博士論文

Tests of Criticality for Branching Process with Immigration

横浜国立大学大学院
国際社会科学府

唐 越之
Tang Yuezhi

2023年3月
March 2023

Tests of Criticality for Branching Process with Immigration

Yuezhi Tang

March 2023

Abstract

For branching processes with immigration, this study considers testing criticality of offspring mean against sub (or super) close-to-criticality alternative hypotheses. The effect of initial values, sample sizes on the test results and the effectiveness of the local alternatives are discussed. First, fixed-sample-size criticality tests (FCT) are advocated. Second, sequential criticality tests (SCT) using stopping times based on observed Fisher information are explored. Third, to clarify the effectiveness of the testing methods for the local hypotheses in the above two tests, comparisons are made with the testing methods for non-local stationary hypotheses: for the FCT, comparisons of power are made, and for the SCT, comparisons of the joint moments of the stopping time and sequential test statistics are made. It is checked whether the simulation results conform to the theoretical values computed from the local model or from the stationary model.

The FCT in this study is analyzed in the same method as the Dickey-Fuller unit root test in autoregressive models. Namely, the limiting test statistic is represented as the ratio of a squared Bessel process (BESQ) and its integral, whose joint density can be derived from the theory of Bessel bridge in Pitman and Yor (1982). Rejection regions for the left-tailed and right-tailed criticality tests are determined and asymptotic powers are computed. It can be seen that as the initial value increases, the power of the test increases. The dimension of a BESQ process determines the critical value of the FCT rejection region. For the estimated dimension, the critical value is determined in the following way. Critical values for sufficiently many dimensions are calculated, and that for the estimated dimension is determined by linear interpolation. Numerical results show that this interpolation method works well for left-sided critical tests with sample sizes above 100 and right-sided critical tests with sample sizes above 50.

The test statistic for the SCT is found to be represented as a time-changed Brownian motion with the local parameter as drift, indicating that the SCT is actually a Z-test and that the initial value have no effect on the rejection region or the power. On the other hand, the stopping time is terminated earlier as the initial value is larger. The SCT also sets an upper bound on the stopping time based on the observed Fisher information. One attempt in this study is that the upper bound is placed at the 99th percentile point of the distribution of the stopping time under the null hypothesis. This prevents the sample size of the sequential test from becoming too large. If the stationary alternative hypothesis is true, the stopping time tend to be terminated later. Therefore, a test is also performed to reject the null hypothesis when the stopping time exceeds the 99 percentile point.

In the case of the FCT, as is well known, the test statistic is normally distributed under the non-local stationary alternatives when the initial value can be neglected. Thus, the powers of the non-local FCT are computed via Z-test, and compared to those of the local FCTs. For the SCT, the joint moments can be computed via the joint Laplace transform of the stopping time and the sequential test statistic. The theoretical values computed from the local model fits well with the simulation outcomes for offspring mean close to 1. It is possible to specify the appropriate

range of offspring means when local alternative hypotheses are valid, and thereby also the range of offspring means when non-local stationary alternative hypotheses are effective.

Contents

1 Introduction **4**

2 Model Setting **5**

I Non-sequential Criticality Tests **6**

3 Limiting approximations of test statistic **6**
3.1 Asymptotic distribution of test statistic 8
3.2 Joint density of $(q_t, \int_0^t q_s ds)$ 9

4 Test power and simulation results **11**
4.1 asymptotic CDFs and simulations of the test statistic 11
4.2 Power of criticality tests 12

5 Conclusions **17**

II Sequential Criticality Tests with Stopping Time **17**

6 Stopping times based on the observed Fisher information **18**
6.1 asymptotic distribution of $\tau_c/\sqrt{c}$ 20

7 Joint asymptotic distribution of sequential test statistic and stopping time **21**
7.1 Joint density of $(B_1 + \delta, U_1)$ 22

8 Stopping time with upper bound **23**

9 Conclusion **25**

III Effectiveness of Local Model **25**

10 Effectiveness of Local Model in Non-Sequential Analysis **26**

11 Effectiveness of Local model in Sequential Analysis **26**
11.1 Joint moment of (ρ_v, U_v) 27
11.2 Comparisons of the joint moments of (B_1, U_1) 29

12 Conclusion **32**

IV Appendix & Reference **32**

Acknowledgement

I would like to express my gratitude to my supervisors at Yokohama National University, especially Prof. Keiji Nagai, who helped me a lot during my six years of master's and doctoral studies, and Prof. Minako Fujio and Associate Prof. Chizuo Furukawa, who helped me to review my doctoral thesis. I also thank Prof. Kotaro Hitomi of Kyoto Institute of Technology and Prof. Yoshihiko Nishiyama of Kyoto University for allowing me to study their theories on the branching process of migration.

1 Introduction

Branching processes serve as a mathematical model of a population in which each individual in generation n produces some random number of offspring in generation $n+1$. This study considers the simplest case in which the offspring of each individual in each generation have the same fixed probability distribution that does not vary from generation to generation, or from individual to individual. Branching processes can be used to model not only population growth, but also cell kinetics, spreads of an infectious disease, etc.

A central problem in the theory of branching processes is whether the population in question is to explode or to extinct after some finite number of generations. It can be shown that starting with one individual in generation zero, the expected size of generation n equals to m^n where m is the expected number of offspring of each individual. If $m < 1$, the expected number of individuals goes rapidly to zero, which implies ultimate extinction with probability 1 by Markov's inequality. This study uses a branching model with nonnegative immigration in each generation, so that the population does not become extinct.

Since the behaviors of branching processes differ depending on whether the offspring mean m is greater or less than 1. When $m = 1$ ($m < 1$, $m > 1$), the process is called a critical (resp. subcritical, supercritical) process. This study considers tests for the criticality of the offspring mean against the local alternatives including both super-critical and sub-critical hypotheses.

This study is organized as follows. Section 2 describes the model and testing hypotheses, and gives the estimator of offspring mean as a test statistic and estimators of other parameters with preliminary asymptotic results.

I with respect to the fixed-sample-size criticality test (FCT) gives the asymptotic distribution of the test statistic, and computes the power of the criticality test including both sub and super critical local alternative hypotheses. Numerical calculations with Mathematica and simulations with R are carried out to validate the asymptotic theory, considering both zero and positive initial values.

II with respect to the sequential criticality test (SCT) adopts a stopping rule based on the observed Fisher information of the offspring mean, and finds that the sequential criticality test is a Z-test via the Dambis-Dubins-Schwarz (D.D.S.) time-changed theorem. Furthermore, to prevent the stopping time from being terminated late, the second stopping rule is defined by truncating the first stopping rule at the 99th quantile of the distribution under the critical hypothesis. Numerical calculations and simulations are conducted to show that the asymptotic theory with respect to the two stopping times fits well with the simulation results. A combined sequential criticality test is proposed.

III validates the effectiveness of the testing methods used in the FCT and the SCT. It is checked whether the simulation results conform to the theoretical values computed from the local model or from the stationary model. It turns out that power performances of the FCT in I are always better than that computed by a Z-test under the stationary alternative hypotheses. As for SCT, the joint moments of the stopping time and sequential test statistic are computed via the joint Laplace transform, and the results show that when the offspring mean is close to 1, the theoretical values of the moments calculated from the local model agree well with the simulation results.. This made it possible to identify the appropriate range of offspring means when local alternative hypotheses are effective.

2 Model Setting

A Branching process $\{Z_n\}$ with immigration $\{Y_n\}$ is given by the following;

$$Z_n = \sum_{k=1}^{Z_{n-1}} \xi_{n,k} + Y_n \tag{1}$$

where Z_n represents the population size of generation n, $\xi_{n,k}$ is the number of offspring of the k-th individual in period n and Y_n is the number of immigration in period n. Let $\{\xi_{n,k}\}$ and $\{Y_n\}$ be independent stochastic sequences consisting of independent random variables, and let each sequence be independently and identically distributed with a non-negative integer-valued distribution; that is, $\xi_{n,k} \sim i.i.d.(m, \sigma^2) \perp\!\!\!\perp Y_n \sim i.i.d.(\lambda, \sigma_Y^2)$. The initial value of Z_0 is a random variable which is independent of $\{\xi_{n,k}\}$ and $\{Y_n\}$.

Main purpose of this study is to test the criticality of offspring mean m. The maximum likelihood estimator (M.L.E.) of m based on a given sample (Z_n, Y_n) of sample size N is obtained as follows (see Lemma 14 in Appendix IV for derivation);

$$\hat{m}_N = \frac{\sum_{n=1}^{N}(Z_n - Y_n)}{\sum_{n=1}^{N} Z_{n-1}} \tag{2}$$

The left-tailed test examines the subcritical null hypothesis against the supercritical alternative hypotheses;

$$H_0 : m \leq 1 \text{ vs } H_1 : m > 1.$$

The right-tailed test confirms the supercritical null hypothesis against the subcritical alternative hypotheses;

$$H_0 : m \geq 1 \text{ vs } H_1 : m < 1.$$

Considering the situation where the offspring mean m approaches 1 as sample size N increases, large, the statistical properties of the test with the MLE are investigated by expressing m using a local parameter δ as:

$$m = 1 + \frac{\delta}{N}. \tag{3}$$

Thus, the hypotheses for the left-tailed sided test and the right-tailed are formulated respectively as follows:

$$\begin{aligned} H_0 :&\delta \leq 0 \text{ vs } H_1 : \delta > 0, \\ H_0 :&\delta \geq 0 \text{ vs } H_1 : \delta < 0. \end{aligned}$$

The test statistic in the criticality test is defined as

$$\hat{\delta}_N \equiv N(\hat{m}_N - 1). \tag{4}$$

The estimators of parameters σ^2, λ and σ_Y^2 are

$$\begin{aligned} s_N^2 =& \frac{1}{N}\sum_{n=1}^{N} \frac{(Z_n - Y_n - \hat{m}_N Z_{n-1})^2}{Z_{n-1}} 1_{\{Z_{n-1}>0\}} \\ \hat{\lambda}_N =& \bar{Y}^{(N)}, \quad \sigma_{Y,N}^2 = \frac{1}{N}\sum_{n=1}^{N}(Y_n - \bar{Y}^{(N)})^2 \end{aligned} \tag{5}$$

where $\bar{Y}^{(N)}$ is the sample mean of $Y_1, \ldots, Y_N$.

Part I
Non-sequential Criticality Tests

This part examines the fixed-sample-size criticality tests (FCT) .

In Section 3, the discrete-time branching process $\{Z_n\}_{n=1,2,...}$ with immigration $\{Y_n\}_{n=1,2,...}$ in (1) with close-to-one offspring mean is first approximated to the continuous-time squared Bessel (BESQ) process q_t with drift δ for $t \in [0,1]$, and the limit of test statistic $\hat{\delta}_N$ in (4) is found to be the ratio of q_1 and integral $\int_0^1 q_s\, ds$. Combining the conditional Laplace transform $E[\exp(-\gamma \int_0^t q_s ds)|q_t]$, called Pitman and Yor's (1982) Bessel bridge, and the transition density of q_t, the joint density of $(q_1, \int_0^1 q_s\, ds)$ can be derived to obtain the asymptotic distribution of $\hat{\delta}_N$.

In Section 4, the theoretical values of the asymptotic distribution of $\hat{\delta}_N$ and the asymptotic power of the FCT are calculated using Mathematica and shown to be consistent with the simulation results. The critical value of the rejection region of the FCT is determined by the initial value and dimension of the BESQ process. The asymptotic test powers are calculated in the following steps. First, the critical values for a sufficiently large number of combinations of initial values and dimensions are calculated in advance. Second, the critical value corresponding to the initial value and estimated dimension obtained from the data is determined by linear interpolation among those values calculated in the first step. Numerical results show that this interpolation method performs successfully for left-sided critical tests with sample sizes greater than 100 and for right-sided critical tests with sample sizes greater than 50. The initial value and sample size affect the power of the FCT, i.e. the power increases as the initial value increases and as the sample size N increases.

3 Limiting approximations of test statistic

Let $\{\mathcal{F}_n\} = \{\sigma(Z_1, Z_2, ..., Z_n, Y_1, Y_2, ..., Y_n)\}$, then

$$E[Z_n|\mathcal{F}_{n-1}] = mZ_{n-1} + \lambda.$$

Define $\varepsilon_n \equiv Z_n - E[Z_n|\mathcal{F}_{n-1}] - (Y_n - \lambda)$, so that

$$\varepsilon_n = Z_n - Y_n - mZ_{n-1} = \sum_{k=1}^{Z_{n-1}} (\xi_k^{(n)} - m)$$

which implies ε_n is a martingale difference with respect to $\mathcal{F}_n$;

$$E[\varepsilon_n|\mathcal{F}_{n-1}] = \sum_{k=1}^{Z_{n-1}} E[\xi_k^{(n)} - m] = 0.$$

The following Lemma 1 provides the fundamental functional central limit theorem for the diffusion approximation of the branching process with immigration derived in Lemma 2. Let $D[0,1]$ be the set of continuous functions with left limits (càdlàg functions) on $[0,1]$ and "$\Rightarrow$" denote weak convergence. See Billingsley (1999) for the weak convergence on $D[0,1]$.

Lemma 1. *(Martingale difference and Brownian motion) Assuming that offspring $\xi_k^{(n)}$ has the fourth moment. Then, as $N \to \infty$*

$$\sum_{n=1}^{\lfloor Nt \rfloor} \frac{\varepsilon_n}{\sigma\sqrt{NZ_{n-1}}} \Rightarrow W_t \text{ on } D[0,1]; \tag{6}$$

where W_t is a Brownian motion for $0 \le t \le 1$.

Proof. This result holds due to Theorem 15: fundamental functional central limit theorem, since the second moment of ε_n is $E[\varepsilon_n^2|\mathcal{F}_{n-1}] = \sigma^2 Z_{n-1}$. □

Assuming that initial value $Z_0/N \to X_0$ as $N \to \infty$, the following Lemma 2 shows that the branching process with immigration can be approximated to a Cox-Ingersoll-Ross (CIR) process with initial value X_0. The existence of such an X_0 is not plausible, but since N is fixed in empirical analysis or simulation, there is no choice but to set $X_0 = Z_0/N$. In doing so, comparisons of the simulations with the theoretical results derived from the following proposition show that the influence of initial values is correctly illustrated. See Definition 16 for the CIR process.

Lemma 2. *(Approximation of branching process) Assume that $Z_0/N \to X_0$ as $N \to \infty$ with L^2 random variable X_0. For $t \in [0,1]$, let $X_t^{(N)} \equiv Z_{\lfloor Nt \rfloor}/N$ and X_t be a CIR process as follows:*

$$X_t = X_0 + \sigma \int_0^t \sqrt{X_s} dW_s + \delta \int_0^t X_s ds + \lambda t.$$

Then, as $N \to \infty$, on $D[0,1]$,

$$\left(X_t^{(N)}, \frac{\sum_{n=1}^{\lfloor Nt \rfloor}(Z_n - Y_n - Z_{n-1})}{N}, \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_{n-1}}{N^2} \right) \tag{7}$$

$$\Rightarrow \left(X_t,\, X_t - X_0 - \lambda t,\, \int_0^t X_s ds \right) \tag{8}$$

where W_t is a Brownian motion constructed in (6).

