$$\begin{aligned}\frac{dP^\theta}{dP^0} &= \exp\left(\int_0^t \theta\sqrt{q_s}d\mathcal{W}_s - \frac{1}{2}\int_0^t \theta^2 q_s ds\right)\\ &= \exp\left(\frac{\theta}{2}\left(q_t - x - t\right) - \frac{\theta^2}{2}\int_0^t q_s ds\right)\end{aligned}$$

Therefore, the joint PDF of $\left(q_t, \int_0^t q_s ds\right)$ with $x \neq 0$ under P_x^δ is

$$f_{q_t,\int_0^t q_s ds}\left(y,v\right) = \exp\left\{\frac{\theta}{2}\left(y-x-t\right) - \frac{\theta^2}{2}v\right\} \mathrm{is}_{-1/2}\left(\nu, t, 0, \frac{x+y}{2}, \frac{\sqrt{xy}}{2}\right)\frac{1}{2}\left(\frac{y}{x}\right)^{-1/4}.$$

The joint PDF of $\left(q_t, \int_0^t q_s ds\right)$ with under P_0^δ is

$$f_{q_t,\int_0^t q_s ds}\left(y,v\right) = \exp\left\{\frac{\theta}{2}\left(y-t\right) - \frac{\theta^2}{2}v\right\}\frac{y^{-\frac{1}{2}}}{\sqrt{2\pi}}\mathrm{es}_v\left(\frac{1}{2},\frac{1}{2},t,0,\frac{y}{2}\right).$$

3 Concise Computation of Powers

The power of the hypothesis test is the probability of rejecting the null hypothesis when the alternative hypothesis is true. So we can compute the power by the CDF of the test statistic. However, the computation of the integral takes a great time and is often not provided by a free software environment. So we propose a concise way to compute the power of the Dickey-Fuller test.

Let $\Psi_\theta\left(z\right)$ to be the CDF of DF's coefficient test statistic with respect to the local parameter θ satisfying $\beta = 1+\theta/N$. Denote $w_\alpha = \Psi_0^{-1}\left(\alpha\right)$, then $\Psi_0\left(w_\alpha\right) = \alpha$. In left tailed test with significance level of α, its critical value is w_α and its power is $\Psi_\theta\left(w_\alpha\right)$. Let $\Phi\left(z\right)$ and $\varphi(z)$ to be the CDF and PDF of the standard normal distribution and denote $z_\alpha = \Phi^{-1}\left(\alpha\right)$. As explained in the abstract, we want to obtain the power of the DF test from the power of the test the null hypothesis of $N(0,1)$ against the alternative $N(\delta,1)$ via a deformation of δ through the relation $\Psi_\theta\left(w_\alpha\right) = \Phi\left(z_\alpha - \delta\right)$. So, define an implicit function $\delta_\alpha\left(\theta\right)\left(= z_\alpha - \delta\right)$ satisfying

$$\Psi_\theta\left(w_\alpha\right) = \Phi\left(\delta_\alpha\left(\theta\right)\right). \tag{9}$$

The Radon–Nikodym derivative of P^θ with respect to P_0 is

$$\frac{dP^\theta}{dP_0}|_{\mathscr{F}_1} = \exp\left(\frac{\theta}{2}\left(q_1 - x - 1\right) - \frac{\theta^2}{2}\int_0^1 q_s ds\right)$$

Let $G\left(\theta, y, v\right) = \exp\left\{\frac{\theta}{2}\left(y-x-1\right) - \frac{\theta^2}{2}v\right\}$, then we can express $\Psi_\theta\left(w_\alpha\right)$ as

$$\begin{aligned}\Psi_\theta\left(w_\alpha\right) &= \int_{\frac{(q_1-x-1)/2}{\int_0^1 q_s ds}\le w_\alpha} \exp\left(\frac{\theta}{2}\left(q_1 - x - 1\right) - \frac{\theta^2}{2}\int_0^1 q_s ds\right)dP_0\\ &= E^0\left[1_{\left\{\frac{(q_1-x-1)/2}{\int_0^1 q_s ds}\le w_\alpha\right\}} G\left(\theta, q_1, \int_0^1 q_s ds\right)\right]\end{aligned} \tag{10}$$

Taking nth order derivative of (10) and tending θ to 0, we have

$$\frac{\partial^n}{\partial\theta^n}\Psi_\theta\left(w_\alpha\right)|_{\theta=0} = E^0\left[1_{\left\{\frac{(q_1-x-1)/2}{\int_0^1 q_s ds}\le w_\alpha\right\}}\frac{\partial^n}{\partial\theta^n}G\left(0, q_1, \int_0^1 q_s ds\right)\right].$$

For simplicity, denote $T_n = \frac{\partial^n}{\partial \theta^n} \Psi_\theta (w_\alpha) |_{\theta=0}$.

Differentiating the right-side in the equation (9) w.r.t θ, we have

$$\begin{cases} \frac{\partial}{\partial \theta} \Psi_\theta (w_\alpha) = \varphi (\delta_\alpha (\theta)) \delta'_\alpha (\theta) \\ \frac{\partial^2}{\partial \theta^2} \Psi_\theta (w_\alpha) = \varphi (\delta_\alpha (\theta)) \delta''_\alpha (\theta) - \varphi (\delta_\alpha (\theta)) \delta_\alpha (\theta) [\delta'_\alpha (\theta)]^2 \\ \frac{\partial^3}{\partial \theta^3} \Psi_\theta (w_\alpha) = \varphi (\delta_\alpha (\theta)) \delta^{(3)}_\alpha (\theta) - 3\varphi (\delta_\alpha (\theta)) \delta_\alpha (\theta) \delta'_\alpha (\theta) \delta''_\alpha (\theta) \\ \qquad +\varphi (\delta_\alpha (\theta)) \delta^2_\alpha (\theta) [\delta'_\alpha (\theta)]^3 - \varphi (\delta_\alpha (\theta)) [\delta'_\alpha (\theta)]^3 \\ \qquad \vdots \end{cases}$$

Since $\Psi_0 (w_\alpha) = \alpha = \Phi (\delta_\alpha (0))$, $\delta_\alpha (0) = z_\alpha$. Let θ tend to 0, we have

$$\begin{cases} T_1 = \varphi (z_\alpha) \delta'_\alpha (0) \\ T_2 = \varphi (z_\alpha) \delta''_\alpha (0) - \varphi (z_\alpha) z_\alpha [\delta'_\alpha (0)]^2 \\ T_3 = \varphi (z_\alpha) \delta^{(3)}_\alpha (0) - 3\varphi (z_\alpha) z_\alpha \delta'_\alpha (0) \delta''_\alpha (0) + \varphi (z_\alpha) z^2_\alpha [\delta'_\alpha (0)]^3 - \varphi (z_\alpha) [\delta'_\alpha (0)]^3 \\ \qquad \vdots \end{cases}$$

Solve the system of equations, we can obtain the derivatives of $\delta (\theta)$ at 0 for each order.

$$\begin{cases} \delta'_\alpha (0) = \frac{T_1}{\varphi(z_\alpha)} \\ \delta''_\alpha (0) = \frac{T_2}{\varphi(z_\alpha)} + z_\alpha \frac{T_1^2}{[\varphi(z_\alpha)]^2} \\ \delta^{(3)}_\alpha (0) = \frac{T_3}{\varphi(z_\alpha)} + 3z_\alpha \frac{T_1 T_2}{[\varphi(z_\alpha)]^2} + 2z^2_\alpha \frac{T_1^3}{[\varphi(z_\alpha)]^3} + \frac{T_1^3}{[\varphi(z_\alpha)]^3} \\ \qquad \vdots \end{cases}$$

Therefore, a Taylor expansion of $\delta_\alpha (\theta)$ is obtained. Once the values of T_n for each critical value w_α are precomputed from the null distribution, the power for different θ can be calculated quickly. In fact, we will see that third-order Taylor expansions are sufficient for accurate calculations.

$$\delta_\alpha (\theta) = z_\alpha + \sum_{n=1}^{\infty} \frac{\theta^n}{n!} \frac{\partial^n}{\partial \theta^n} \delta_\alpha (0)$$

The same method can be used to approximate the power of the DF t-test statistic. The difference is that we should substitute $\Psi_\theta (w_\alpha)$ with the CDF of t test statistic $\Psi^t_\theta (w_\alpha)$ above.