Proof. Using a telescoping sum and Lemma 1,

$$\begin{aligned}
\frac{Z_{\lfloor Nt \rfloor}}{N} &= \frac{Z_0}{N} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_n - Y_n - Z_{n-1}}{N} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Y_n}{N} \\
&= \frac{Z_0}{N} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_n - Y_n - mZ_{n-1}}{N} + (m-1)\frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_{n-1}}{N} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Y_n}{N} \\
&= \frac{Z_0}{N} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} \sum_{k=1}^{Z_{n-1}} (\xi_{n,k} - m)}{N} + +\delta \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_{n-1}}{N^2} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Y_n}{N} \\
&= \frac{Z_0}{N} + \sigma \sum_{n=1}^{\lfloor Nt \rfloor} \sqrt{\frac{Z_{n-1}}{N}} \frac{\varepsilon_n}{\sigma\sqrt{NZ_{n-1}}} + \delta \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_{n-1}}{N^2} + \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Y_n}{N} \\
&\Rightarrow X_t = X_0 + \sigma \int_0^t \sqrt{X_s} dW_s + \delta \int_0^t X_s ds + \lambda t.
\end{aligned}$$

The above approximation includes

$$\begin{aligned}
&\left(\frac{\sum_{n=1}^{\lfloor Nt \rfloor}(Z_n - Y_n - Z_{n-1})}{N}, \frac{\sum_{n=1}^{\lfloor Nt \rfloor} Z_{n-1}}{N^2} \right) \\
\Rightarrow &\left(\sigma \int_0^t \sqrt{X_s} dW_s + \delta \int_0^t X_s ds, \int_0^t X_s ds \right) \\
= &\left(X_t - X_0 - \lambda t, \int_0^t X_s ds \right).
\end{aligned}$$

□

Furthermore, the CIR process is transformed into a squared Bessel process with drift. The squared Bessel process with drift is a Girsanov transformation of the ordinary squared Bessel (BESQ) process. See Definition Definition 17 for the BESQ process. Thus, $q_t \equiv 4X_t/\sigma^2$ is a $4\lambda/\sigma^2$-dimensional squared Bessel process with drift δ;

$$\begin{aligned} q_t &= \frac{4X_0}{\sigma^2} + 2\int_0^t \sqrt{\frac{4X_s}{\sigma^2}} dW_s + \delta \int_0^t \frac{4X_s}{\sigma^2} ds + \frac{4\lambda}{\sigma^2} t \\ &= q_0 + 2\int_0^t \sqrt{q_s} dW_s + \delta \int_0^t q_s ds + \frac{4\lambda}{\sigma^2} t \end{aligned} \tag{9}$$

where $q_0 \equiv 4X_0/\sigma^2$. Under the null hypothesis,

$$q_t = q_0 + 2\int_0^t \sqrt{q_s} dW_s + (4\lambda/\sigma^2)t$$

which is denoted as $q_t \sim \text{BESQ}^d_{q_0}$. The testing hypotheses are reduced to the followings in continuous time:

$$\text{Under} P^0 : dq_t = 2\sqrt{q_t} dW_t + 4\lambda/\sigma^2 \, dt, \tag{10}$$

$$\text{Under} P^\delta : dq_t = 2\sqrt{q_t} dW_t + (4\lambda/\sigma^2 + \delta q_t) dt. \tag{11}$$

3.1 Asymptotic distribution of test statistic

The estimator of the local parameter δ is used as a test statistic. It can be seen that the test statistic is expressed as the ratio of the squared Bessel process with drift and its integral.

Lemma 2 implies that as $N \to \infty$, the test statistic $\hat{\delta}_N$ in (4) has the following limiting approximation;

$$\begin{aligned} \hat{\delta}_N &= N(\hat{m}_N - 1) \\ &= \frac{\sum_{n=1}^N (Z_n - Y_n - Z_{n-1})}{\sum_{n=1}^N Z_{n-1}} \\ &\Rightarrow \frac{X_1 - X_0 - \lambda}{\int_0^1 X_s ds} = \frac{\sigma \int_0^1 \sqrt{X_s} dW_s}{\int_0^1 X_s ds} + \delta. \end{aligned} \tag{12}$$

Then using $q_t = 4X_t/\sigma^2$, the ratio representation is obtained;

$$\hat{\delta}_N \Rightarrow \frac{q_1 - q_0 - d}{\int_0^1 q_s ds} \tag{13}$$

where $d = 4\lambda/\sigma^2$ is the dimension of q_t.

Thus, once the joint density of $(q_t, \int_0^t q_s ds)$ with respect to the BESQ process q_t with drift δ and initial value q_0 is obtained, the limiting CDF of the test statistic $\hat{\delta}_N$ is derived as follows:

$$\begin{aligned} F(z) &= P\left(\frac{q_1 - q_0 - d}{\int_0^1 q_s ds} \le z\right) \\ &= \begin{cases} \int_0^{q_0+d} \int_0^{(y-q_0-d)/z} f_{q_1, \int_0^1 q_s ds}(y, v) \, dv dy & z < 0 \\ \int_0^{q_0+d} f_{q_1}(y) \, dy & z = 0 \\ 1 - \int_{q_0+d}^{\infty} \int_0^{(y-q_0-d)/z} f_{q_1, \int_0^1 q_s ds}(y, v) \, dv dy & z > 0 \end{cases} \end{aligned} \tag{14}$$

where $f_{q_1}(y)$ is the PDF of q_1, given in (17).

3.2 Joint density of $(q_t, \int_0^t q_s ds)$

The following is function is used to derive the joint density of $(q_t, \int_0^t q_s ds)$ under H_0 : $\delta = 0$;

$$\mathrm{is}_y(\nu, t, r, z, x) = \mathscr{L}_\gamma^{-1}\left(\frac{\sqrt{2\gamma}}{\sinh t\sqrt{2\gamma}} \exp\left\{-r\sqrt{2\gamma} - \frac{z\sqrt{2\gamma}\cosh t\sqrt{2\gamma}}{\sinh t\sqrt{2\gamma}}\right\} I_\nu\left(\frac{2x\sqrt{2\gamma}}{\sinh t\sqrt{2\gamma}}\right)\right)(y)$$

where $\mathscr{L}_\gamma^{-1}\left(\int_0^\infty e^{-\gamma y} f(y)dy\right)(y) = f(y)$ is the inverse Laplace transform, $\nu = d/2 - 1 > -1/2$ is the index of BESQ process with dimension $d > 0$, and I_v is the modified Bessel function defined as

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+\nu+1)}\left(\frac{x}{2}\right)^{2k+\nu}. \tag{15}$$

For the definition of the is, see Definition 19 in Appendix and Borodin & Selminen (2002)[6].

Theorem 3. *The joint density of $(q_t, \int_0^t q_s ds)$ with respect to BESQ_x^d with positive initial value $x > 0$ under the null hypothesis has the following representation with $\nu = d/2 - 1$;*

$$\begin{aligned}
f_{q_t, \int_0^t q_s ds}(y, v) =& \frac{1}{2}\left(\frac{y}{x}\right)^{\nu/2} \mathrm{is}_v(\nu, t, 0, \frac{x+y}{2}, \frac{\sqrt{xy}}{2}). \qquad (16)\\
=& \frac{1}{2}(\frac{y}{x})^{\nu/2} 2^\nu \sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(\sqrt{xy}/2)^{\nu+2l}(-(x+y)/2)^k}{\sqrt{2\pi} j!k!l!\Gamma(\nu+l+1)}\\
&\times \frac{\exp(-((\nu+2l+1+k)t + (x+y)/2 + kt + 2jt)^2/4v)\Gamma(\nu+2l+1+k+j)}{v^{1+(\nu+2l+1+k)/2}\Gamma(\nu+2l+1+k)}\\
&\times D_{2+\nu+2l+k}\left(\frac{(\nu+2l+1+k)t + (x+y)/2 + kt + 2jt}{\sqrt{v}}\right)
\end{aligned}$$

Proof. According to Pitman&Yor(1982) [15], BESQ process q_t with initial value $x > 0$ and order $v > -1/2$ has a density in y equal to

$$f_{q_t}(y) = q_t^d(x, y) = \frac{1}{2t}\left(\frac{y}{x}\right)^{\nu/2} \exp\left(-\frac{x+y}{2t}\right) I_\nu\left(\frac{\sqrt{xy}}{t}\right), \qquad t > 0 \tag{17}$$

Pitman&Yor(1982) [15] also gives the Bessel bridge of BESQ process, which is the conditional Laplace transformation of $\int_0^t q_s ds$ given q_t;

$$E_x^d[\exp\left(-\frac{b^2}{2}\int_0^t q_s ds\right)|q_t = y] \tag{18}$$

$$= \frac{bt}{\sinh\, bt} \exp\left\{\frac{x+y}{2t}(1 - bt\coth\, bt)\right\} I_\nu\left(\frac{\sqrt{xy}b}{\sinh\, bt}\right) / I_\nu\left(\frac{\sqrt{xy}}{t}\right). \tag{19}$$

Letting $b^2/2 = \gamma$ in the the Bessel bridge and multiplying it by $q_t^d(x, y)$, the expression of the Laplace transformation of the joint density $f_{q_t, \int_0^t q_s ds}(y, v)$ is obtained;

$$\begin{aligned}
&E_x^d[\exp(-\gamma\int_0^t q_s ds)|q_t = y] \cdot f_{q_t}(y)\\
=& \int_0^\infty e^{-\gamma v} f_{\int_0^t q_s ds|q_t}(v|y) f_{q_t}(y)\, dv \qquad (20)\\
=& \int_0^\infty e^{-\gamma v} f_{q_t, \int_0^t q_s ds}(y, v)\, dv\\
=& \frac{1}{2}\left(\frac{y}{x}\right)^{\nu/2} \frac{\sqrt{2\gamma}}{\sinh\, \sqrt{2\gamma}t} \exp\left(-\frac{x+y}{2}\sqrt{2\gamma}\coth\, \sqrt{2\gamma}t\right) I_\nu\left(\frac{\sqrt{xy}\sqrt{2\gamma}}{\sinh\, \sqrt{2\gamma}t}\right) \qquad (21)
\end{aligned}$$

Inverting this Laplace transform with respect to γ yields the first expression in (16), and the last expression is obtained from the formulas in Definition 19. □

For BESQ process with initial value $x = 0$, the joint density $f_{q_t,\int_0^t q_s ds}(y,v)$ is represented via the following es function;

$$\mathrm{es}_y(\mu,\nu,t,x,z) = \mathscr{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\sinh^\nu t\sqrt{2\gamma}}\exp(-x\sqrt{2\gamma} - z\sqrt{2\gamma}\coth\, t\sqrt{2\gamma})\right)(y)$$

For details of es function, see Definition 19 and Borodin & Selminen (2002) [6].

Theorem 4. *Joint density of* $(q_t, \int_0^t q_s ds)$ *with respect to* BESQ_x^d *with initial value* $x = 0$ *under the null hypothesis has the following expression with* $\nu = d/2 - 1$;

$$\begin{aligned} f_{q_t,\int_0^t q_s ds}(y,v) =& \frac{y^\nu 2^{-(\nu+1)}}{\Gamma(\nu+1)}\mathrm{es}_v(\nu+1,\nu+1,t,0,y/2) \qquad (22)\\ =& \frac{y^\nu}{\Gamma(\nu+1)}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-y)^k\exp(-((\nu+1+k)t+y/2+kt+2jt)^2/4v)\Gamma(\nu+1+k+j)}{\sqrt{2\pi}j!k!v^{1+(\nu+1+k)/2}\Gamma(\nu+1+k)}\\ =& \times D_{2+\nu+k}\left(\frac{(\nu+1+k)t+y/2+kt+2jt}{\sqrt{v}}\right). \end{aligned}$$

Proof. The modified Bessel function (15) reduces to the following;

$$\lim_{x\to 0} I_\nu(x) = \lim_{x\to 0}\left(\frac{x}{2}\right)^\nu \frac{1}{\Gamma(\nu+1)}$$

so that the density of q_t in (17) reduces to

$$q_t^d(0,y) = \frac{y^\nu(2t)^{-(\nu+1)}}{\Gamma(\nu+1)}\exp(-\frac{y}{2t}).$$

The Laplace transformation of the joint density in (16) reduces to the following when $x = 0$;

$$\begin{aligned} &\int_0^\infty e^{-\gamma v} f_{\int_0^t q_s ds|q_t}(v|y) f_{q_t}(y)dv\\ =& \frac{2^{-\frac{\nu}{2}-\frac{1}{2}}y^\nu\gamma^{(\nu+1)/2}\exp\left(-\frac{\sqrt{\gamma}y}{\sqrt{2}\coth\sqrt{2}\sqrt{\gamma}t}\right)}{\Gamma(\nu+1)\sinh^{\nu+1}\sqrt{2}\sqrt{\gamma}t}\\ =& \frac{y^\nu(2\gamma)^{(\nu+1)/2}}{\Gamma(\nu+1)\sinh^{\nu+1}t\sqrt{2\gamma}}\exp\left(-\frac{y\sqrt{\gamma}}{\sqrt{2}\coth\, t\sqrt{2\gamma}}\right). \end{aligned}$$

The first expression in (22) follows by inverting this Laplace transform with respect to γ, and the last expression is obtained from the formulas in Definition 19. □

As it has been shown in (11), the BESQ process has a drift δ under the alternative hypothesis. Girsanov transform theorem is used to eliminate the drift and thus obtain the joint density of $(q_t, \int_0^t q_s ds)$ under the alternative hypothesis. Details of Girsanov transform theorem is given in Theorem 20.

Theorem 5. *Joint density of* $(q_t, \int_0^t q_s ds)$ *with respect to* BESQ_x^d *under the alternative hypothesis has the following relationship with that under the null hypothesis;*

$$f^\delta_{q_t,\int_0^t q_s ds}(y,v) = \exp\left\{\frac{\delta}{4}(y - q_0 - d) - \frac{\delta^2}{8}v\right\}\cdot f^0_{q_t,\int_0^t q_s ds}(y,v). \qquad (23)$$

Proof. Letting W_t to be a Brownian motion under P^0, define a Girsanov transformation $d\tilde{W}_t = dW_t - \frac{\delta}{2}\sqrt{q_t}dt$ so that $\tilde{W}_t$ is a Brownian motion under P^δ which has the Radon–Nikodym derivative with respect to P^0;

$$\begin{aligned} L_t \equiv \frac{dP^\delta}{dP^0}|_{\mathcal{F}_t} &= \exp\left(\frac{\delta}{2}\int_0^t \sqrt{q_s}dW_s - \frac{\delta^2}{8}\int_0^t q_s ds\right) \\ &= \exp\left(\frac{\delta}{4}(y - q_0 - d) - \frac{\delta^2}{8}v\right). \end{aligned}$$

□

4 Test power and simulation results

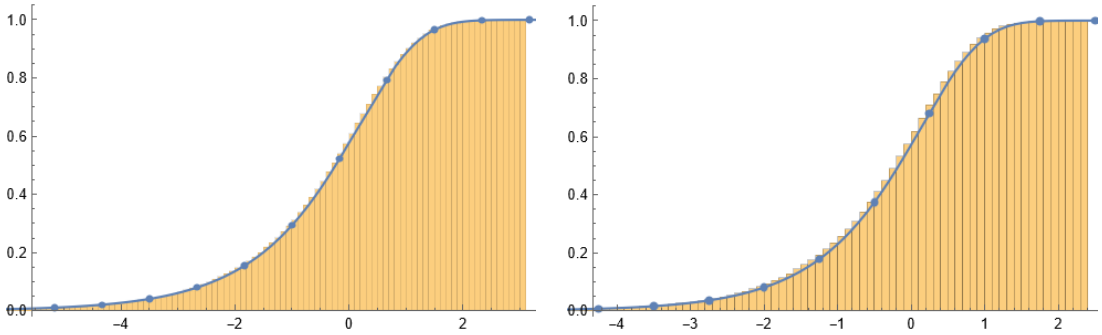
4.1 asymptotic CDFs and simulations of the test statistic

The asymptotic CDFs of the test statistic $\hat{\delta}_N$ in (14) under H_0 with $x > 0$, H_0 with $x = 0$, and H_1 can be computed numerically from (16)(22) and (23) respectively.

Numerical results in Figure 1, Figure 2 and Figure 3 are calculated with parameter $\sigma^2 = 2$, $\lambda = 3$, the initial value of CIR process $X_0 = 0, 1$ respectively. It can be seen that they conform to the simulation outcomes under both the null hypothesis $H_0 : \delta = 0$, the supercritical alternative hypothesis $H_1 : \delta = -2$ and supercritical alternative hypothesis $H_1 : \delta = 2$.