4 Exact Distribution of Dickey-Fuller's Coefficient Test Statistic

According to Imhof(1961), the quadratic form of normal variables has an explicit expression. We apply its more straightforward expression for computing the exact distribution of Dickey-Fuller's coefficient test statistic.

Lemma 3. *Let $x = (x_1, \cdots, x_n)'$ be a column random vector which follows a multidimensional normal law with mean vector O and covariance matrix I_n, $\mu = (\mu_1, \cdots, \mu_n)'$ be a constant vector. The quadratic form $Q = (x + \mu)' A (x + \mu)$ follows*

$$P (Q > x) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin \{\theta (u)\}}{u \rho (u)} du$$

where

$$\theta (u) = \frac{1}{2} \sum_{r=1}^{n} \left[\arctan (\lambda_r u) + \frac{\mu_r^2 \lambda_r u}{1 + \lambda_r^2 u^2} \right] - \frac{1}{2} xu$$

$$\rho(u) = \prod_{r=1}^{n} \left(1+\lambda_r^2 u^2\right)^{\frac{1}{4}} \exp\left(\sum_{r=1}^{N} \frac{(\mu_r \lambda_r u)^2}{2\left(1+\lambda_r^2 u^2\right)}\right)$$

and λ_r are eigenvalues of A.

4.1 Constant Initial Value

If the initial value is a fixed value, define the N-dimensional vectors,

$$\boldsymbol{x} = (x_1, \cdots, x_N)'$$

$$\boldsymbol{\mu} = \left(\beta x_0, \beta^2 x_0, \cdots, \beta^N x_0\right)'$$

$$\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_N)'.$$

Using a recursion formula for AR(1) process, we have the relation $\boldsymbol{x} = A\boldsymbol{\varepsilon} + \boldsymbol{\mu}$ where

$$A = \begin{pmatrix} 1 & & & & \\ \beta & 1 & & O & \\ \beta^2 & \beta & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \beta^{N-1} & \beta^{N-2} & \beta^{N-3} & \cdots & 1 \end{pmatrix}$$

Let $\zeta = -\left(z/N+1\right)^{-1}/2$, then for $z \geq -N$, $N\left(\hat{\beta}_N - 1\right) \leq z$ is equivalent to

$$\sum_{t=1}^{N} x_{t-1}^2 + 2\zeta \sum_{t=1}^{N} x_t x_{t-1} \geq 0 \tag{11}$$

Imhof (1961) provides a formula for computing the quadratic form of normally distributed variables. However, it is not possible to use this formula directly since the polynomial in (11) is inhomogeneous because of the linear term $x_0 x_1$. However, if we adopt a linear transformation, then the effect of this linear term can be removed. Using a similar approach to the Lagrange method to complete the square, one can express the polynomial in (6) into a canonical form.

$$\sum_{t=1}^{N} x_{t-1}^2 + 2\zeta \sum_{t=1}^{N} x_t x_{t-1} = \sum_{n=1}^{N-1} \left[c_n \left(x_n + b_n x_{n+1} + \gamma_n\right)^2 + d_n\right] + c_N \left(x_N + \gamma_N\right)^2$$

where c_n, d_n, b_n and γ_n are derived by the following recurrence relations.

$$\begin{cases} c_1 = 1, \gamma_1 = \zeta x_0 & \\ b_1 = \zeta, d_1 = x_0^2 & \\ d_n = -c_{n-1}\gamma_{n-1}^2 & n = 2, \cdots, N-1 \\ c_n = 1 - c_{n-1} b_{n-1}^2 & n = 2, \cdots, N-1 \\ \gamma_n = -\frac{b_{n-1}\gamma_{n-1}c_{n-1}}{c_n} & n = 2, \cdots, N-1 \\ b_n = \frac{\zeta}{c_n} & n = 2, \cdots, N-1 \\ c_N = -c_{N-1} b_{N-1}^2 & \\ \gamma_N = \frac{\gamma_{N-1}}{b_{N-1}} & \end{cases}$$

By making $y_n = x_n + b_n x_{n+1} + \gamma_n$, $\boldsymbol{y} = (y_1, \cdots, y_N)'$ and $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_N)'$, we get the desired linear transformation $\boldsymbol{y} = B\boldsymbol{x} + \boldsymbol{\gamma}$ where

$$B = \begin{pmatrix} 1 & b_1 & & & & \\ & 1 & b_2 & & O & \\ & & 1 & \ddots & & \\ & & & \ddots & b_{N-2} & \\ & O & & & 1 & b_{N-1} \\ & & & & & 1 \end{pmatrix}$$

After this linear transformation, a quadratic form of normally distributed variables is extracted from (6), which enables us to compute the exact distribution of Dickey-Fuller's test statistic through Imhof's formula. In the CDF of test statistic, writing the inequality in a matrix form, we have

$$\begin{aligned} P\left(N\left(\hat{\beta}-1\right) \le z\right) &= P\left(\sum_{n=1}^{N} c_n y_n^2 + \sum_{n=1}^{N-1} d_n \ge 0\right) \\ &= P\left(\boldsymbol{y}'C\boldsymbol{y} \ge -\sum_{n=1}^{N} d_n\right) \\ &= P\left(\left(BA\boldsymbol{\varepsilon} + B\boldsymbol{\mu} + \boldsymbol{\gamma}\right)' C \left(BA\boldsymbol{\varepsilon} + B\boldsymbol{\mu} + \boldsymbol{\gamma}\right) \ge -\sum_{n=1}^{N-1} d_n\right) \\ &= P\left(\left(\frac{\boldsymbol{\varepsilon}}{\sigma} + \frac{(BA)^{-1}(B\boldsymbol{\mu} + \boldsymbol{\gamma})}{\sigma}\right)' (BA)' CBA \left(\frac{\boldsymbol{\varepsilon}}{\sigma} + \frac{(BA)^{-1}(B\boldsymbol{\mu} + \boldsymbol{\gamma})}{\sigma}\right) \ge -\sum_{n=1}^{N-1} \frac{d_n}{\sigma^2}\right) \end{aligned}$$

where $C = \operatorname{diag}(c_1, c_2, \cdots, c_N)$.