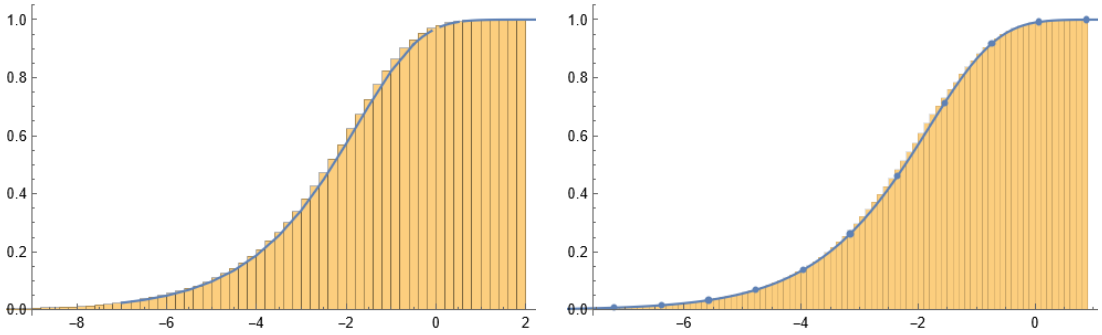
Simulations are conducted with replication times 10000, generation size N=100, offsprings $\xi_{n,k} \sim i.i.d.$Negative Binomial$(k; p)$ and immigrations $Y_n \sim i.i.d.$Poisson(λ). Note that the parameter (k, p) of Negative Binomial distribution correspond to (m, σ^2) through the relationship $p = m/\sigma^2$ and $k = mp/(1 - p)$.

Figure 1: asymptotic CDFs and simulations of $\hat{\delta}_N$ with N=100, $\delta = 0$



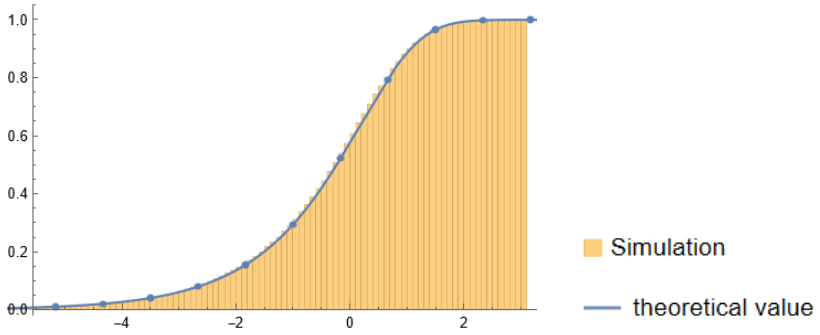
$X_0 = 0$ $X_0 = 1$

Figure 2: asymptotic CDFs and simulations of $\hat{\delta}_N$ with N=100, $\delta = -2$



$X_0 = 0$ $X_0 = 1$

Figure 3: asymptotic CDFs and simulations of $\hat{\delta}_N$ with N=100, $\delta = 2$



$X_0 = 0$

4.2 Power of criticality tests

From (16) and (22), it can be seen that the asymptotic distribution of test statistic $\hat{\delta}_N$ under the null hypothesis depends only on the initial value x and dimension d of the BESQ process.

For $x \geq 0$, $d \geq 0$, the critical value of the reject region for left tailed test δ_* and δ^* for right tailed test at significance level α are defined as

$$\begin{aligned} \text{Left-tailed FCT: } & P^0\left\{\hat{\delta}_N < \delta_*(d,x)\right\} = \alpha, \\ \text{Right-tailed FCT: } & P^0\left\{\hat{\delta}_N > \delta^*(d,x)\right\} = \alpha. \end{aligned} \tag{24}$$

The asymptotic power of each FCT is the probability of rejecting the null hypothesis when the alternative hypothesis is true and can be computed from the theoretical computations of the CDFs.

$$\begin{aligned} \text{Left-tailed FCT: } & P^\delta\left\{\hat{\delta}_N < \delta_*\right\}, \\ \text{Right-tailed FCT: } & P^\delta\left\{\hat{\delta}_N > \delta^*\right\}. \end{aligned}$$

(d, x) is determined on (λ, σ^2) by the relationship $x = 4X_0/\sigma^2$, $d = 4\lambda/\sigma^2$, where the initial value X_0 of CIR process X_t is set as $X_0 = Z_0/N$ and considered to be a known constant. Figure 4 and Figure 5 are the theoretical and simulation results of asymptotic power at significance level $\alpha = 0.05$, with parameters $\lambda = 3, \sigma^2 = 2\,(d = 6)$ and $X_0 = 0, 1\,(x = 0, 2)$ respectively. It turns out that the simulations conform to the asymptotic powers for both left-tailed and right-tailed FCT, and the initial value affects the power in such a way that the power increases as the initial value increases.

Figure 4: Effect of initial value for left tailed test, with $\alpha = 0.05\, d = 6,\ \sigma^2 = 2,\ N = 50$

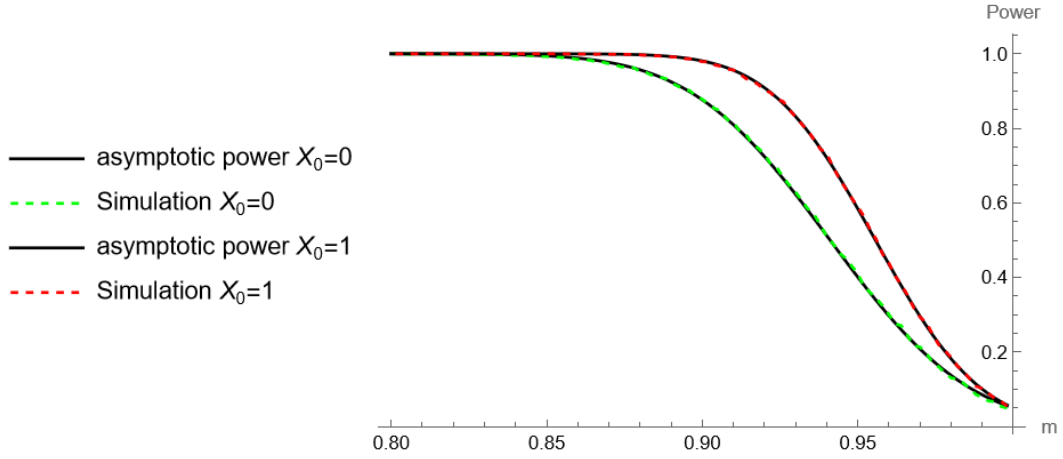
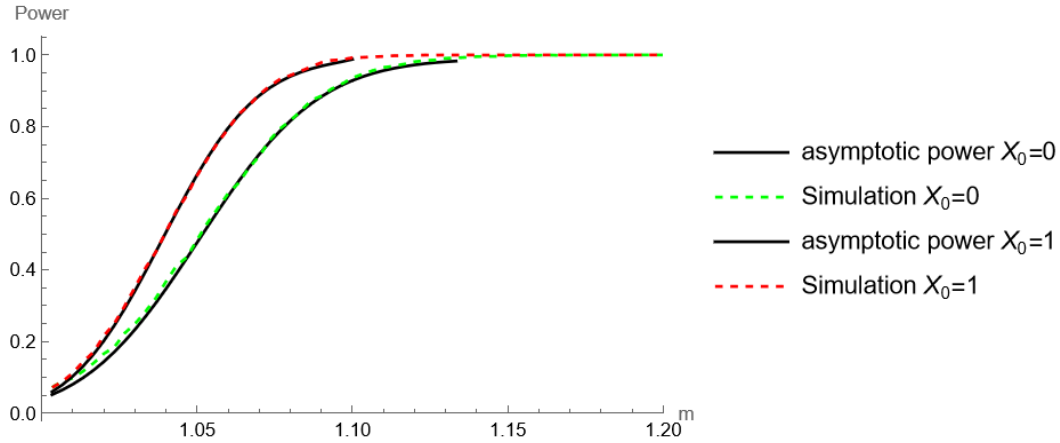


Figure 5: Effect of initial value for left tailed test, with $\alpha = 0.05\, d = 6,\ \sigma^2 = 2\, N = 30$



For a fixed X_0, it has been explained that (d, x) is determined on (λ, σ^2). Therefore the asymptotic distribution of $\hat{\delta}_N$ is actually dependent on the pair (λ, σ^2) under H_0 and so is the critical value of the reject region. The difficulty in empirical analysis is that λ and σ^2 are usually unknown. This study provides a linear interpolation solution via using the estimator $\hat{d}_N = 4\hat{\lambda}_N/s_N^2$ and $\hat{\sigma}_N^2 = s_N^2$ at (5) in the following way. First, the critical values for sufficiently many pairs of $(d, 4X_0/\sigma^2)$ with a fixed X_0 are calculated in advance. Second, the critical value corresponding to $(\hat{d}_N, 4X_0/s_N^2)$ is determined by linear interpolation among those values calculated in the first step.

For simplicity, the work below considers $x = 4X_0/\sigma^2$ to be a known constant. Figure 6 and Figure 7 illustrate the critical values corresponding to a sequence of d starting from 0.1 to 10 in steps of 0.1 at significance level $\alpha = 0.05$, with $\sigma^2 = 2$ and $X_0 = 0, 1$ respectively.

Figure 6: Critical values for left tailed test with $\alpha = 5\%$, $\sigma^2 = 2$

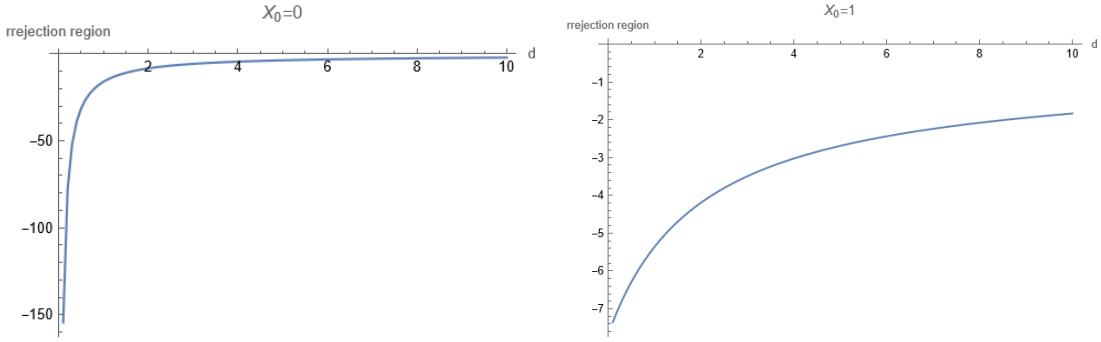


Figure 7: Critical values for right tailed test with $\alpha = 5\%$, $\sigma^2 = 2$

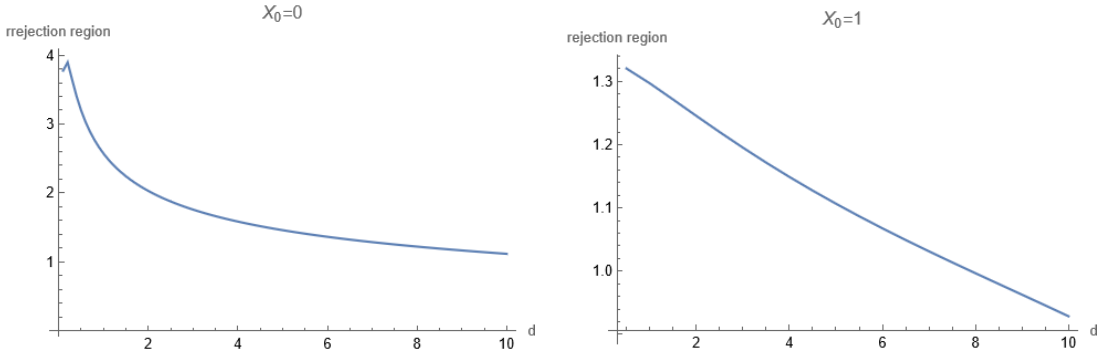


Table 1: Critical values at $\alpha = 5\%, \sigma^2 = 2$

	$X_0 = 0$		$X_0 = 1$	
d	left side	right side	left side	right side
0.5	-31.44715345	1.320331553	-6.289399114	3.197180412
1	-16.0782248	1.296952928	-5.359260793	2.57091506
1.5	-10.95319619	1.271413653	-4.693935738	2.241002715
2	-8.389253303	1.245477305	-4.19434231	2.027988701
2.5	-6.849814339	1.21998296	-3.805456556	1.874612134
3	-5.82265827	1.195320733	-3.494157512	1.756639538
3.5	-5.088303624	1.171650416	-3.239301313	1.661642672
4	-4.537000737	1.149006693	-3.026718935	1.582836035
4.5	-4.107733937	1.127358214	-2.846575108	1.515866565
5	-3.763867919	1.106688336	-2.691822707	1.457923658
5.5	-3.482081135	1.086815601	-2.557308038	1.406993444
6	-3.246808719	1.067689923	-2.439195698	1.361725024
6.5	-3.047279313	1.049233803	-2.334539508	1.321055693
7	-2.875772728	1.031327844	-2.24108737	1.284185846
7.5	-2.726679282	1.01384849	-2.157040181	1.250488304
8	-2.595768712	0.996683356	-2.080977064	1.219455005
8.5	-2.479819484	0.979706422	-2.011762285	1.190666883
9	-2.376330796	0.962796169	-1.948457502	1.16376197
9.5	-2.283327774	0.945825821	-1.890298884	1.138405365
10	-2.199235442	0.928670154	-1.836640777	1.114391217

Since the critical value is monotonic with respect to d, the estimator $\hat{d}_N$ is used to determine the corresponding critical value $\delta_*(\hat{d}_N)$ and $\delta^*(\hat{d}_N)$ by linear interpolation, and thus the rejection regions of FCT are determined as follows;

$$\text{Left-tailed FCT: Reject } H_0 \text{ when } \hat{\delta}_N < \delta_*(\hat{d}_N),$$
$$\text{Right-tailed FCT: Reject } H_0 \text{ when } \hat{\delta}_N > \delta^*(\hat{d}_N).$$

The following figures are the numerical results of the asymptotic powers and simulations of the linear interpolation method.

Figure 8: Power of left tailed test by linear interpolation method, with $\alpha = 5\%$, $d = 6$, $X_0 = 0$

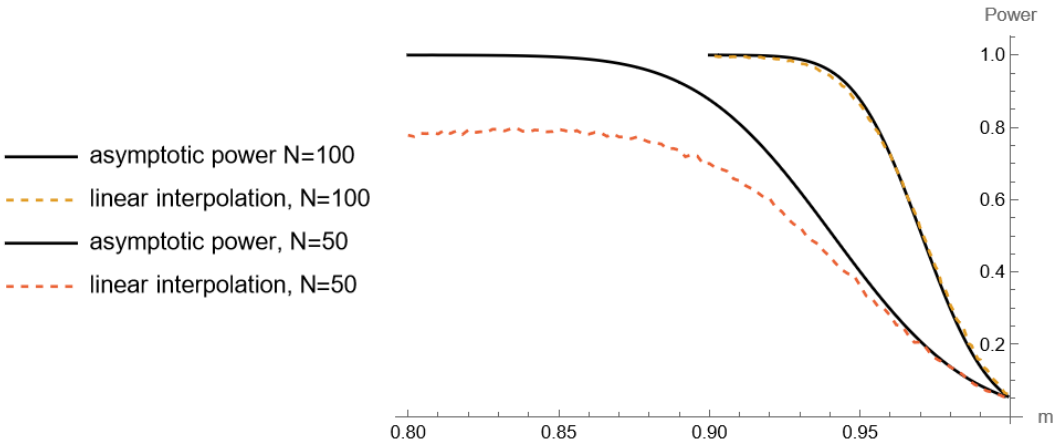


Figure 9: Power of left tailed test by linear interpolation method, with $\alpha = 5\%$, $d = 6$, $X_0 = 1$

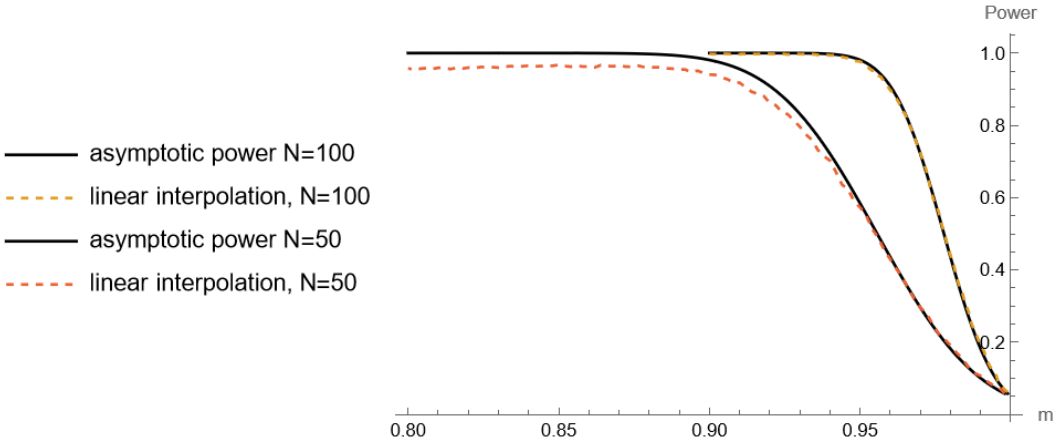


Table 2: Power of the left tailed test with $\alpha = 5\%$, $d = 6$, $N = 100$

	$X_0 = 0$			$X_0 = 1$		
δ	asymptotic power	simulation	linear interpolation	asymptotic power	simulation	linear interpolation
-1	13.76%	14.12%	15.87%	18.21%	17.75%	18.55%
-2	29.75%	30.72%	32.41%	43.59%	44.63%	44.82%
-3	51.27%	50.52%	51.82%	72.18%	72.36%	71.83%
-4	72.44%	72.97%	72.32%	90.90%	91.02%	90.38%
-5	87.67%	87.70%	86.25%	98.09%	98.13%	97.41%
-6	95.70%	95.76%	93.99%	99.74%	99.61%	99.30%
-7	98.83%	98.74%	97.77%	99.98%	99.98%	99.80%
-8	99.75%	99.74%	99.03%	100.00%	99.99%	99.77%
-9	99.96%	99.94%	99.58%	100.00%	100.00%	99.76%
-10	99.99%	100.00%	99.68%	99.99%	100.00%	99.75%

Figure 10: Power of right tailed test by linear interpolation method, with $\alpha = 5\%$, $d = 6$, $X_0 = 0$

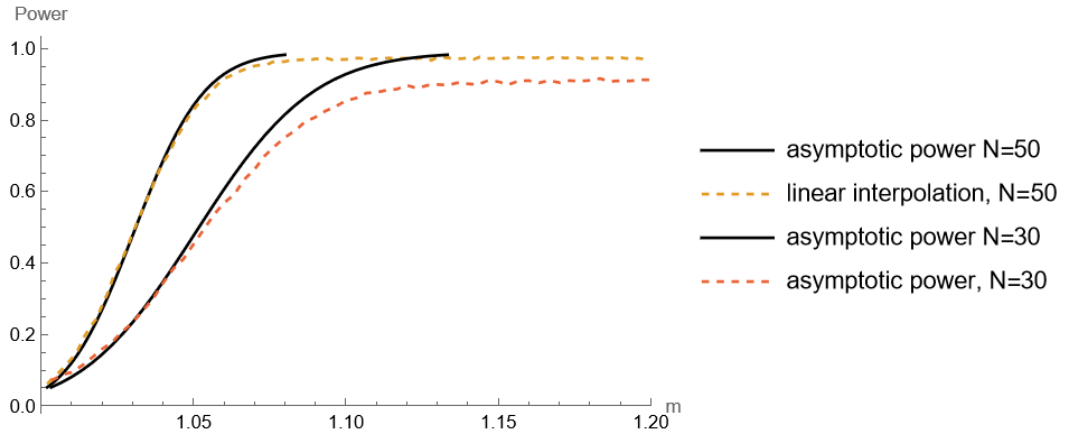


Figure 11: Power of right tailed test by linear interpolation method, with $\alpha = 5\%$, $d = 6$, $X_0 = 1$

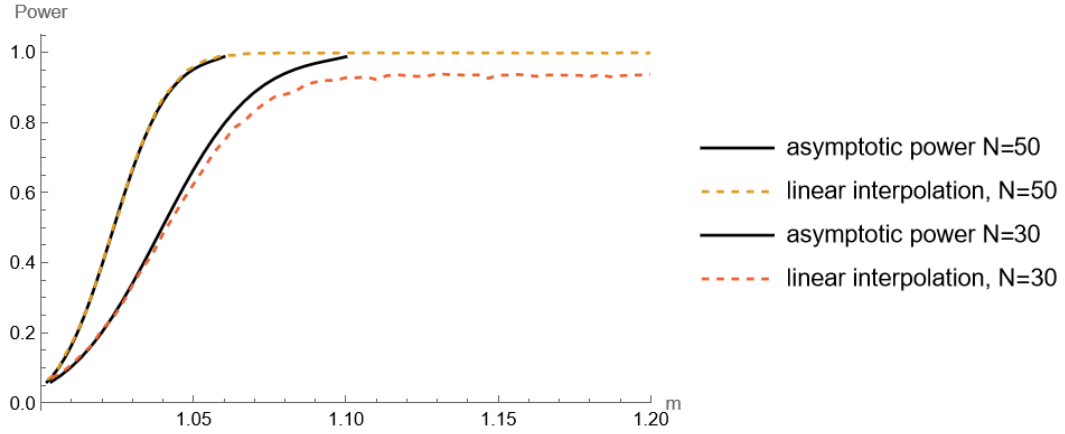


Table 3: Power of the right-tailed tests with $\alpha = 5\%$, $d = 6$, $N = 100$

	$X_0 = 0$			$X_0 = 1$		
δ	asymptotic power	simulation	linear interpolation	asymptotic power	simulation	linear interpolation
0.5	11.99%	13.89%	14.04%	16.43%	16.22%	16.34%
1	26.91%	27.19%	27.48%	39.43%	39.10%	39.11%
1.5	47.62%	48.18%	48.53%	66.52%	66.70%	67.01%
2	68.59%	69.87%	70.24%	86.08%	87.02%	87.23%
2.5	84.12%	84.93%	85.01%	95.17%	95.93%	95.85%
3	92.86%	93.53%	93.56%	98.77%	99.06%	98.94%
3.5	96.78%	97.75%	97.66%			
4	98.28%	99.27%	99.18%			

As it is shown in Figure 8~ Figure 11, for all m, the power of FCT increases as the sample size N increases. The linear interpolation method works well under the subcritical alternative hypothesis and under the supercritical alternative hypothesis, especially in the neighborhood of $m = 1$. More specifically, the powers calculated by the linear interpolation fit with the theoretical results over the entire range of m under the subcritical alternative hypothesis $H_1 : \delta < 0$ with sample size N more than 100, or under the supercritical alternative hypothesis $H_1 : \delta > 0$ with sample size N more than 50. As for the circumstances when sample size N less than 50 for left-tailed test, or when sample size N less than 30 for right-tailed tests, the powers calculated by the linear interpolation do not fit with the theoretical values when δ drift away from 0, but it still fits the theoretical results well at $m \in (0.95, 1.05)$, which is sufficient for the needs of criticality tests. Otherwise a non-local stationary alternative hypothesis should be used in such cases when $m < 0.95$.

If the initial value is not negligible, the figures indicate that the linear interpolation method is valid in either the right-tailed or left-tailed test not only for local parameters close to 1, but also for non-local parameters away from 1.

5 Conclusions

This part succeeds in deriving the asymptotic distribution of the test statistics $\hat{\delta}_N$ via the joint density of BESQ process q_t and its integral $\int_0^t q_s\, ds$. Numerical calculations are conducted via Mathematica and found to be consistent with the simulation outcomes. In addition, a linear interpolation method is implemented to calculate the asymptotic power of the fixed-sample-size criticality tests. The effectiveness of the interpolation method is discussed through numerical results with respect to the influence of the initial value and the sample size. Especially, when the initial value is not negligible, the linear interpolation method is effective in either the right-tailed or left-tailed test.

Part II
Sequential Criticality Tests with Stopping Time

This part explores the sequential criticality tests (SCT) using a stopping time τ_c based on observed Fisher information of m.

In Section 6, the sequential test statistic $\hat{\delta}_{\tau_c}$ for the SCT is found to be represented as a time-changed Brownian motion B_1 with the local parameter δ as drift, therefore the SCT is actually a Z-test and the initial value has no effect on the rejection region or the power. On the other hand, the stopping time τ_c is terminated earlier as the initial value is larger.

The asymptotic distribution of τ_c is right-skewed, and tends to be terminated later under the subcritical alternative hypothesis $H_1 : \delta < 0$ than under the (super) critical hypothesis. Therefore Section 8 sets an upper bound π_c at the 99th percentile point of the distribution of stopping time τ_c under the null hypothesis. Further, one attempt in this study is to propose a combined sequential criticality test using the sequential test statistics $\hat{\delta}_{\tau_c}$ and estimator $\hat{\pi}_c$ of π_c, where the test procedure is to reject the null hypothesis immediately after the stopping time exceeds $\hat{\pi}_c$. This prevents the stopping time τ_c from being terminated late so that sample size of the SCT does not become too large.

Numerical calculations and simulations are conducted to show the asymptotic distributions of τ_c, $\hat{\tau}_c$ and $\hat{\pi}_c$.

6 Stopping times based on the observed Fisher information

For any fixed $c > 0$, define stopping times τ_c and $\hat{\tau}_c$ based on the observed Fisher information $I_n = \sum_{n=1}^{N} Z_{n-1}/\sigma^2$ of m derived at Lemma 14 in Appendix;

$$\tau_c = \inf\left\{N > 1 : \sum_{n=1}^{N} Z_{n-1}/\sigma^2 \geq c\right\}; \qquad \hat{\tau}_c = \inf\left\{N > 1 : \sum_{n=1}^{N} Z_{n-1}/s_N^2 \geq c\right\}. \quad (25)$$

where s_N^2 is the estimator of σ^2 given in (5).

Lemma 6. *s_N^2 is the consistent estimator of σ^2, i.e.*

$$S_N^2 \to^p \sigma^2 \text{ as } N \to \infty.$$

Proof. Use the notation $\varepsilon_n = Z_n - Y_n - mZ_{n-1} = \sum_{k=1}^{Z_{n-1}} (\xi_k^{(n)} - m)$ in Section 3, so that

$$\begin{aligned}
s_N^2 - \sigma^2 =& \frac{1}{N}\sum_{n=1}^{N} \frac{(Z_n - Y_n - \hat{m}_N Z_{n-1})^2}{Z_{n-1}} 1_{\{Z_{n-1}>0\}} - \sigma^2 \\
=& \frac{1}{N}\sum_{n=1}^{N} \frac{(\varepsilon_n + (m - \hat{m}_N) Z_{n-1})^2}{Z_{n-1}} 1_{\{Z_{n-1}>0\}} - \sigma^2 \\
=& \frac{1}{N}\sum_{n=1}^{N} (\frac{\varepsilon_n^2}{Z_{n-1}} - \sigma^2) 1_{\{Z_{n-1}>0\}} + 2(m - \hat{m}_N) \sum_{n=1}^{N} \sqrt{\frac{Z_{n-1}}{N}} \frac{\varepsilon_n}{\sqrt{N Z_{n-1}}} 1_{\{Z_{n-1}>0\}} \\
&+ N(m - \hat{m}_N)^2 \sum_{n=1}^{N} \frac{Z_{n-1}}{N^2} 1_{\{Z_{n-1}>0\}}
\end{aligned}$$

The convergence of the second and third term can be concluded from Lemma 2, it remains

to investigate that of the first term;

$$
\begin{aligned}
& P\left\{ \left| \frac{1}{N} \sum_{n=1}^{N} \left(\frac{\varepsilon_n^2}{Z_{n-1}} - \sigma^2 \right) 1_{\{Z_{n-1}>0\}} \right| > \epsilon \right\} \\
\leq & E\left[\left(\sum_{n=1}^{N} \left(\frac{\varepsilon_n^2}{Z_{n-1}} - \sigma^2 \right) \right)^2 / N^2 \epsilon^2 \right] \\
\leq & \sum_{n=1}^{N} E\left[\left(\frac{\varepsilon_n^2}{Z_{n-1}} - \sigma^2 \right)^2 / N^2 \epsilon^2 \right] \to 0, \ N \to \infty.
\end{aligned}
$$

The above result yields from $E[\varepsilon_n^2 | \mathcal{F}_{n-1}] = \sigma^2 Z_{n-1}$. □

In sequential analysis, it is assumed that that initial value $Z_0/\sqrt{c} \to X_0$ as $c \to \infty$. Since c is fixed in empirical analysis or simulation, X_0 is actually set as $X_0 = Z_0/\sqrt{c}$. The following Lemma 7 approximate the branching process with immigration to the Cox-Ingersoll-Ross (CIR) process with initial value X_0.

Lemma 7. *Assume that $Z_0/\sqrt{c} \to X_0$ as $c \to \infty$ with L^2 random variable X_0. For $t \in [0, \infty)$, let $X_t^{(c)} \equiv Z_{\lfloor \sqrt{c}t \rfloor}/\sqrt{c}$ and X_t to be a CIR process as follows;*

$$
X_t = X_0 + \sigma \int_0^t \sqrt{X_s} dW_s + \delta \int_0^t X_s ds + \lambda t \tag{26}
$$

Then, as $c \to \infty$, on $D[0, \infty)$,

$$
\begin{aligned}
& \left(X_t^{(c)}, \ \frac{\sum_{n=1}^{\lfloor \sqrt{c}t \rfloor} (Z_n - Y_n - Z_{n-1})}{\sqrt{c}}, \ \frac{\sum_{n=1}^{\lfloor \sqrt{c}t \rfloor} Z_{n-1}}{c} \right) \\
\Rightarrow & \left(X_t, \ \sigma \int_0^t \sqrt{X_s} dW_s + \delta \int_0^t X_s ds, \ \int_0^t X_s ds \right)
\end{aligned}
$$

where "$\Rightarrow$" stands for weak convergence, W is a standard Brownian motion.

Proof. This is an extension of Lemma 2 in I, and can be proved in the same way. □

Lemma 6 and Lemma 7 implies that τ_c and $\hat{\tau}_c$ have the following asymptotic distribution as $c \to \infty$, denoted as U_1;

$$
\tau_c/\sqrt{c} \Rightarrow U_1 \equiv \left\{ t : \int_0^t \frac{X_s}{\sigma^2} \, ds = 1 \right\}. \tag{27}
$$

For $t \in [0, \infty)$, let $q_t \equiv 4X_t/\sigma^2$ so that q_t is the squared Bessel process (BESQ) with drift as in (9). Therefore, U_1 can be represented as

$$
U_1 = \left\{ t : \int_0^t \frac{q_s}{4} \, ds = 1 \right\}.
$$

The distribution of $\int_0^t q_s \, ds$ with respect to the ordinary BESQ_x^d is studied in Pitman&Yor(1982) [15] and is represented as the following ec function;

$$
\begin{aligned}
f_{\int_0^t q_s ds}(v) &= \mathscr{L}_\gamma^{-1} \left(\cosh^{-d/2} \sqrt{2\gamma} t \exp(-\frac{x\sqrt{2\gamma} \tanh \sqrt{2\gamma} t}{2}) \right) \\
&= \mathrm{ec}_v(0, \nu + 1, t, 0, x/2).
\end{aligned}
$$

where

$$\mathrm{ec}_y(\mu,\nu,t,x,z) = \mathscr{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\cosh^\nu t\sqrt{2\gamma}}\exp(-x\sqrt{2\gamma} - z\sqrt{2\gamma}\tanh t\sqrt{2\gamma})\right).$$

For details of the ec function, see Definition 19 in Appendix and Borodin & Selminen (2002) [6].