Eigendecompose the symmetric matrix $(BA)' CBA$ into $Q'\Lambda Q$, where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_N)$ is a diagonal matrix composed of eigenvalues of $(BA)' CBA$, and Q is composed of their respective orthogonal eigenvectors. Then we can turn the quadratic form of normally distributed variables into a standard form.

$$P\left(N\left(\hat{\beta}-1\right) \le z\right) = P\left(\left(\frac{Q\boldsymbol{\varepsilon}}{\sigma} + \frac{Q(BA)^{-1}(B\boldsymbol{\mu} + \boldsymbol{\gamma})}{\sigma}\right)' \Lambda \left(\frac{Q\boldsymbol{\varepsilon}}{\sigma} + \frac{Q(BA)^{-1}(B\boldsymbol{\mu} + \boldsymbol{\gamma})}{\sigma}\right) \ge -\sum_{n=1}^{N-1} \frac{d_n}{\sigma^2}\right)$$

Following Imhof (1961), the exact distribution of Dickey-Fuller's test statistic has a explicit expression.

$$P\left(N\left(\hat{\beta}_N-1\right) \le z\right) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\sin\{\theta(u)\}}{u\rho(u)} du$$

where

$$\begin{aligned} \theta(u) &= \frac{1}{2}\sum_{r=1}^{N}\left[\arctan(\lambda_r u) + \frac{\delta_r^2 \lambda_r u}{1+\lambda_r^2 u^2}\right] + \sum_{n=1}^{N-1} \frac{d_n u}{2\sigma^2} \\ \rho(u) &= \prod_{r=1}^{N}\left(1+\lambda_r^2 u^2\right)^{\frac{1}{4}} \exp\left(\sum_{r=1}^{N} \frac{(\delta_r \lambda_r u)^2}{2(1+\lambda_r^2 u^2)}\right) \end{aligned}$$

and δ_r is the rth element of $Q(BA)^{-1}(B\boldsymbol{\mu} + \boldsymbol{\gamma})/\sigma$.

4.2 Normally Distributed Initial Value

In order to avoid the computation of extracting the quadratic form, one can treat the fixed initial value x_0 to be a normally distributed variable $x_0 + \eta\varepsilon_0$, where η is a very small value and $\varepsilon_0 \sim N(0, \sigma^2)$ is independent to $\varepsilon_1, \cdots, \varepsilon_N$. This treatment loses little precision in numerical computation.

Put $\boldsymbol{x} = (x_0, x_1, \cdots, x_N)'$, $\boldsymbol{\mu} = (x_0, \beta x_0, \beta^2 x_0, \cdots, \beta^N x_0)'$, and $\boldsymbol{\varepsilon} = (\varepsilon_0, \varepsilon_1, \cdots, \varepsilon_N)'$. Using the recursion formula for the AR(1) process, we have the relation $\boldsymbol{x} = A\boldsymbol{\varepsilon} + \boldsymbol{\mu}$ where

$$A = \begin{pmatrix} \eta & & & & & \\ \eta\beta & 1 & & & O & \\ \eta\beta^2 & \beta & 1 & & & \\ \eta\beta^3 & \beta^2 & \beta & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \eta\beta^N & \beta^{N-1} & \beta^{N-2} & \beta^{N-3} & \cdots & 1 \end{pmatrix}.$$

We can express the least squares estimator of the regression coefficient in a matrix form[1].

$$\hat{\beta}_N = \frac{\sum_{t=1}^{N} x_t x_{t-1}}{\sum_{t=1}^{N} x_{t-1}^2} = \frac{\boldsymbol{x}' \left[\frac{1}{2} \begin{pmatrix} \mathbf{0} & I_N \\ 0 & \mathbf{0}' \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbf{0}' & 0 \\ I_N & \mathbf{0} \end{pmatrix} \right] \boldsymbol{x}}{\boldsymbol{x}' \begin{pmatrix} I_N & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} \boldsymbol{x}}$$

Therefore, the CDF of Dickey-Fuller's test statistic can be written through a quadratic form of normally distributed variables, enabling us to compute its exact distribution by applying Imhof's formula directly.

$$\begin{aligned} P\left(N\left(\hat{\beta}-1\right) \le y\right) &= P\left(\hat{\beta} \le \frac{y}{N} + 1\right) \\ &= P\left(\frac{1}{2}\boldsymbol{x}' \left[\begin{pmatrix} \mathbf{0} & I_N \\ 0 & \mathbf{0}' \end{pmatrix} + \begin{pmatrix} \mathbf{0}' & 0 \\ I_N & \mathbf{0} \end{pmatrix} \right] \boldsymbol{x} \le \left(\frac{y}{N}+1\right) \boldsymbol{x}' \begin{pmatrix} I_N & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} \boldsymbol{x}\right) \\ &= P\left(\boldsymbol{x}' \left\{ \left(\frac{y}{N}+1\right) \begin{pmatrix} I_N & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} - \left[\frac{1}{2} \begin{pmatrix} \mathbf{0} & I_N \\ 0 & \mathbf{0}' \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbf{0}' & 0 \\ I_N & \mathbf{0} \end{pmatrix} \right] \right\} \boldsymbol{x} \ge 0\right) \\ &= P\left(\left(\frac{\boldsymbol{\varepsilon}}{\sigma} + \frac{A^{-1}\boldsymbol{\mu}}{\sigma}\right)' A'BA \left(\frac{\boldsymbol{\varepsilon}}{\sigma} + \frac{A^{-1}\boldsymbol{\mu}}{\sigma}\right) \ge 0\right) \end{aligned}$$

where

$$B = \left(\frac{y}{N}+1\right) \begin{pmatrix} I_N & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} - \left[\frac{1}{2} \begin{pmatrix} \mathbf{0} & I_N \\ 0 & \mathbf{0}' \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbf{0}' & 0 \\ I_N & \mathbf{0} \end{pmatrix} \right]$$

Eigendecompose the symmetric matrix $A'BA$ into $P'\Lambda P$, where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_{N+1})$ is composed of the eigenvalues of $A'BA$ and P is composed of their respective orthogonal eigenvectors. Then we have

$$P\left(N\left(\hat{\beta}-1\right) \le y\right) = P\left(\left(\frac{P\boldsymbol{\varepsilon}}{\sigma} + \frac{PA^{-1}\boldsymbol{\mu}}{\sigma}\right)' \Lambda \left(\frac{P\boldsymbol{\varepsilon}}{\sigma} + \frac{PA^{-1}\boldsymbol{\mu}}{\sigma}\right) \ge 0\right)$$

Following Imhof (1961), the exact distribution of Dickey-Fuller's test statistic is

$$P\left(N\left(\hat{\beta}-1\right) \le y\right) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin\{\theta(u)\}}{u\rho(u)} du$$

[1] I_N is a identity matrix of size N

where

$$\theta(u) = \frac{1}{2}\sum_{r=1}^{N+1}\left[\arctan(\lambda_r u) + \frac{\delta_r^2 \lambda_r u}{1+\lambda_r^2 u^2}\right]$$

$$\rho(u) = \prod_{r=1}^{N+1}\left(1+\lambda_r^2 u^2\right)^{\frac{1}{4}} \exp\left(\sum_{r=1}^{N+1}\frac{(\delta_r \lambda_r u)^2}{2\left(1+\lambda_r^2 u^2\right)}\right)$$

and δ_r is the rth element of $\frac{P' A^{-1}\mu}{\sigma}$.