Therefore, the distribution of U_1 under the null hypothesis $H_0 : \delta = 0$ is derived.

Theorem 8. *The CDF of U_1 under the null hypothesis is*

$$P\{U_1 \geq t\} = P\{\int_0^t X_s/\sigma^2\, ds \leq 1\} = P\{\int_0^t q_s\, ds \leq 4\}. \tag{28}$$

6.1 asymptotic distribution of $\tau_c/\sqrt{c}$

Figure 12 are the numerical results and simulations outcomes of the distributions of U_1 computed from (28), with $\delta = 0$, $\sigma^2 = 5$, $d = 6$, $X_0 = 0, 1$ respectively. Simulations are conducted with replication times 10000 and $c = 500^2$. Figure 13 show that τ_c is terminated earlier as the initial value is larger.

Figure 12: Distribution of U_1 under $H_0 : \delta = 0$ with $\sigma^2 = 5$, $d = 6\, c = 500^2$

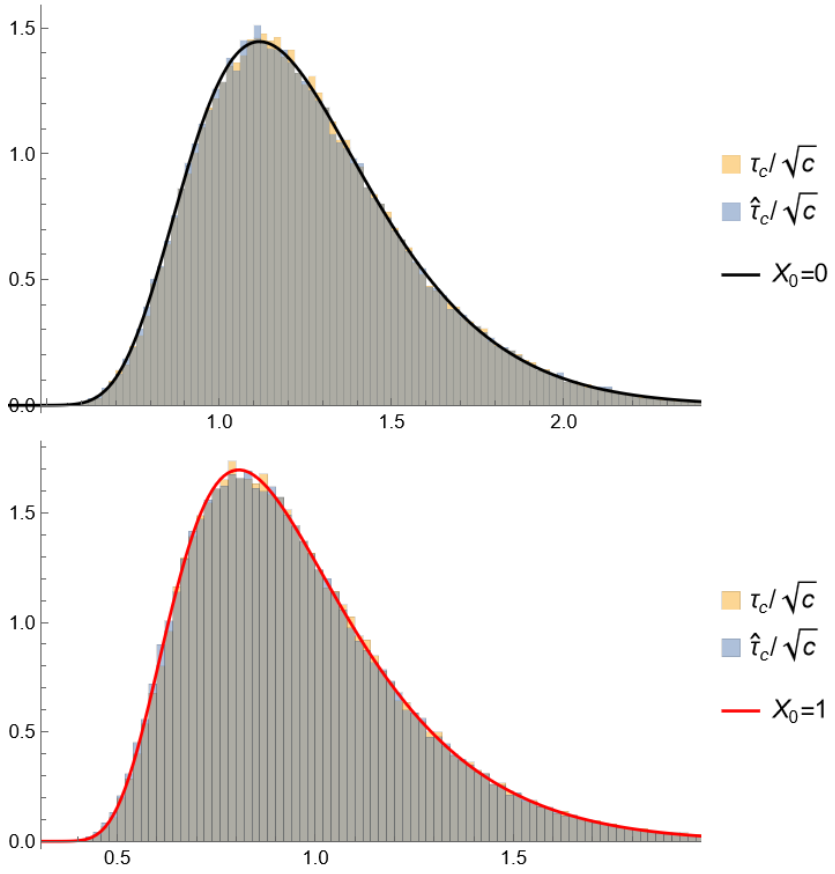
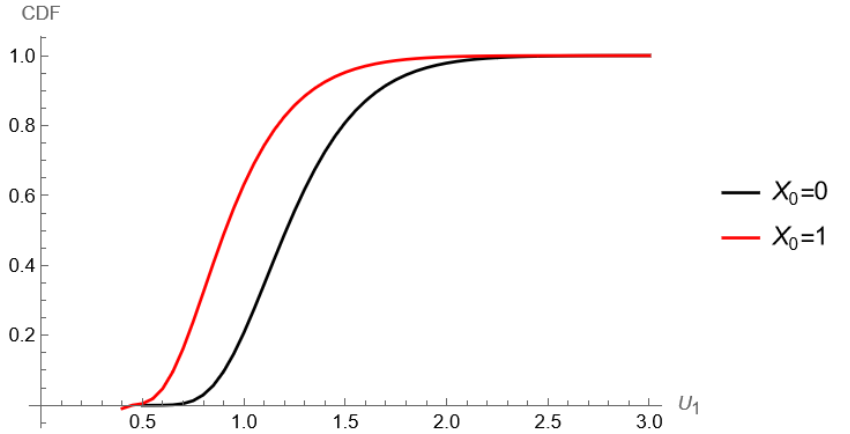


Figure 13: CDF of U_1 under $H_0 : \delta = 0$ with $\sigma^2 = 5$, $d = 6\,c = 500^2$



7 Joint asymptotic distribution of sequential test statistic and stopping time

Define a martingale M_t together with its quadratic variation as follows;

$$M_t \equiv \int_0^t \sqrt{\frac{X_s}{\sigma^2}}\, dW_s, \qquad \langle M \rangle_t = \int_0^t \frac{X_s}{\sigma^2}\, ds. \tag{29}$$

Extend the stopping time U_1 to stopping times U_v for $v \geq 0$;

$$U_v \equiv \inf \left\{ t \geq 0 : \langle M \rangle_t = v \right\}.$$

According to Theorem 23: Dambis-Dubins-Schwarz's Time-change Theorem, $B_v \equiv M_{U_v}$ is a time-changed Brownian motion and $\langle M \rangle_{U_v} = \langle B \rangle_v = v$.

Lemma 9. *Letting $\rho_0 = X_0/\sigma^2$, $\rho_v = X_{U_v}/\sigma^2$ is a Bessel (BES) process with drift δ and dimension $2\lambda/\sigma^2 + 1 (= d/2 + 1)$;*

$$\rho_v = \rho_0 + B_v + \delta v + \frac{d}{4} \int_0^v \frac{1}{\rho_s}\, ds.$$

Furthermore, U_v has the following representation;

$$U_v = \int_0^v \frac{1}{\rho_s}\, ds. \tag{30}$$

Proof. The definition of U_v implies $v = \int_0^{U_v} X_s/\sigma^2\, ds$, and $dv/du = X_u/\sigma^2$. Then the inverse function theorem, $du/dv = 1/(X_{U_v}/\sigma^2)$ implies (30). Dividing both sides of (26) by σ^2 and substituting t with U_v yield the result;

$$\begin{aligned}
\rho_v = X_{U_v}/\sigma^2 &= X_0/\sigma^2 + \int_0^{U_v} \sqrt{\frac{X_s}{\sigma^2}}\, dW_s + \delta \int_0^{U_v} \frac{X_s}{\sigma^2}\, ds + \frac{\lambda}{\sigma^2} U_v \\
&= \rho_0 + M_{U_v} + \delta \langle M \rangle_{U_v} + \frac{d}{4} U_v \\
&= \rho_0 + B_v + \delta v + \frac{d}{4} U_v \\
&= \rho_0 + B_v + \delta v + \frac{d}{4} \int_0^v \frac{1}{\rho_s}\, ds.
\end{aligned} \tag{31}$$

□

Theorem 10. *The sequential test statistics $\hat{\delta}_{\tau_c}$ and stopping time $\tau_c/\sqrt{c}$ have the following joint limit as $c \to \infty$.*

$$(\hat{\delta}_{\tau_c}, \tau_c/\sqrt{c}) \Rightarrow (B_1 + \delta, U_1). \tag{32}$$

Proof. The limit of $\tau_c/\sqrt{c}$ in (27) and Lemma 7 imply that

$$\begin{aligned}
\hat{\delta}_{\tau_c} &\Rightarrow \frac{\sigma \int_0^{U_1} \sqrt{X_s}\, dW_s}{\int_0^{U_1} X_s\, ds} + \delta \\
&= \frac{\int_0^{U_1} \sqrt{X_s/\sigma^2}\, dW_s}{\int_0^{U_1} X_s/\sigma^2\, ds} + \delta \\
&= \frac{M_{U_1}}{\langle M \rangle_{U_1}} + \delta \\
&= B_1 + \delta.
\end{aligned} \tag{33}$$

□

Therefore the test procedure of left sided SCT with rejection rate 5% is to reject H_0 when $\hat{\delta}_{\tau_c} < -1.645$, and that of the right sided SCT is to reject H_0 when $\hat{\delta}_{\tau_c} > 1.645$.

7.1 Joint density of $(B_1 + \delta, U_1)$

Lemma 11. *Joint density of (ρ_v, U_v) under the null hypothesis is derived from the following relationship with that of $(q_t, \int_0^t q_s ds)$;*

$$f_{\rho_v, U_v}(y, t) = 16y\, f_{q_t, \int_0^t q_s ds}(4y, 4v). \tag{34}$$

Proof. Let $u = U_v$ so that $v = \int_0^u X_s/\sigma^2\, ds = \int_0^u q_s/4\, ds$, and $dv = q_u/4du$. Then the Laplace transform of the joint CDF is

$$\begin{aligned}
\int_0^\infty e^{-\gamma v} P^0 \{\rho_v \le y, U_v \le t\}\, dv &= E^0 \left[\int_0^\infty e^{-\gamma v} 1_{\{\rho_v \le y, U_v \le t\}}\, dv \right] \\
&= E^0 \left[\int_0^\infty e^{-\gamma \int_0^u q_s/4\, ds} 1_{\{q_u/4 \le y, u \le t\}}\, q_u/4\, du \right] \\
&= \int_0^t E \left[e^{-\gamma/4 \int_0^u q_s ds} 1_{\{q_u \le 4y\}} q_u/4 \right] du.
\end{aligned}$$

Take the derivative with respect to t by both sides;

$$\begin{aligned}
\int_0^\infty e^{-\gamma v} \frac{\partial}{\partial t} P^0 \{\rho_v \le y, U_v \le t\}\, dv &= E^0 \left[e^{-\gamma/4 \int_0^t q_s ds} 1_{\{q_t \le 4y\}} q_t/4 \right] \\
&= \int_0^{4y} \int_0^v e^{-\gamma v/4} f_{q_t, \int_0^t q_s ds}(4y, v) y\, dv dy
\end{aligned}$$

and take the the derivative with respect to y by both sides;

$$\int_0^\infty e^{-\gamma v} \frac{\partial}{\partial t \partial y} P^0 \{\rho_v \le y, U_v \le t\}\, dv = \int_0^\infty e^{-\gamma v/4} f_{q_t, \int_0^t q_s ds}(4y, v) 4y\, dv$$

The result follows by inverting this Laplace transform with respect to v. □

Since ρ_v has the expression of $\rho_v = \rho_0 + B_v + \frac{d}{4}U_v$ under the null hypothesis, $f_{B_v,U_v}(z,t)$ is derived as

$$f_{B_v,U_v}(z,t) = f_{\rho_v,U_v}(\rho_0 + z + \frac{d}{4}t, t) \tag{35}$$

where z satisfies $z \geq -\rho_0 - \frac{d}{4}t$.

Theorem 12. *Like that in the non-sequential theory, the drift δ can be removed via the Girsanov transform, so that the joint density of (ρ_v, U_v) under alternative hypothesis is*

$$f^{\delta}_{B_1+\delta,U_1}(z,t) = \exp\left(\delta z - \frac{\delta^2}{2}v\right) f_{B_1,U_1}(z,t). \tag{36}$$

Proof. Under Alternative hypothesis, ρ_v has the expression (31): $\rho_v = X_0/\sigma^2 + \tilde{B}_v + \delta v + \frac{d}{4}U_v$. Suppose B_v is Brownian motion under P^0, define a Girsanov transformation $d\tilde{B}_v = dB_v - \delta dv$ so that $\tilde{B}_v$ is also Brownian motion under P^{δ}. Then ρ_v has the expression of $\rho_v = X_0/\sigma^2 + B_v + \frac{d}{4}U_v$ under the null hypothesis and the Radon–Nikodym derivative is represented as

$$\begin{aligned} L_v \equiv \frac{dP^{\delta}}{dP^0}|_{\mathcal{F}_v} &= \exp\left(\int_0^v \delta\, dB_s - \frac{\delta^2}{2}\int_0^v d\langle B\rangle_s\right) \\ &= \exp\left(\delta z - \frac{\delta^2}{2}v\right) \end{aligned}$$

□

Therefore, the marginal distribution of U_1 is obtain from $f_{B_1,U_1}(z,t)$;

$$f_{U_1}(t) = \int_{-\infty}^{\infty} f_{B_1,U_1}(z,t)\, dz.$$

8 Stopping time with upper bound

There is a possibility that the estimator s_N^2 underestimates σ^2 when N is small, so that the sampling procedure prematurely stops before sufficient data is collected. To prevent such cases, a minimum sample size N_0 should be set. On the other hand, τ_c or $\hat{\tau}_c$ in (25)tend to be terminated later when the subcritical alternative hypothesis $H_1 : \delta < 0$ is true. This study proposes the following combined sequential tests with the pair $(\tau_c, \hat{\delta}_{\tau_c})$ to prevent τ_c from being terminated too late. That is, if the dimension $d = 4\lambda/\sigma^2$ and the initial value $x = 4X_0/\sigma^2$ of the BESQ process and are known, the upper bound π_c is determined by

$$\pi_c = \left\lfloor \sqrt{c}u^*(d,x) \right\rfloor \text{ with } P^0\left\{U_1 < u^*(d,x)\right\} = 0.99. \tag{37}$$

Then, the left tailed test ϕ rejects H_0 when τ_c exceeds π_c and, otherwise, when $\hat{\delta}_{\tau_c}$ is below z_*. The right tailed test ψ rejects H_0 when $\hat{\delta}_{\tau_c}$ exceeds z^*, i.e.

$$\phi = \begin{cases} 1 & \text{if } \tau_c \geq \pi_c \text{ or } \left(\tau_c < \pi_c,\, \hat{\delta}_{\tau_c} < z_*\right) \\ 0 & \text{otherwise} \end{cases}$$

$$\psi = \begin{cases} 1 & \text{if } \hat{\delta}_{\tau_c} > z^* \\ 0 & \text{otherwise.} \end{cases}$$

The critical values $u^* = u^*(d,x)$, $z_* = z_*(d,x)$ and $z^* = z^*(d,x)$ are determined in the following way via the joint density of (B_1, U_1) in (35).

$$\begin{aligned} \phi :& P^0\{U_1 \geq u^*(d,x)\} = 0.01, \\ & P^0\{U_1 < u^*(d,x),\, B_1 < z_*(d,x)\} = \alpha - 0.01; \\ \psi \ :& P^0\{B_1 > z^*(d,x)\} = \alpha. \end{aligned} \tag{38}$$

(d,x) is determined on (λ, σ^2) by the relationship $x = 4X_0/\sigma^2$, $d = 4\lambda/\sigma^2$, where the initial value X_0 of CIR process X_t is set as $X_0 = Z_0/\sqrt{c}$ and considered to be a known constant. Thus u^* is actually dependent on the pair (λ, σ^2). The difficulty in empirical analysis is that λ and σ^2 are usually unknown. This study provide a linear interpolation solution via using the estimator $\hat{d}_N = 4\hat{\lambda}_N/s_N^2$ and $\hat{\sigma}_N^2 = s_N^2$ given in (5) in the following way. First, u^*, z_* and z^* in (38) for sufficiently many pairs of $(d, 4X_0/\sigma^2)$ are calculated with a fixed X_0 in advance. Second, u^*, z_* and z^* corresponding to $(\hat{d}_{\hat{\tau}_c}, 4X_0/s_{\hat{\tau}_c}^2)$ is determined by linear interpolation among those values calculated in the first step.

Define upper bound $\hat{\pi}_c$ for $\hat{\tau}_c$ using $u^*(d,x)$,

$$\hat{\pi}_c = \inf\left\{N > N_0 : N \geq \sqrt{c}u^*(\hat{d}_{\hat{\tau}_c}, 4X_0/s_{\hat{\tau}_c}^2)\right\}.$$

The left tailed test $\hat{\phi}$ rejects H_0 when τ_c exceeds $\hat{\pi}_c$ and, otherwise, when $\hat{\delta}_{\tau_c}$ is below z_*. The right tailed test ψ rejects H_0 when $\hat{\delta}_{\tau_c}$ exceeds z^*, i.e.