5 Simulation and Numerical Computation

5.1 Density of Dickey-Fuller's Coefficient Test Statistic

We conduct numerical computations for limiting approximation and exact distributions of Dickey-Fuller's coefficient test statistic and for stationary and explosive alternative hypotheses. We apply a small sample size of $N = 25$. For the exact distribution, there is little difference between the fixed initial value and the normally distributed initial value if η is small enough. We treat the initial value as a fixed value in the numerical computation and compute the PDF of Dickey-Fuller's test statistic by numerical differentiation. We also implement simulations for verification. As shown in figures 1 and 2, the limiting and exact densities are almost the same for both stationary and explosive alternative hypotheses.

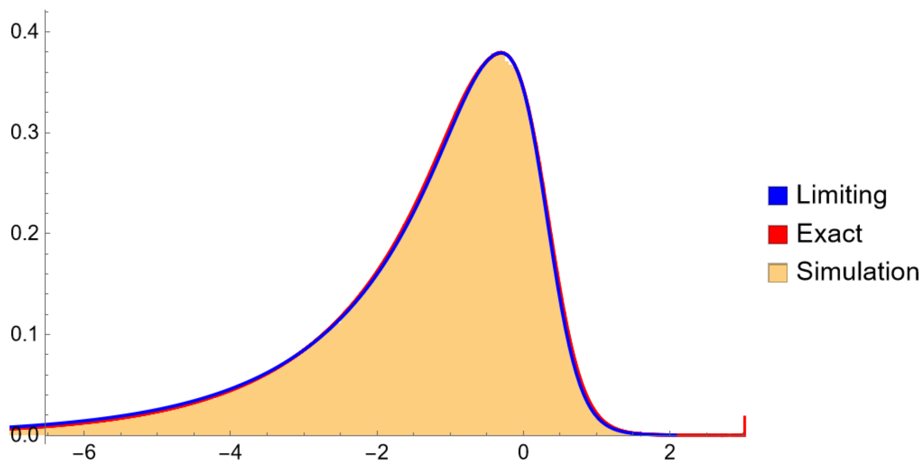

Fig 1: $\beta = 0.98$, $X_0 = 1$, $N = 25$

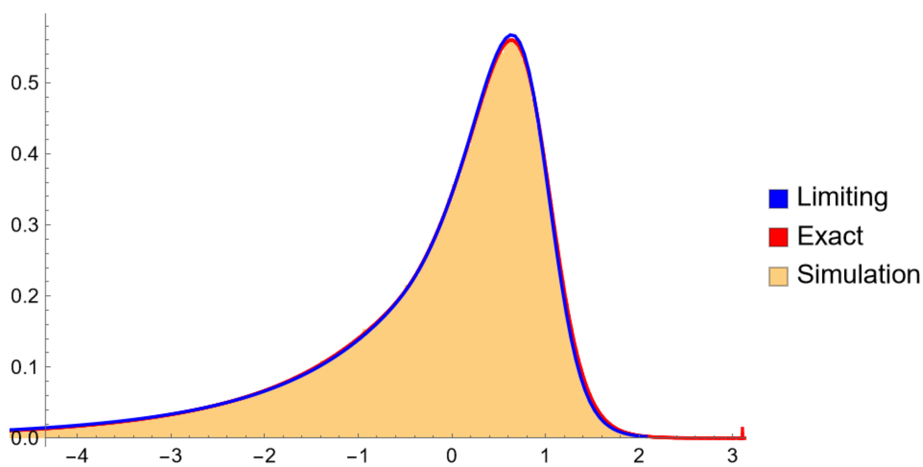

Fig 2: $\beta = 1.02$, $X_0 = 1$, $N = 25$

5.2 Power of Dickey-Fuller's Coefficient Test

In the case of a stationary alternative hypothesis $\beta < 0$, the Dickey-Fuller test is a left-tailed test. The critical region is $R_1 = \left\{N\left(\hat{\beta}_N - 1\right) < w_\alpha\right\}$ at a significance level of α, where w_α is computed via the CDF of Dickey-Fuller's test statistic under the null hypothesis and it satisfies

$$P_0\left\{N\left(\hat{\beta}_N - 1\right) < w_\alpha\right\} = \alpha$$

The power of a hypothesis test is the probability of rejecting the null hypothesis when the alternative hypothesis is true. So we can compute the statistical powers with different regression coefficients β, which is $P^\theta\left\{N\left(\hat{\beta}_N - 1\right) < w_\alpha\right\}$. For significance levels of 5% and a sample size of 25, the powers are shown in the following tables with different initial values and regression coefficients. The critical values and powers in table 1 is computed through the limiting distribution, and table 2 uses the exact distribution. We can see that the power increases significantly as the initial value grows, especially for a smaller β.

x_0	0	1	2	3	4	5	6	7	8	9	10
$w_{0.05}$	-8.039	-7.730	-6.930	-5.911	-4.902	-4.020	-3.295	-2.716	-2.258	-1.896	-1.609
$\beta = 0.99$	5.6%	5.6%	5.7%	5.9%	6.0%	6.3%	6.6%	7.0%	7.5%	8.1%	8.7%
$\beta = 0.98$	6.3%	6.4%	6.5%	6.8%	7.3%	7.9%	8.6%	9.6%	10.9%	12.4%	14.2%
$\beta = 0.97$	7.0%	7.1%	7.4%	7.9%	8.7%	9.7%	11.1%	12.9%	15.2%	18.1%	21.7%
$\beta = 0.96$	7.9%	8.0%	8.4%	9.2%	10.3%	11.8%	14.0%	16.8%	20.6%	25.4%	31.2%
$\beta = 0.95$	8.7%	8.9%	9.5%	10.5%	12.1%	14.3%	17.4%	21.5%	27.0%	33.9%	42.1%
$\beta = 0.94$	9.7%	9.9%	10.7%	12.0%	14.1%	17.1%	21.3%	26.9%	34.3%	43.4%	53.7%
$\beta = 0.93$	10.7%	11.0%	12.0%	13.7%	16.3%	20.2%	25.6%	32.9%	42.2%	53.2%	65.0%
$\beta = 0.92$	11.8%	12.2%	13.4%	15.5%	18.8%	23.6%	30.4%	39.4%	50.5%	63.0%	75.2%
$\beta = 0.91$	13.0%	13.5%	14.9%	17.4%	21.5%	27.4%	35.6%	46.2%	58.8%	71.9%	83.5%
$\beta = 0.90$	14.3%	14.8%	16.5%	19.5%	24.4%	31.4%	41.1%	53.2%	66.7%	79.7%	89.7%

Tab 1: (Lim)Stationary alternative: Power of DF Test at significance level of 5%, $N = 25$

x_0	0	1	2	3	4	5	6	7	8	9	10
$w_{0.05}$	-7.371	-7.085	-6.357	-5.443	-4.544	-3.757	-3.106	-2.581	-2.160	-1.824	-1.555
$\beta = 0.99$	5.6%	5.6%	5.7%	5.9%	6.1%	6.4%	6.7%	7.1%	7.6%	8.2%	8.9%
$\beta = 0.98$	6.3%	6.4%	6.6%	6.9%	7.4%	8.0%	8.8%	9.9%	11.1%	12.7%	14.6%
$\beta = 0.97$	7.0%	7.2%	7.5%	8.1%	8.9%	10.0%	11.4%	13.3%	15.7%	18.8%	22.5%
$\beta = 0.96$	7.9%	8.0%	8.5%	9.4%	10.6%	12.3%	14.6%	17.6%	21.5%	26.4%	32.4%
$\beta = 0.95$	8.8%	9.0%	9.7%	10.8%	12.5%	14.9%	18.2%	22.6%	28.3%	35.4%	43.8%
$\beta = 0.94$	9.7%	10.0%	10.9%	12.4%	14.7%	18.0%	22.4%	28.4%	36.0%	45.3%	55.8%
$\beta = 0.93$	10.8%	11.2%	12.3%	14.2%	17.2%	21.4%	27.2%	34.8%	44.4%	55.6%	67.3%
$\beta = 0.92$	12.0%	12.4%	13.8%	16.2%	19.9%	25.2%	32.4%	41.7%	53.1%	65.5%	77.3%
$\beta = 0.91$	13.2%	13.7%	15.4%	18.3%	22.8%	29.3%	38.0%	48.9%	61.6%	74.4%	85.3%
$\beta = 0.90$	14.5%	15.2%	17.2%	20.7%	26.0%	33.7%	43.8%	56.2%	69.6%	82.0%	91.2%

Tab 2: (Exact)Stationary alternative: Power of DF Test at significance level of 5%, $N = 25$

The power of DF's classical test can be approximated by the CDF of normal distribution. Table 3 gives the Taylor coefficient of $\delta_\alpha(\theta)$ up to 5th order for $\alpha = 0.05$ and different initial values.

	$\delta_\alpha(0)$	$\delta'_\alpha(0)$	$\delta''_\alpha(0)$	$\delta^{(3)}_\alpha(0)$	$\delta^{(4)}_\alpha(0)$	$\delta^{(5)}_\alpha(0)$
$x_0 = 0$	-1.645	-0.229552	0.001297	0.000171	0.000017	0.000001
$x_0 = 5$	-1.645	-0.459086	0.005200	0.001383	0.000286	0.000059
$x_0 = 10$	-1.645	-1.14227	0.031158	0.020773	0.009873	0.004891

Tab 3: Taylor coefficient for $\delta_\alpha(\theta)$

In figures 3-4, we compared the powers from the limiting distribution and the approximation through normal distribution. We see that it is good enough to use 3rd order Taylor expansion of $\delta_\alpha(\theta)$ for different initial values.

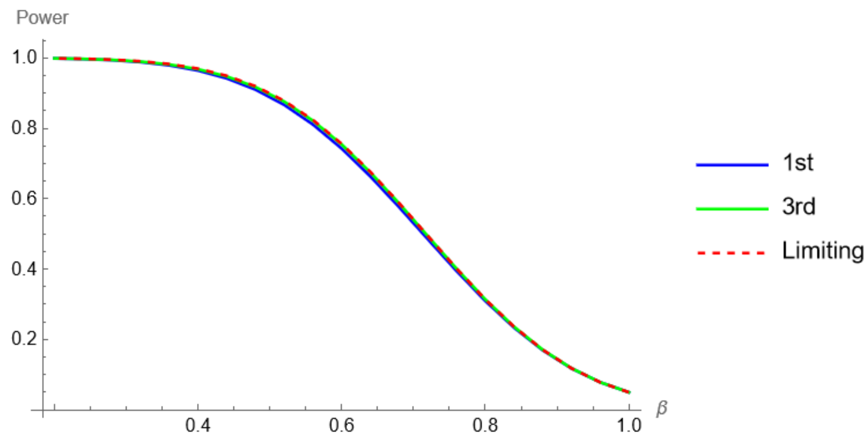

Fig 3: DF coefficient test: Approximation by normal distribution ($\alpha = 0.05, x_0 = 0, N = 25$)

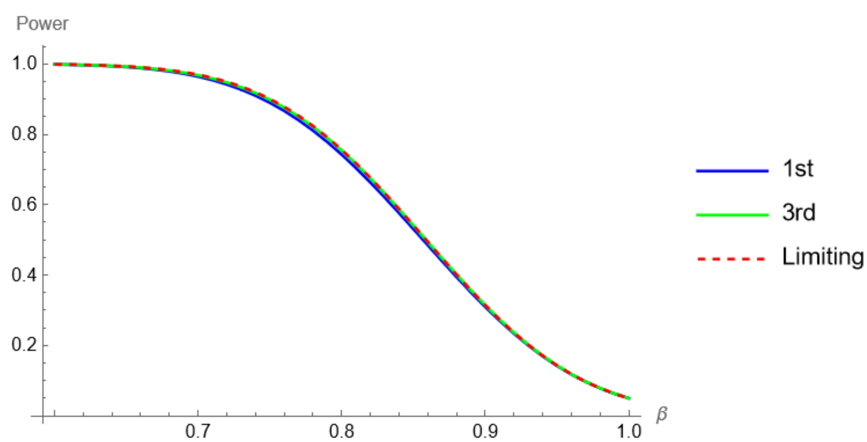

Fig 4: DF coefficient test: Approximation by normal distribution ($\alpha = 0.05, x_0 = 5, N = 25$)

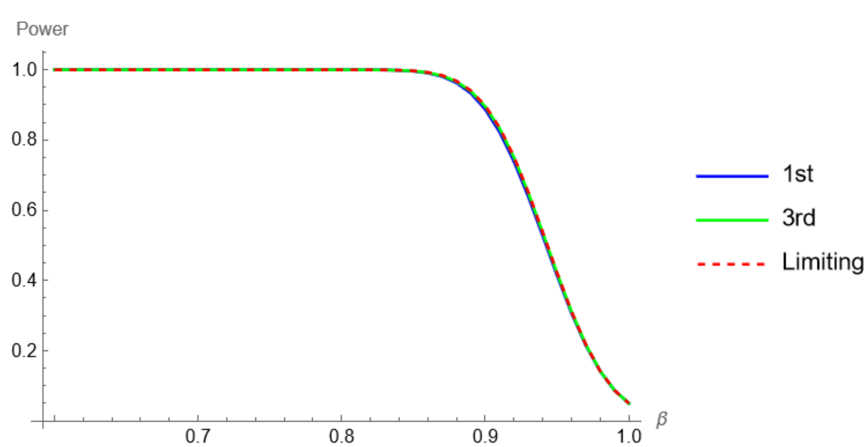

Fig 5: DF coefficient test: Approximation by normal distribution ($\alpha = 0.05, x_0 = 10, N = 25$)

We apply the critical value of limiting distribution in the computation of power. The number of iteration times is 100000 in the simulation. Figures 6-8 show that the powers computed from simulation and exact distribution coincide for any cases.