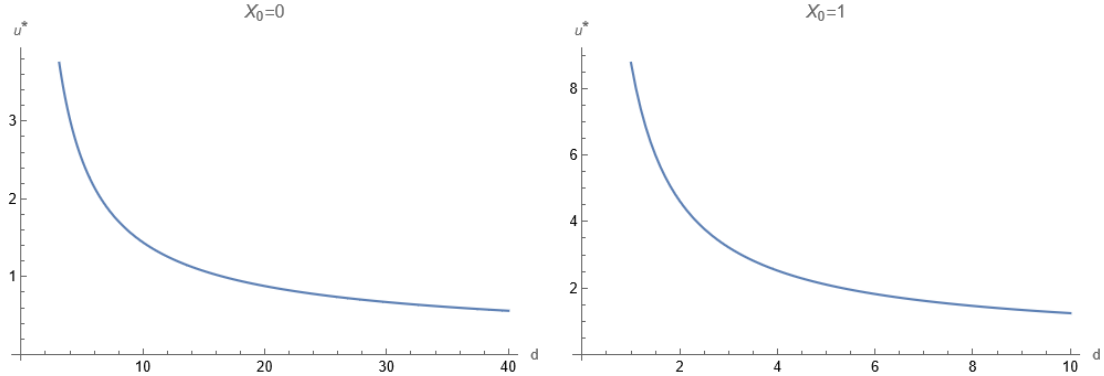
$$\hat{\phi} = \begin{cases} 1 & \text{if } \hat{\tau}_c \geq \hat{\pi}_c \text{ or } \left(\hat{\tau}_c < \hat{\pi}_c,\, \hat{\delta}_{\tau_c} < \hat{z}_*\right) \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{\psi} = \begin{cases} 1 & \text{if } \hat{\delta}_{\tau_c} > \hat{z}^* \\ 0 & \text{otherwise} \end{cases}$$

where critical values are $\hat{z}_* = \hat{z}_*(\hat{d}_{\hat{\tau}_c}, 4X_0/s_{\hat{\tau}_c}^2)$ and $\hat{z}^* = \hat{z}^*(\hat{d}_{\hat{\tau}_c}, 4X_0/s_{\hat{\tau}_c}^2)$. The powers and the expected sample sizes of these tests can be computed from (36).

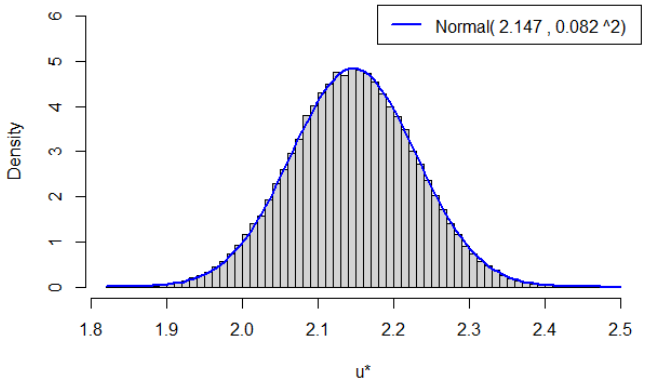
For simplicity, the work below considers $x = 4X_0/\sigma^2$ to be a known constant. Figure 14 illustrate the value of u^* calculated with corresponding to a sequence of d starting from 0.1 to 10 in steps of 0.1, with $\sigma^2 = 2$, and $X_0 = 0, 1$ $(x = 0, 2)$ respectively.

Figure 14: 99th percentile point of U_1 under the null hypothesis, $\sigma^2 = 2$



Since u^* is monotonic with respect to d, the estimator $\hat{d}_N$ is used to determine the corresponding u^*. Figure 15 is the simulation result of u^* by the linear interpolation. It indicates that $\hat{\pi}_c$ is normally distributed as $c \to \infty$.

Figure 15: asymptotic distribution of u^*, $X_0 = 0$, $d = 6$, $c = 500^2$



9 Conclusion

This part considers the stopping time τ_c based on observed Fisher information of m and succeeds in deriving the asymptotic distribution of $\tau_c/\sqrt{c}$ via the density of the integral $\int_0^t q_s\, ds$ of BESQ process q_t. A time change method is implemented to investigate the asymptotic distribution of the sequential test statistic $\hat{\delta}_{\tau_c}$ for the sequential criticality test. It is found to be a time-changed Brownian motion B_1 with the local parameter δ as drift, therefore the SCT is actually a Z-test. The joint density $f_{B_1+\delta,U_1}$ is obtained from $f_{q_t,\int_0^t q_s ds}$.

In addition, an upper bound u^* is set at the 99th percentile point of the distribution U_1 under the null hypothesis. An attempt in this part is to propose a combined sequential test with upper bound $\hat{\pi}_c$ of $\hat{\tau}_c$ to prevent $\hat{\tau}_c$ from being terminated too late so that sample size of the SCT does not become too large. The combined test also rejects the null hypothesis when the stopping time $\hat{\tau}_c$ exceeds the upper bound $\hat{\pi}_c$. The operating characteristics of the combined tests can be computed from the joint density $f_{B_1+\delta,U_1}$.

Part III
Effectiveness of Local Model

This study localize the parameter by $m = 1 + \frac{\delta}{N}$ in I and $m = 1 + \frac{\delta}{\sqrt{c}}$ in II to carry out the FCT and the SCT. In order to clarify the effectiveness of the local hypotheses in the above two tests, it is checked whether the simulation results conform to the theoretical values computed from the local model or from the stationary model.

For the FCT, the powers under the non-local stationary alternatives $H_1 : m < 1$ are computed as a Z-test since the sequential test statistic is normally distributed in the case that the initial value can be neglected. Comparisons of the powers from non-local stationary test and the powers from local FCT Section 4 are made.

For the SCT, the test statistics $\hat{m}_{\tau_c}$ are normally distributed under both the non-local stationary alternatives and the local alternatives, thus comparisons of the joint moments of the stopping time τ_c and sequential test statistic $\hat{\delta}_{\tau_c}$ are made. In detail, the theoretical value of $E[\tau_c]$, $SE[\tau_c]$ and $cov[\tau_c, \hat{\delta}_{\tau_c}]$ under local alternatives are computed for a sequence of m starting at 0.8 to 1.2 in steps of 0.1 in the following way. First, the modified Laplace transform $E\left[\exp(-\alpha\rho_v - \beta U_v)/\rho_v\right]$ under the null hypothesis is obtained. Girsanov transformation is used to obtain the joint Laplace transform under alternative hypothesis.

The simulation results conform to the theoretical values computed from the local model only when m is close to 1, otherwise they conform to those computed from the stationary model. This study helps to decide when to adapt local models and when to adapt non-local stationary models. Those decisions differ depending on the value of c.

10 Effectiveness of Local Model in Non-Sequential Analysis

For a non-local stationary branching model with initial values $Z_0 = 0$, or assuming $Z_0/N \to 0$ as $N \to \infty$, the test statistic $\hat{m}_N$ are normally distributed; see Sriram, Basawa & Huggins (1991) [18] for details.

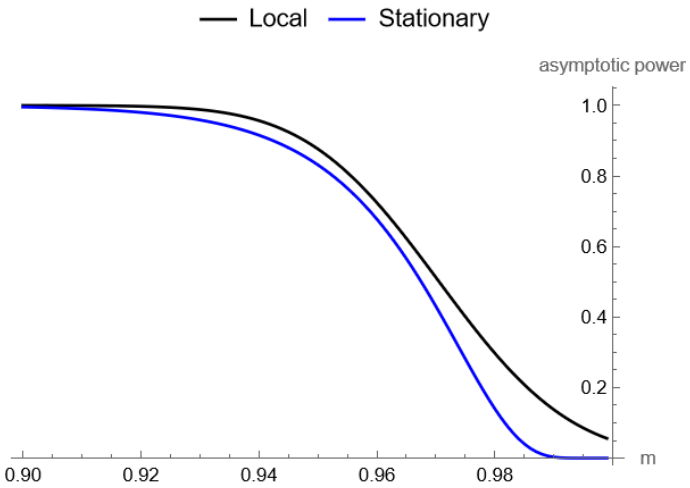
$$\sqrt{N}(\hat{m}_N - m) \Rightarrow N(0, (1-m)\sigma^2/\lambda)$$

Therefore the powers in such cases can be computed as a Z-test by the following;

$$\begin{aligned} & P^m\left\{\hat{m}_N < \frac{\delta_*}{N} + 1\right\} \\ =& \Phi\left(\frac{\delta_*}{N} + 1;\, m, \frac{(1-m)\sigma^2}{\lambda N}\right) \end{aligned}$$

where P^m is the probability measure under $H_1 : m < 1$, δ_* is the critical value of the reject region for left tailed test, computed from (24) in I Section 4, $\Phi(x;\, \mu, \sigma^2)$ is the CDF of Normal distribution $N(\mu, \sigma^2)$.

Figure 16: Comparison of theoretical powers of non-local stationary and local alternatives when $Z_0 = 0$, $d = 6$



It is clear that the power of FCT calculated under the local alternatives is greater then that under the non-local stationary alternatives for all m, but the difference between these two decreases as m drift away from 1.

11 Effectiveness of Local model in Sequential Analysis

The test statistics $\hat{m}_{\tau_c}$ are normally distributed under both the non-local stationary alternatives and the local alternatives, thus comparisons of the joint moments of the stopping time τ_c and sequential test statistic $\hat{\delta}_{\tau_c}$ are made. $E[\tau_c]$, $SE[\tau_c]$ and $cov[\tau_c, \hat{\delta}_{\tau_c}]$ under local alternatives are computed in the following way. First, to obtain the modified

joint Laplace transform $E\left[\exp(-\alpha\rho_v - \beta U_v)/\rho_v\right]$, from which the joint moment $E[U_1], E[U_1^2], E[\rho_1 U_1]$ are obtained. Second, Girsanov transformation is used to obtain joint Laplace transform under alternative hypothesis.

11.1 Joint moment of (ρ_v, U_v)

Theorem 13. *Modified Laplace transform under the null hypothesis has the expression of*

$$E^0\left[\frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\right] = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}\frac{x^n\alpha^j\beta^l}{n!j!l!}\int_0^1 J(s,v,n,j,l)\,ds \tag{39}$$

where $x^{(m)}$ stands for the factorial power $x^{(m)} = x(x-1)...(x-(m-1))$, ${}_2F_1$ stands for the hyper-geometric function ${}_2F_1(a,b,c;z) = \sum_{k=0}^{\infty}\frac{a^{(k)}b^{(k)}}{c^{(k)}}\frac{z^k}{k!}$ in the following explicit expression of J.

$$\begin{aligned}J(s,v,n,j,l) = &\frac{(-1)^n s^{\frac{\nu-3}{4}}\left(1-\sqrt{s}\right)^{j+n}\log^l(s)(-n-\nu-1)^{(j)}\left(\sqrt{s}+1\right)^{-j-\nu-n-1}}{\Gamma\left(\frac{1}{2}(j+l-n+1)\right)}\\ &\times 2^{\frac{1}{2}(-j-3l-3n-1)+\nu}v^{\frac{1}{2}(j+l-n-1)}\,{}_2F_1\left(-j,-n;-j-n-\nu;\frac{\left(\sqrt{s}+1\right)^2}{\left(\sqrt{s}-1\right)^2}\right).\end{aligned}$$

Proof. Recall that under the null hypothesis, $f_{q_t,\int_0^t q_s ds}(y,v)$ has the expression (21) that

$$\mathscr{L}_\gamma(f_{q_t,\int_0^t q_s ds}(y,v)) = \frac{1}{2}\left(\frac{y}{x}\right)^{\nu/2}\frac{\sqrt{2\gamma}}{\sinh\sqrt{2\gamma}t}\exp\left(-\frac{x+y}{2}\sqrt{2\gamma}\coth\sqrt{2\gamma}t I_\nu\left(\frac{\sqrt{xy}\sqrt{2\gamma}}{\sinh\sqrt{2\gamma}t}\right)\right)$$

Taking the Laplace transform with respect to y gives the modified Laplace transform of $(q_t, \int_0^t q_s ds)$:

$$\begin{aligned}&E^0\left[\exp(-\alpha q_t - \gamma\int_0^t q_s ds)\right]\\ =&\int_0^\infty\int_0^\infty e^{-\alpha y}e^{-\gamma v}f_{q_t,\int_0^t q_s ds}(y,v)\,dvdy\\ =&\int_0^\infty e^{-\alpha y}\mathscr{L}_\gamma\left(f_{q_t,\int_0^t q_s ds}(y,v)\right)dy\\ =&2^{\frac{\nu+1}{2}}\gamma^{\frac{\nu+1}{2}}\exp\left(-\frac{\gamma x + \frac{\sqrt{\gamma}t}{\sqrt{2}}\alpha x\coth\frac{\sqrt{\gamma}t}{\sqrt{2}}}{2\alpha + 2\sqrt{2\gamma}\coth\frac{\sqrt{\gamma}t}{\sqrt{2}}}\right)\left(\alpha\sinh\frac{\sqrt{\gamma}t}{\sqrt{2}} + \sqrt{2\gamma}\cosh\frac{\sqrt{\gamma}t}{\sqrt{2}}\right)^{-\nu-1}\end{aligned}$$

Take the following Laplace transform of the modified joint moment generating function of (ρ_v, U_v), and change the integral variable by letting $u = U_v$, so that $v = \int_0^{U_v} X_s/\sigma^2\,ds = \int_0^u q_s/4\,ds$, and $du = 1/\rho_s\,dv$. Therefore

$$\begin{aligned}&\int_0^\infty e^{-\gamma v}E^0\left[\frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\right]dv\\ =&E^0\left[\int_0^\infty e^{-\gamma v}\exp(-\alpha\rho_v - \beta U_v)\,du\right]\\ =&E^0\left[\int_0^\infty e^{-\beta u}\exp(-\alpha q_v/4 - \gamma\int_0^u q_s/4\,ds)\,du\right]\\ =&\int_0^\infty e^{-\beta u}E^0\left[\exp(-\alpha q_v/4 - \gamma\int_0^u q_s/4\,ds)\right]du\end{aligned}$$

Since this expression can not be computed directly, Taylor expansion is implemented at the neighborhood of $x = 0$. Then replace the integral variable by $u = \exp\left(-2\sqrt{2}\sqrt{\gamma}s\right)$;

$$\begin{aligned}
&\int_0^\infty e^{-\gamma v} E^0\left[\frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\right] dv \\
=&\sum_{n=0}^{\infty} \frac{x^n}{n!} \int_0^1 (-1)^n \gamma^{\nu/2} 2^{\frac{3\nu}{2}-2n} s^{\frac{\beta}{2\sqrt{2}\sqrt{\gamma}}+\frac{\nu+1}{4}-1} \times \frac{\left(\sqrt{2}\alpha\sqrt{\gamma}\left(\sqrt{s}+1\right) - 2\gamma\left(\sqrt{s}-1\right)\right)^n}{\left(\alpha\left(1-\sqrt{s}\right) + \sqrt{2}\sqrt{\gamma}\left(\sqrt{s}+1\right)\right)^{\nu+n+1}} ds.
\end{aligned}$$

Apply Taylor expansion again for multivariate variables α, β at the neighborhood of $\alpha = 0$, $\beta = 0$;

$$\int_0^\infty e^{-\gamma v} E^0\left[\frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\right] dv = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} \frac{x^n \alpha^j \beta^l}{n!j!l!} \int_0^1 K(s, v, n, j, l)\, ds$$

where

$$\begin{aligned}
K(s, v, n, j, l) = &\frac{(-1)^n s^{\frac{\nu-3}{4}} \left(1-\sqrt{s}\right)^{j+n} \log^l(s)(-n-\nu-1)^{(j)} \left(\sqrt{s}+1\right)^{-j-\nu-n-1}}{\Gamma\left(\frac{1}{2}(j+l-n+1)\right)}. \\
&\times 2^{\frac{1}{2}(-j-3l-3n-1)+\nu} v^{\frac{1}{2}(j+l-n-1)}\, {}_2F_1\left(-j, -n; -j-n-\nu; \frac{\left(\sqrt{s}+1\right)^2}{\left(\sqrt{s}-1\right)^2}\right)
\end{aligned}$$

The explicit expression of the modified moment generating function is obtained by inverting this Laplace transform respect to γ, and $J(s, v, n, j, l)$ given in (39) is the inverse Laplace transformation of $K(s, v, n, j, l)$. □

Therefore, the moment of U_1 and ρ_1 is computed by

$$\begin{aligned}
E^0[U_1] &= \frac{\partial^2}{\partial\alpha\partial\beta} E^0\left[\frac{\exp(-\alpha\rho_1 - \beta U_1)}{\rho_1}\right]|_{\alpha=0,\beta=0} \\
&= \sum_{n=0}^{\infty} \frac{x^n}{n!} \int_0^1 J(s, 1, n, 1, 1)\, ds; \\
E^0[U_1^2] &= -\frac{\partial^3}{\partial\alpha\partial\beta^2} E^0\left[\frac{\exp(-\alpha\rho_1 - \beta U_1)}{\rho_1}\right]|_{\alpha=0,\beta=0} \\
&= \sum_{n=0}^{\infty} \frac{x^n}{n!} \int_0^1 J(s, 1, n, 1, 2)\, ds; \\
E^0[\rho_1 U_1] &= -\frac{\partial^3}{\partial\alpha^2\partial\beta} E^0\left[\frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\right]|_{\alpha=0,\beta=0} \\
&= \sum_{n=0}^{\infty} \frac{x^n}{n!} \int_0^1 J(s, v, n, 2, 1)\, ds.
\end{aligned}$$

Since $\rho_1 = \rho_0 + B_1 + \frac{d}{4}U_1$ under the null hypothesis, $E[B_1U_1]$ has the following relationship with $E[\rho_1U_1]$;

$$E[\rho_1 U_1] = \rho_0 E[U_1] + E[B_1 U_1] + \frac{d}{4} E[U_1^2].$$

Under Alternative hypothesis, ρ_v has the expression (31) that : $\rho_v = \rho_0 + \tilde{B}_v + \delta v + \frac{d}{4}U_v = \rho_0 + B_v + \frac{d}{4}U_v$. Suppose B_v is Brownian motion under P^0, define a Girsanov transformation $d\tilde{B}_v = dB_v - \delta dv$ so that $\tilde{B}_v$ is also Brownian motion under P^δ, and the

Radon–Nikodym derivative is represented as

$$L_v \equiv \frac{dP^\delta}{dP^0}|_{\mathcal{F}_v} = \exp\left(\int_0^v \delta dB_s - \frac{\delta^2}{2}\int_0^v d\langle B\rangle_s\right)$$
$$= \exp\left(\delta(\rho_v - \rho_0 - \frac{d}{4}U_v) - \frac{\delta^2}{2}v\right).$$

Therefore

$$\begin{aligned}
&E^\delta\left[\frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\right] \\
&= \int_\Omega \frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v}\, dP^\delta \\
&= \int_\Omega \frac{\exp(-\alpha\rho_v - \beta U_v)}{\rho_v} \exp(\delta(\rho_v - \rho_0 - \frac{d}{4}U_v) - \frac{\delta^2}{2}v)\, dP^0 \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2}v)\int_\Omega \frac{\exp(-(\alpha-\delta)\rho_v - (\beta + \frac{d\delta}{4})U_v}{\rho_v}\, dP^0 \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2}v)E^0\left[\frac{\exp(-(\alpha-\delta)\rho_v - (\beta + \frac{d\delta}{4})U_v}{\rho_v}\right] \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2}v)\sum_{n=0}^\infty\sum_{j=0}^\infty\sum_{l=0}^\infty \frac{x^n(\alpha-\delta)^j(\beta+\frac{d\delta}{4})^l}{n!j!l!}\int_0^1 J(s,v,n,j,l)\, ds \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2}v)\sum_{n=0}^\infty\sum_{M=0}^\infty\sum_{j=0}^M \frac{x^n(\alpha-\delta)^j(\beta+\frac{d\delta}{4})^{M-j}}{n!j!(M-j)!}\int_0^1 J(s,v,n,j,M-j)\, ds \quad (40)
\end{aligned}$$

The moment of U_1 and ρ_1 can be computed in the same way as under the null hypothesis;

$$\begin{aligned}
E^\delta[U_1] &= \frac{\partial^2}{\partial\alpha\partial\beta}E^\delta\left[\frac{\exp(-\alpha\rho_1 - \beta U_1)}{\rho_1}\right]|_{\alpha=0,\beta=0} \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2})\sum_{n=0}^\infty\sum_{M=0}^\infty\sum_{j=0}^M \frac{x^n M^\delta(-1)^j(\frac{d}{4})^{M-j}}{n!j!(M-j)!}\int_0^1 J(s,v,n,j+1,M-j+1)\, ds; \\
E^\delta[U_1^2] &= -\frac{\partial^3}{\partial\alpha\partial\beta^2}E^\delta\left[\frac{\exp(-\alpha\rho_1 - \beta U_1)}{\rho_1}\right]|_{\alpha=0,\beta=0} \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2})\sum_{n=0}^\infty\sum_{M=0}^\infty\sum_{j=0}^M \frac{x^n M^\delta(-1)^j(\frac{d}{4})^{M-j}}{n!j!(M-j)!}\int_0^1 J(s,v,n,j+1,M-j+2)\, ds; \\
E^\delta[\rho_1 U_1] &= -\frac{\partial^3}{\partial\alpha^2\partial\beta}E^\delta\left[\frac{\exp(-\alpha\rho_1 - \beta U_1)}{\rho_1}\right]|_{\alpha=0,\beta=0} \\
&= \exp(-\delta\rho_0 - \frac{\delta^2}{2})\sum_{n=0}^\infty\sum_{M=0}^\infty\sum_{j=0}^M \frac{x^n M^\delta(-1)^j(\frac{d}{4})^{M-j}}{n!j!(M-j)!}\int_0^1 J(s,v,n,j+2,M-j+1)\, ds.
\end{aligned}$$

Since $\rho_1 = \rho_0 + \delta + B_1 + \frac{d}{4}U_1$ under the null hypothesis, $E[B_1U_1]$ has the following relationship with $E[\rho_1U_1]$;

$$E[\rho_1U_1] = (\rho_0+\delta)E[U_1] + E[B_1U_1] + \frac{d}{4}E[U_1^2].$$

11.2 Comparisons of the joint moments of (B_1, U_1)

The theory of the sequential analysis for stationary branching model are given as follows;

1. $\frac{\tau_c}{c} \to \frac{(1-m)\sigma^2}{\lambda}$;
2. $\sqrt{c}(\hat{m}_{\tau_c} - m) \Rightarrow N(0,1)$;
3. $\sqrt{c}\left(\frac{\tau_c}{c} - \frac{(1-m)\sigma^2}{\lambda}\right) \Rightarrow N\left(0, \frac{\sigma^4+(1-m)\sigma^2}{\lambda^2}\right)$.

The following figures are the theoretical values and simulation results of the moments $E[\tau_c]$, $SE[\tau_c]$ and $cov[\tau_c, \hat{\delta}_{\tau_c}]$. Computations are conducted as m varies from 0.8 to 1.2, with $\sigma^2 = 1.2, d = 4$, iteration times=10000, $c = 50^2, 100^2$ and $X_0 = 0, 1$ respectively.

Figure 17: effectiveness of local parameter: Compare $E[\tau_c]$;
$\sigma^2 = 1.2, d = 4, X_0 = 0$, iteration times10000, $c = 50^2, 100^2$

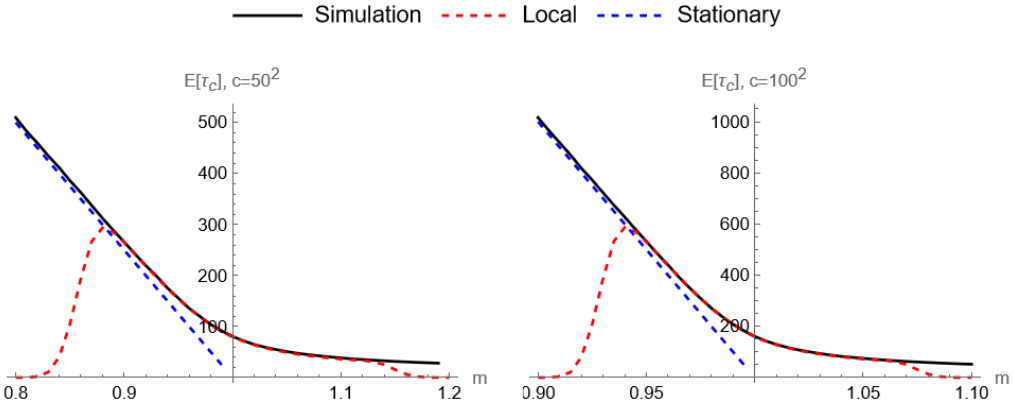


Figure 18: effectiveness of local parameter: Compare $SE[\tau_c]$;
$\sigma^2 = 1.2, d = 4, X_0 = 0$, iteration times10000, $c = 50^2, 100^2$

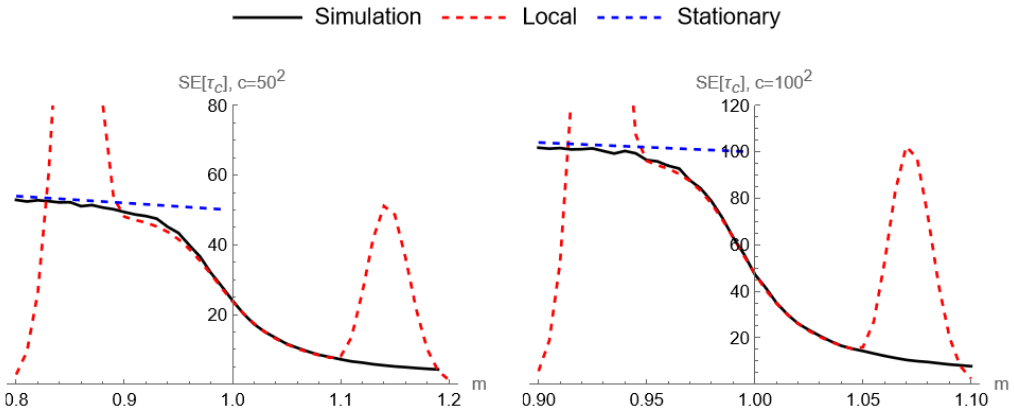


Figure 19: effectiveness of local parameter: Compare $Cov[\delta_{\tau_c}, \tau_c]$;
$\sigma^2 = 1.2, d = 4, X_0 = 0$, iteration times10000, $c = 50^2, 100^2$

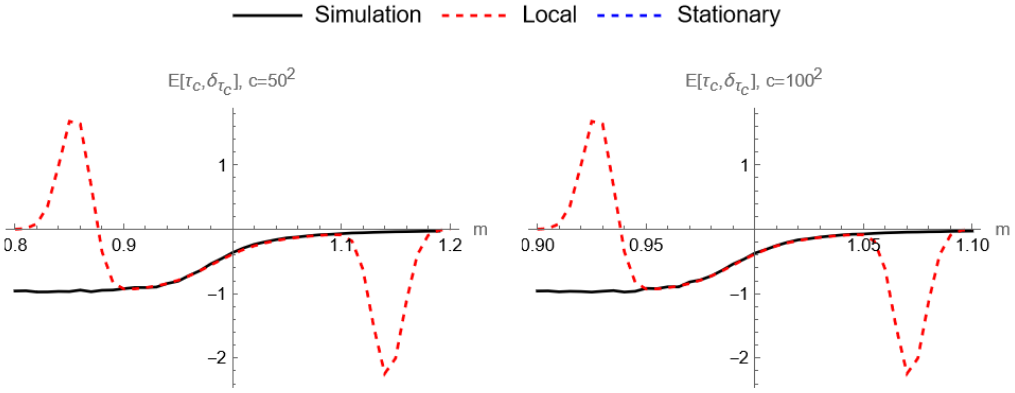


Figure 20: effectiveness of local parameter: Compare $E[\tau_c]$;
$\sigma^2 = 1.2, d = 4, X_0 = 1$, iteration times10000, $c = 50^2, 100^2$

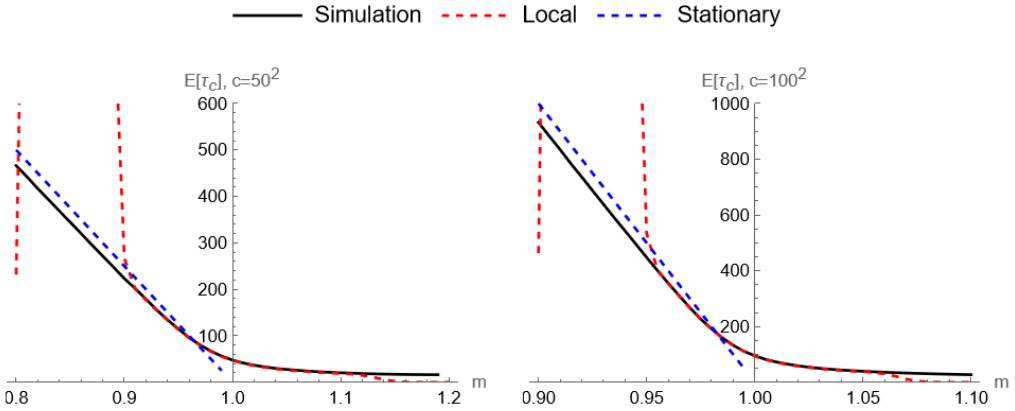


Figure 21: effectiveness of local parameter: Compare $SE[\tau_c]$;
$\sigma^2 = 1.2, d = 4, X_0 = 1$, iteration times10000, $c = 50^2, 100^2$

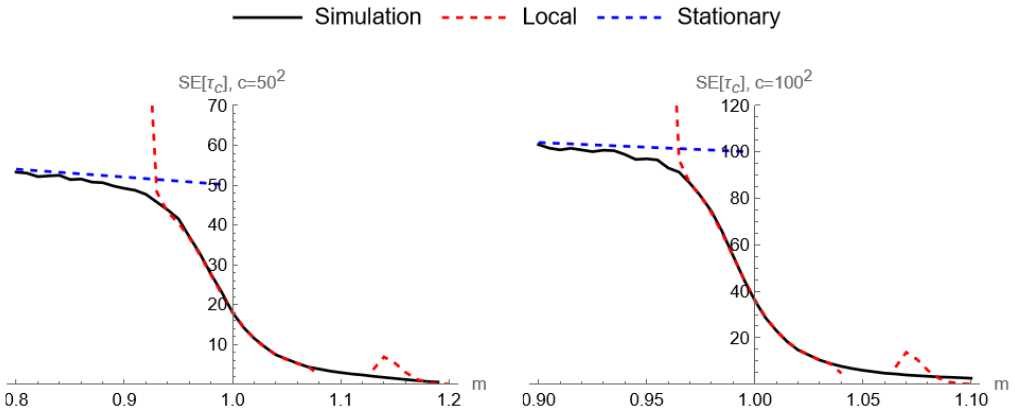
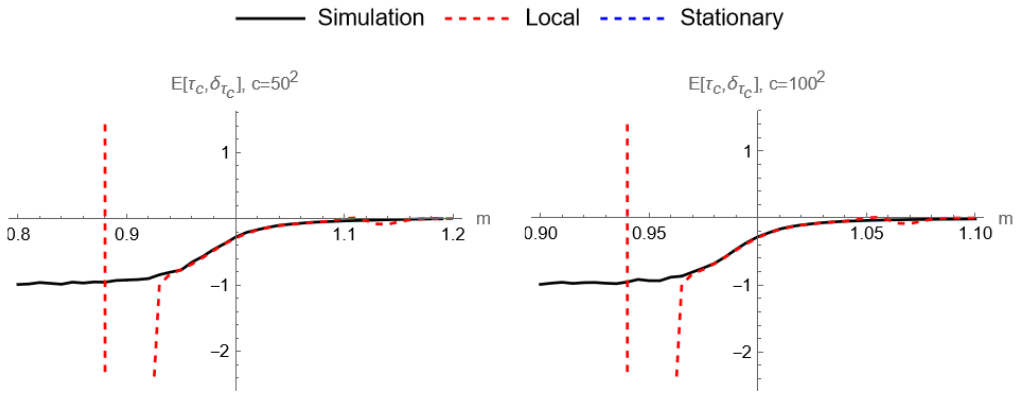


Figure 22: effectiveness of local parameter: Compare $Cov[\delta_{\tau_c}, \tau_c]$;
$\sigma^2 = 1.2, d = 4, X_0 = 1$, iteration times10000, $c = 50^2, 100^2$



As it is shown in the figures, the moments $E[\tau_c]$, $SE[\tau_c]$ and $cov[\tau_c, \hat{\delta}_{\tau_c}]$ calculated from the local model $\delta = \sqrt{c}(m-1)$ fit well with the simulation results in the neighborhood of $m = 1$. More specifically, in $m \in (0.9, 1.1)$ when $c = 50^2$, or in $m \in (0.95, 1.05)$ when $c = 100^2$. At the same time, moments calculated by the stationary model $m < 1$ fit with the simulation results in $m < 0.9$ when $c = 50^2$, or in $m < 0.95$ when $c = 100^2$. These results imply that when the level c of the observed Fisher information is large enough, the stationary model could perform well even at the neighborhood of $m = 1$. However, when the c is small, the range in which the stationary model performs well becomes narrow. In other words, the local model performs better than the stationary model at the neighborhood of $m = 1$ especially when c is small.