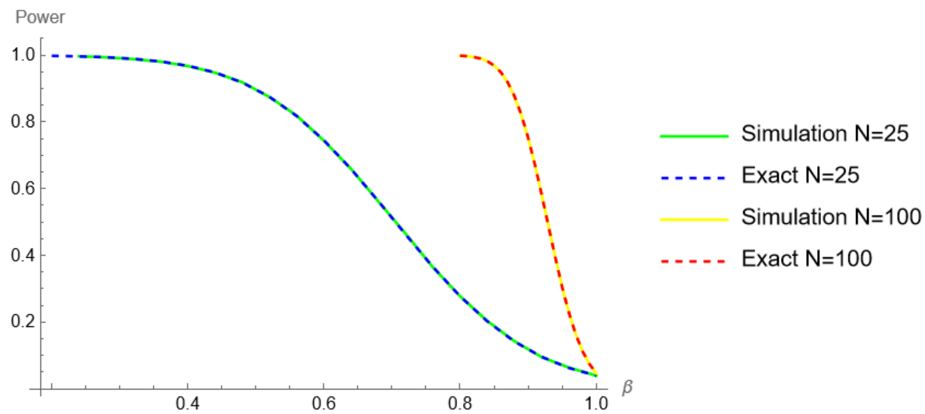

Fig 6: DF coefficient test: Simulation and Exact power ($\alpha = 0.05, x_0 = 0$)

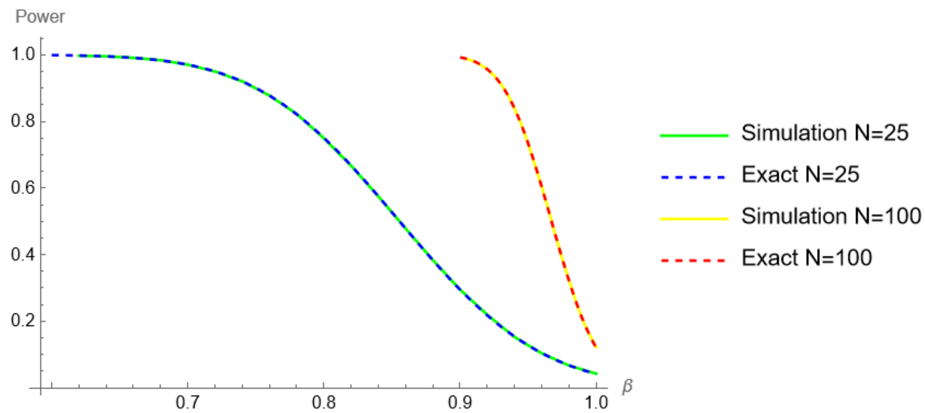

Fig 7: DF coefficient test: Simulation and Exact power ($\alpha = 0.05, x_0 = 5$)

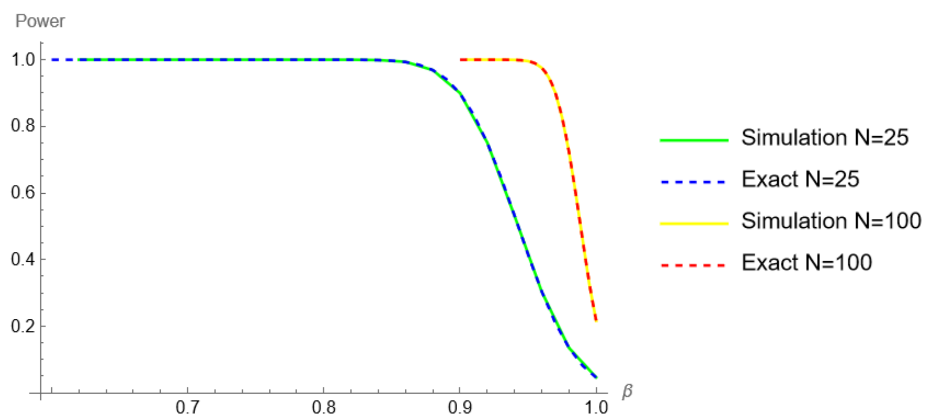

Fig 8: DF coefficient test: Simulation and Exact power ($\alpha = 0.05, x_0 = 10$)

In figures 9-11, we see that the powers of the DF test get closer to the powers computed from limiting distribution as the sample size grows. It implies the consistency of limiting distribution holds.

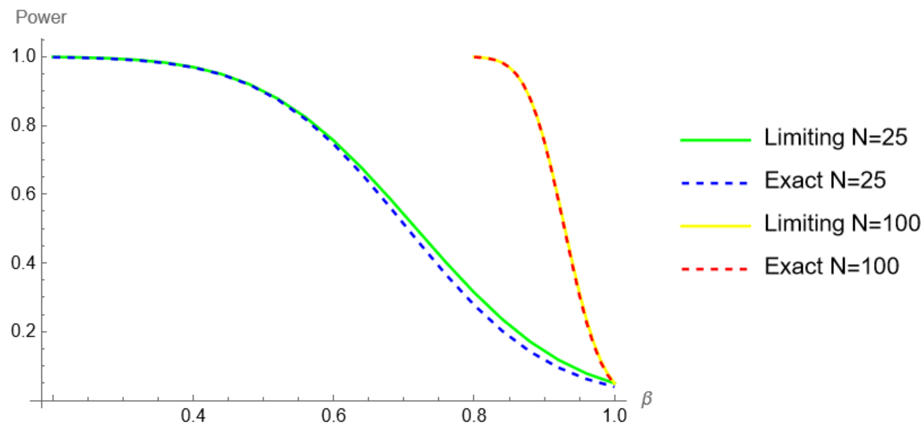

Fig 9: DF coefficient test: Exact and Limiting power($\alpha = 0.05, x_0 = 0$)

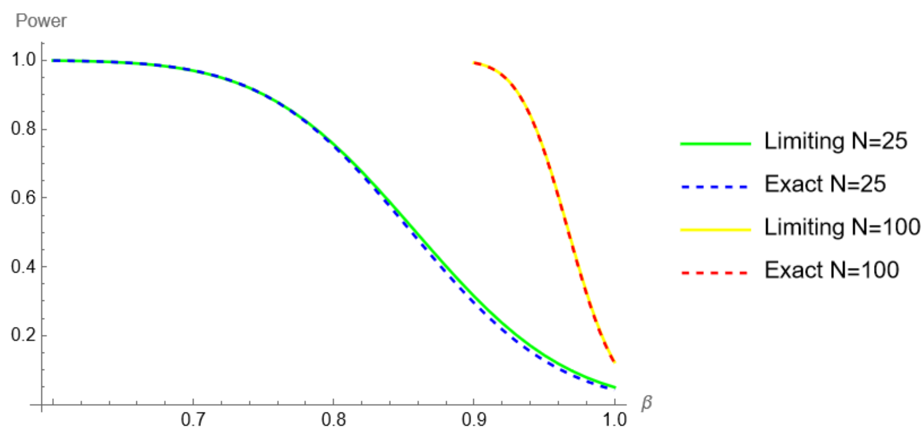

Fig 10: DF coefficient test: Exact and Limiting power($\alpha = 0.05, x_0 = 5$)

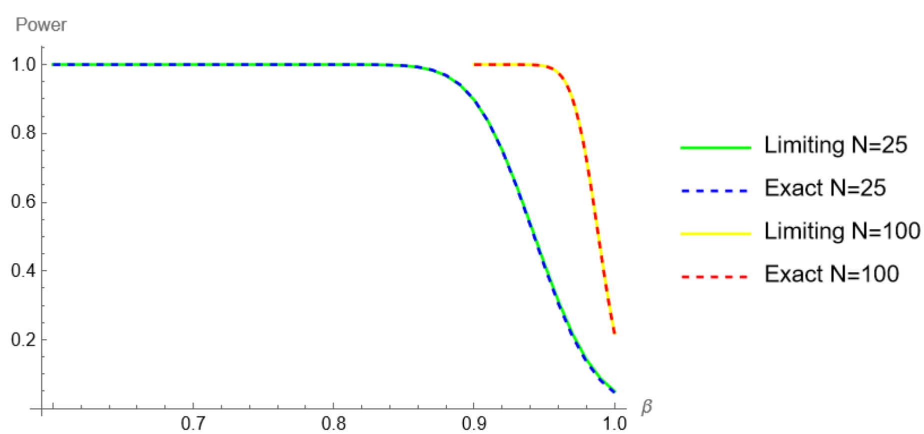

Fig 11: DF coefficient test: Exact and Limiting power($\alpha = 0.05, x_0 = 10$)

5.3 Power of Dickey-Fuller's T Test

In the simulation and numerical computation, we take the variance of the error terms $\sigma^2 = 1$, but we use its estimator s^2 for the t statistic in the simulation. The power of the DF t test can also be approximated by the CDF of normal distribution. Table 3 gives the Taylor coefficient of $\delta_\alpha(\theta)$ up to the fifth order for a significance level $\alpha = 0.05$. Since the coefficients depend on the initial value of the squared Bessel process, they are the same for $\{x_0 = 5, N = 25\}$ and $\{x_0 = 10, N = 100\}$.

	$\delta_\alpha(0)$	$\delta'_\alpha(0)$	$\delta''_\alpha(0)$	$\delta^{(3)}_\alpha(0)$	$\delta^{(4)}_\alpha(0)$	$\delta^{(5)}_\alpha(0)$
$x_0 = 0$	-1.645	-0.232544	0.001133	0.000244	0.000014	0.000003
$x_0 = 5, N = 25$	-1.645	-0.465088	0.004532	0.001954	0.000228	0.000106
$x_0 = 5, N = 100$	-1.645	-0.29068	0.001770	0.000477	0.000035	0.000010
$x_0 = 10, N = 25$	-1.645	-1.15804	0.026764	0.030745	0.007743	0.008912
$x_0 = 10, N = 100$	-1.645	-0.465088	0.004532	0.001954	0.000228	0.000106

Tab 4: Taylor coefficient for $\delta_\alpha(\theta)$

Figure 12-14 shows that1st order Taylor expansion for $\delta_\alpha(\theta)$ has good approximation to the power computed by the limiting distribution for different initial value,

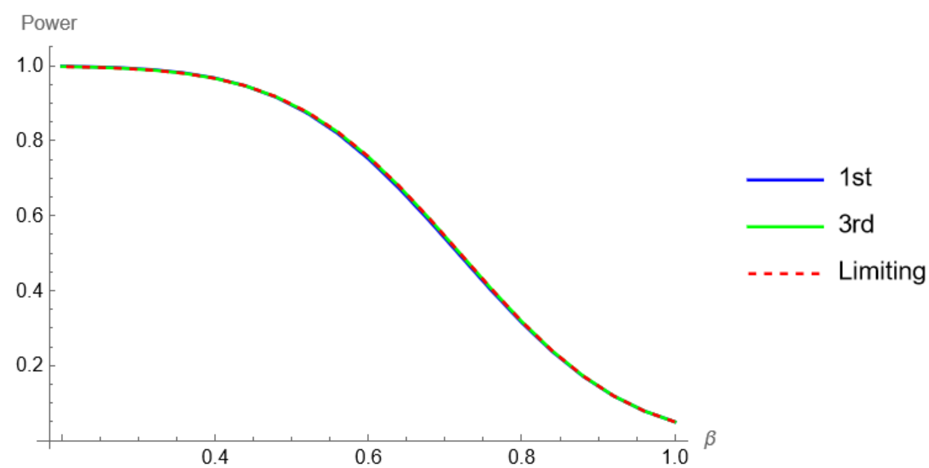

Fig 12: DF's t test: Approximation by normal distribution ($\alpha = 0.05, x_0 = 0, N = 25$)

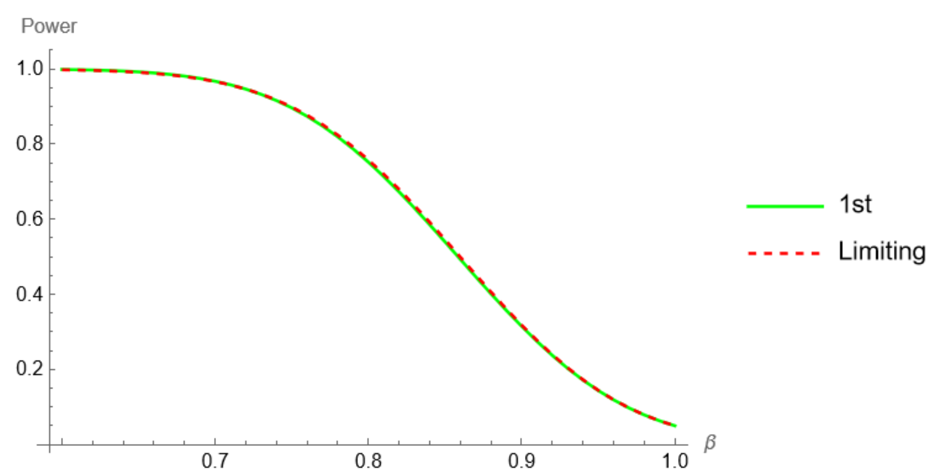

Fig 13: DF's t test: Approximation by normal distribution ($\alpha = 0.05, x_0 = 5, N = 25$)

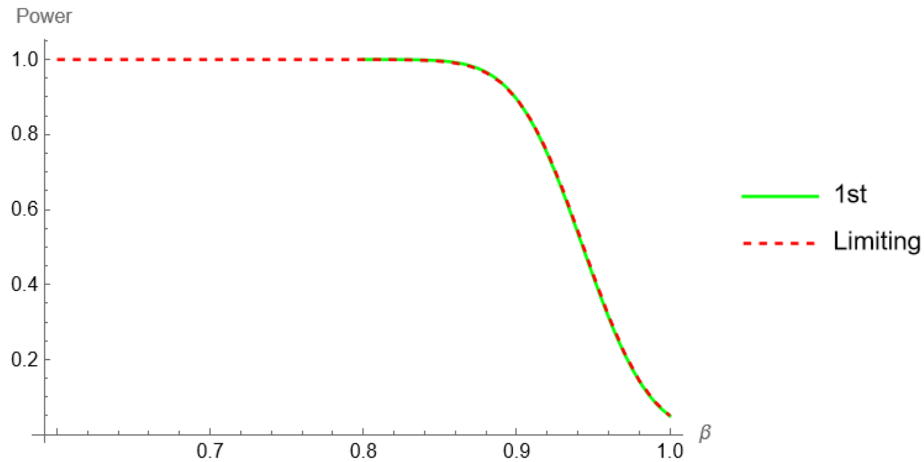

Fig 14: DF's t test: Approximation by normal distribution ($\alpha = 0.05, x_0 = 10, N = 25$)

As the sample size grows, the simulation results get closer to the limiting distribution. Figures 15-17 show the consistency of the limiting distribution for different initial value. In the simulation works, we take 100000 times iterations, and we use the estimated variance s^2 for the t statistic.

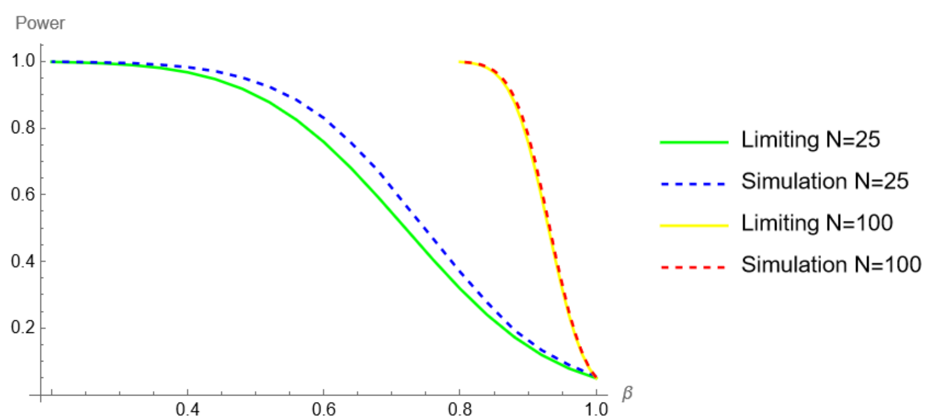

Fig 15: DF's t test: Simulation and Limiting power($\alpha = 0.05, x_0 = 0$)

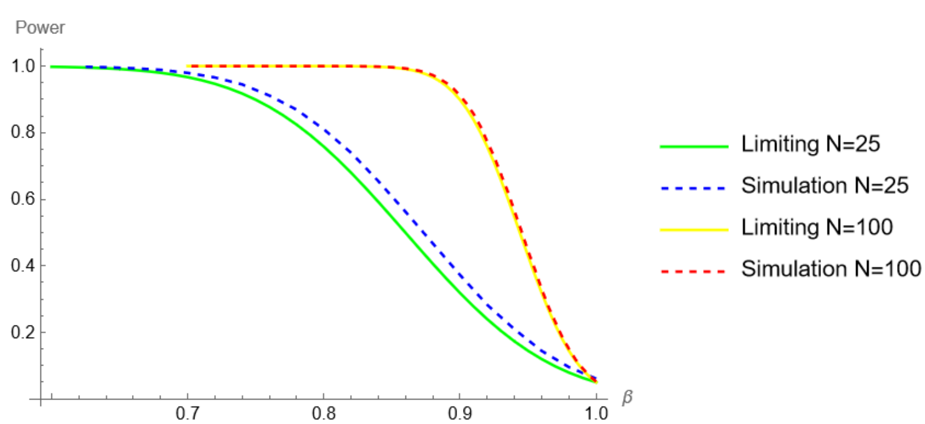

Fig 16: DF's t test: Simulation and Limiting power($\alpha = 0.05, x_0 = 5$)

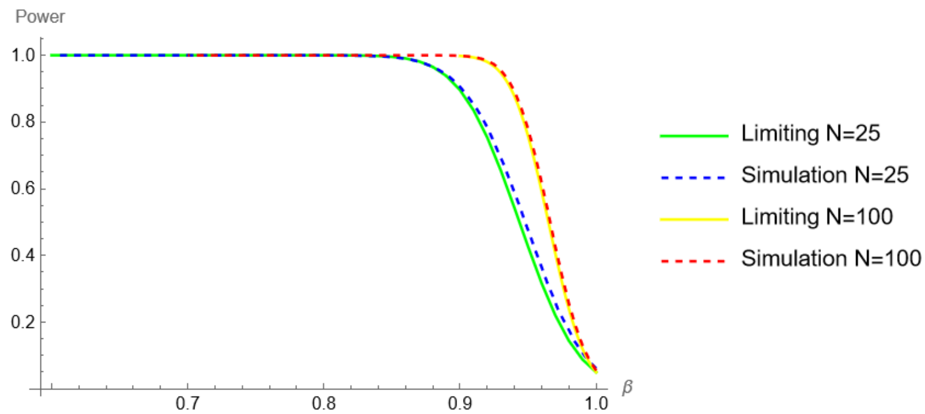

Fig 17: DF's t test: Simulation and Limiting power($\alpha = 0.05, x_0 = 10$)

6 Conclusion

The joint density of the squared Bessel process and its integral provides a new way to compute the limiting distribution of Dickey Fuller's test statistics for the AR(1) model with an initial value. Further, for independent normal errors with mean 0 and variance σ^2, Imhof's formula allows us to compute the exact distribution of Dickey Fuller's coefficient test statistic, and the computation is quick as long as the sample size is small. In comparing these two approaches, the limiting distribution from the first approach fits well with the exact distribution from the second approach even for a small sample. The larger the sample size, the higher the goodness of fit. The first approach needs the time-consuming computation of double integrals of triple sums of the functions including parabolic cylinder functions. We propose a concise method for computing powers using the CDF of the normal distribution. The problem of the unit root test is transformed into the problem of testing for a shift in the location parameter of the normal distribution. It is always good enough to compute the first order Taylor expansion of the shift for stationary alternative hypotheses.

References

[1] Abadir, K. M. (1995a). The Limiting Distribution of the t Ratio Under a Unit Root. *Econometric Theory* **11**(4), 775–793. http://doi.org/10.1017/s0266466600009737

[2] Abadir, K. M. (1995b). The joint density of two functionals of a Brownian motion. Mathematical Methods of Statistics, 4(4), 449-462.

[3] Bobkoski, M. J. (1983). Hypothesis testing in nonstationary time series. Ph.D. thesis, Univ. Wisconsin-Madison.

[4] Borodin, A.N. and Salminen, P. (2002) Handbook of Brownian Motion—Facts and Formulae. 2nd Edition, Birkhauser, Basel. http://dx.doi.org/10.1007/978-3-0348-8163-0

[5] Cavanagh, C. (1985). Roots Local To Unity. *Manuscript, Department of Economics, Harvard University.*

[6] Chan, N. H. (1988). The parameter inference for nearly nonstationary time series. *J. Amer. Statist. Assoc.* **83**(403), 857-862. https://doi.org/10.1080/01621459.1988.10478674

[7] Chan, N. H. and Wei, C. Z. (1987). Asymptotic Inference For Nearly Nonstationary AR(1) Processes. *Ann. Statist.* **15**(3), 1050-1063. https://doi.org/10.1214/aos/1176350492

[8] Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association, 74(366a), 427-431.

[9] Gil-Pelaez, J. (1951). Note on the Inversion Theorem. Biometrika, 38(3/4), 481–482.

[10] Hitomi, K., Nagai, K., Nishiyama, Y. and Tao, J. (2021). Joint asymptotic properties of stopping times and sequential estimators in stationary first-order autoregressive models. KIER Discussion Paper 1060, Kyoto University

[11] Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. Biometrika, 48(3/4), 419-426.

[12] Nagai, K., Nishiyama, Y., & Hitomi, K. (2018). Sequential test for unit root in AR (1) model. KIER Discussion Paper, 1003, 1-27.

[13] Perron, P. (1991). A continuous time approximation to the unstable first-order autoregressive process: The case without an intercept. Econometrica: Journal of the Econometric Society, 211-236.

[14] Phillips, P. C. B. (1987a). Time Series Regression with a Unit Root. *Econometrica* **55**(2), 277-301. https://doi.org/10.2307/1913237

[15] Phillips, P. C. B. (1987b). Towards a Unified Asymptotic Theory for Autoregression. *Biometrika* **74**(3), 535-547. https://doi.org/10.1093/biomet/74.3.535

[16] Pitman, J., & Yor, M. (1982). A decomposition of Bessel bridges. Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete, 59(4), 425-457.

[17] Revuz, D., & Yor, M. (2013). Continuous martingales and Brownian motion (Vol. 293). Springer Science & Business Media.

[18] Tanaka, K. (1996). Time series analysis: nonstationary and noninvertible distribution theory. Wiley, New York.

[19] White, J.S. (1958). The Limiting Distribution of the Serial Correlation Coefficient in the Explosive Case. *Ann. Math. Statist.* **29**(4), 1188-1197. https://doi.org/10.1214/aoms/1177706450