12 Conclusion

This part validates the effectiveness of the local model used for the FCT in I and the SCT in II. It is confirmed that the local models perform better than the stationary model at the neighborhood of $m = 1$. In detail, the power of FCT calculated under the local alternatives is greater then that under the non-local stationary alternatives for all m, but the difference between these two decreases as m drift away from 1. As for SCT, the local model performs better than the stationary model at the neighborhood of $m = 1$ especially when the level c of the observed Fisher information is small.

Part IV
Appendix & Reference

Lemma 14. *(Maximum likelihood estimation of offspring mean m with power series distribution) Suppose the number of offspring $\{\xi_{n,k}\}$ and the number of immigration $\{Y_n\}$ are non-negative, integer-valued and independent random variables.*

Let the offspring distribution of $\{\xi_{n,k}\}$ and the immigration distribution of $\{Y_n\}$ be the identical power series distribution with non-negative real sequence $\{a_i\}$ and $\{b_j\}$ respectively;

$$p(i, \theta) = P\{\xi_{n,k} = i\} = \frac{a_i \theta^i}{g(\theta)}, \ i = 0, 1, ...$$

$$q(j, \eta) = P\{Y_n = j\} = \frac{b_j \eta^j}{h(\eta)}, \ l = 0, 1, ...$$

where θ and η are real parameters; $g(\theta)=\sum_{i=0}^{\infty}a_i\theta^i$, $h(\eta)=\sum_{j=0}^{\infty}b_j\eta^j$.

The sum $\xi_{n,1}+\xi_{n,2}+...+\xi_{n,k}$ also follows the power series distribution, with non-negative sequence $\{a_l^{(k)}\}$ satisfying $\sum_{l=0}^{\infty}a_l^{(j)}\theta^l=g^j(\theta)$ as the coefficients, so that $P\{\xi_{n,1}+\xi_{n,2}+...+\xi_{n,k}=l\}=a_l^{(k)}\theta^l/g^k(\theta)$.

The following likelihood $L_n(\theta,\eta)$ of θ and η is given in [4]

$$\begin{aligned}L_n(\theta,\eta)&=\prod_{i=1}^{n}p(\xi_{i,Z_{i-1}},\theta)\cdot\prod_{j=1}^{n}q(Y_j,\eta)\\&=\prod_{i=1}^{n}\frac{a_{Z_i-Y_i}^{(Z_{i-1})}\theta^{Z_i-Y_i}}{g^{Z_{i-1}}(\theta)}\frac{b_j\eta^j}{h(\eta)}.\end{aligned}$$

The maximum likelihood estimation (M.L.E.) of offspring mean m is given to be

$$\hat{m}_n=\sum_{i=1}^{n}\frac{Z_i-Y_i}{Z_{i-1}}$$

by letting

$$\begin{aligned}\frac{\partial\log L_n(\theta,\eta)}{\partial m}&=\frac{\partial\log L_n(\theta,\eta)}{\partial\theta}\frac{\partial\theta}{\partial m}\\&=\sum_{i=1}^{n}\left((Z_i-Y_i)\frac{\partial\log\theta}{\partial m}-Z_{i-1}\frac{\partial\log g(\theta)}{\partial m}\right)\\&=\sum_{i=1}^{n}\left((Z_i-Y_i)\frac{1}{\sigma^2}-Z_{i-1}\frac{m}{\sigma^2}\right)\\&=0.\end{aligned}$$

Further more, the observed Fisher information of m is

$$\begin{aligned}I_n(m)&=-\frac{\partial^2\log L_n(\theta,\eta)}{\partial m^2}\\&=\frac{\sum_{i=1}^{n}Z_{i-1}}{\sigma^2}.\end{aligned}\tag{41}$$

Theorem 15. *(Functional Central Limit Theorem for martingale difference [2]) For a positive integer n, let $\{\varepsilon_{n1},\varepsilon_{n2},...\}$ be a martingale difference with respect to the σ-fields $\{\mathcal{F}_0^n,\mathcal{F}_1^n,...\}$. Suppose ε_{nk} is a martingale difference with respect to $\mathcal{F}_k^n$, i.e. ε_{nk} is $\mathcal{F}_k^n$-measurable and $E[\varepsilon_{nk}|\mathcal{F}_{k-1}^n]=0$. Suppose the second moments of ε_{nk} exist and let $\sigma_{nk}^2\equiv E[\varepsilon_{nk}^2|\mathcal{F}_{k-1}^n]$. Then $\sum_{k\le nt}\varepsilon_{nk}\Rightarrow W$ in the sense of $D[0,\infty)$, where "$\Rightarrow$" stands for weak convergence and W is a standard Brownian motion, if the following hold.*

1.$\sum_{k\le nt}\sigma_{nk}^2\Rightarrow_n t$ for every t,

2.$\sum_{k\le nt}E[\varepsilon_{nk}^2 1_{\{|\varepsilon_{nk}|\}>\epsilon}]\to_n 0$ for every t and ϵ.

Definition 16. (CIR process) For any $a\ge 0$, $b\in\mathbb{R}$ and $\sigma>0$, the CIR model X_t is given by the following stochastic differential equation:

$$dX_t=(a-bX_t)\,dt+\sigma\sqrt{X_t}\,dW_t\tag{42}$$

where W_t is a Brownian motion.

Definition 17. (squared Bessel process) For every $d \geq 0$ and $x \geq 0$, the unique strong solution of the equation

$$q_t = x + 2\int_0^t \sqrt{q_s} dW_s + dt \tag{43}$$

is defined as the d-dimensional squared Bessel process started at x and denoted by BESQ_x^d.

The number $v = d/2 - 1$ is the index of the corresponding process, and denote $\mathrm{BESQ}_x^{(\nu)}$ instead of BESQ_x^d when v is used instead of d.

Theorem 18. *(Itô's formula) Suppose f is of differentiability class C^2 i.e. f' and f'' exist and are continuous. For a semimartingale X_t, the following hold with probability one for all $t \geq 0$;*

$$f(X_t) - f(X_0) = \int_0^t f'(X_s) dX_s + \frac{1}{2}\int_0^t f''(X_s) d_s.$$

Definition 19. (Special inverse Laplace transform [14]) Here are some special form of inverse Laplace transformation results which are used in Theorem 3 and Lemma 8 for calculations of the joint density of $(q_1, \int_0^1 q_s ds)$ and the density of $\int_0^1 q_s ds$. For $t \geq 0, \nu > -1/2$

$$\begin{aligned}
\mathrm{is}_y(\nu, t, r, z, x) &:= \mathscr{L}_\gamma^{-1}\left\{\frac{\sqrt{2\gamma}}{\sinh t\sqrt{2\gamma}} \exp\left(-r\sqrt{2\gamma} - \frac{z\sqrt{2\gamma}\cosh t\sqrt{2\gamma}}{\sinh t\sqrt{2\gamma}}\right) I_\nu\left(\frac{2x\sqrt{2\gamma}}{\sinh t\sqrt{2\gamma}}\right)\right\} \\
&= \sum_{l=0}^{\infty} \frac{x^{\nu+2l}}{l!\Gamma(\nu+l+1)} \mathrm{es}_y(\nu+2l+1, \nu+2l+1, t, r, z) \\
&\text{for} \nu \geq -1, t + \nu t + r + z > 0, t > 0 \\
\mathrm{es}_y(\mu, \nu, t, x, z) &:= \mathscr{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\sinh^\nu t\sqrt{2\gamma}} \exp(-x\sqrt{2\gamma} - z\sqrt{2\gamma}\coth t\sqrt{2\gamma})\right) \\
&= \sum_{k=0}^{\infty} \frac{(-z)^k}{k!} \mathrm{s}_y(\mu+k, \nu+k, t, x+z+kt) \\
&\text{for} \nu \geq 0, \nu t + x + z > 0, t > 0 \\
\mathrm{s}_y(\mu, \nu, t, z) &:= \mathscr{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\sinh^\nu t\sqrt{2\gamma}} \exp(-z\sqrt{2\gamma})\right) \\
&= 2^\nu \sum_{j=0}^{\infty} \frac{\Gamma(\nu+j)\exp\left(-\frac{(\nu t+z+2jt)^2}{4y}\right)}{\sqrt{2\pi} j! y^{1+\mu/2}\Gamma(\nu)} D_{\mu+1}\left(\frac{\nu t + z + 2jt}{\sqrt{y}}\right) \\
&\text{for} \nu \geq 0, \nu t + z > 0, t > 0
\end{aligned}$$

$$\begin{aligned}
\mathrm{ec}_y(\mu,\nu,t,x,z) &= \mathscr{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\cosh^\nu t\sqrt{2\gamma}}\exp(-x\sqrt{2\gamma}-z\sqrt{2\gamma}\tanh t\sqrt{2\gamma})\right)\\
&= \sum_{k=0}^{\infty}\frac{z^k}{k!}c_y(\mu+k,\nu+k,t,x+z+kt)\\
&\text{for}\nu\geq 0,\nu t+x+z>0,t>0\\
\mathrm{c}_y(\mu,\nu,t,z) &= \mathscr{L}_\gamma^{-1}\left(\frac{(2\gamma)^{\mu/2}}{\cosh^\nu t\sqrt{2\gamma}}\exp(-z\sqrt{2\gamma})\right)\\
&= 2^\nu\sum_{j=0}^{\infty}\frac{(-1)^j\Gamma(\nu+j)\exp\left(-\frac{(\nu t+z+2jt)^2}{4v}\right)}{\sqrt{2\pi}v^{1+\mu/2}\Gamma(\nu)j!}D_{\mu+1}\left(\frac{\nu t+z+2jt}{\sqrt{v}}\right)\\
&\text{for}\nu\geq 0,\nu t+z>0,t>0.
\end{aligned}$$

Theorem 20. *(change of measure and Girsanov Theorem) Let $\{W_t\}$ be a standard Brownian motion under the probability measure P, and X_t be a measurable process adapted to the natural filtration of the Brownian motion $\{\mathcal{F}_t^W\}$; assume that the usual conditions have been satisfied . Given X_t as a continuous local martingale starting from zero , define Z_t to be the exponential semi-martingale associated with X_t, i.e.*

$$Z_t = \mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}\langle X\rangle_t\right),$$

where $\langle X\rangle_t$ denotes the quadratic variation of X.

By the stability property of stochastic integration, that Z_t is also a continuous local martingale, and a new probability measure Q can be defined on by letting Radon-Nikodym derivative be

$$\frac{dQ}{dP}|_{\mathcal{F}_t} = \mathcal{E}(X)_t.$$

Then for each t the measure Q restricted to the augmented sigma fields$\mathcal{F}_t^0$ is equivalent to P restricted to$\mathcal{F}_t^0$. Furthermore if M_t is a local martingale under measure P then the process

$$\widetilde{M}_t = M_t - \langle M, X\rangle_t$$

is a Q-local martingale on the filtered probability space $(\Omega, F, Q, \{\mathcal{F}_t^W\})$.

Corollary 21. *Girsanov ' s theorem describes the distribution of the stochastic process W_t under this new probability measure. Define*

$$\widetilde{W}_t = W_t - \int_0^t f_s\,ds,$$

then the stochastic process $\widetilde{W}_t$ is a standard Wiener process under the probability measure Q by letting

$$\frac{dQ}{dP}|_{\mathcal{F}_t} = \exp\left(\int_0^t f_s dW_s - \frac{1}{2}\int_o^t f_s^2 ds\right). \tag{44}$$

Remark. The fact that$\widetilde{W}_t$ is continuous is trivial; by Girsanov's theorem it is a Q local martingale, and by computing $\left\langle\widetilde{W}\right\rangle_t = \langle W\rangle_t = t$, it follows by Levy's characterization of Brownian Motion that this is a Q-Brownian Motion.

Lemma 22. *(Lévy's Characterization of Brownian Motion) Let B be a continuous local martingale with $B_0 = 0$. Then B is standard Brownian motion on the underlying filtered probability space is and only if B has quadratic variation $\langle B \rangle_t = t$.*

Theorem 23. *(Dambis-Dubins-Schwarz's Time-change theorem) Let $\{M_t\}$ be a continuous local martingale with respect to filtration $\mathcal{F}_t$ such that $M_0 = 0$ and $\langle M \rangle_\infty = \infty$ almost surely. For all $t \geq 0$, define a stopping time τ_s based on the quadratic variation $\langle M \rangle_t$:*

$$\tau_s = \inf \{t \geq 0, \langle M \rangle_t > s\} = \langle M \rangle_s^{-1}.$$

Then $B_s \equiv M_{\tau_s}$ is a time-changed Brownian motion with respect to filtration $\mathcal{G}_s = \mathcal{F}_{\tau_s}$, and $B_{\langle M \rangle_t} = M_t$.

References

[1] Athreya, K. B., Ney, P. E., & Ney, P. E. (2004). Branching processes. *Courier Corporation.*

[2] Billingsley, P. (1999). Convergence of probability measures 2nd. *John Wiley & Sons.*

[3] Billingsley, P. (1995). Probability and measure 3rd. *John Wiley & Sons.*

[4] Bhat, B. R., & Adke, S. R. (1981). Maximum likelihood estimation for branching processes with immigration. *Advances in Applied Probability, 13(3), 498-509.*

[5] Borodin, A. N., & Salminen, P. (2002). Handbook of Brownian motion-facts and formulae 2nd. *Springer Science & Business Media.*

[6] Borodin, A. N., & Salminen, P. (2002). Inverse Laplace Transforms. In Handbook of Brownian Motion—Facts and Formulae (pp. 649-651). *Birkhäuser, Basel.*

[7] Brockwell, P. J., & Davis, R. A. (2006). Time series: theory and methods 2nd. *Springer science & business media.*

[8] Chan, N. H. (1988). The parameter inference for nearly nonstationary time series. *Journal of the American Statistical Association, 83(403), 857-862.*

[9] Cohn, D. L. (2010). Measure theory 2nd. *New York: Birkhäuser.*

[10] Feller, W. (1951). Two singular diffusion problems. *Annals of mathematics, 173-182.*

[11] Ikeda, N., & Watanabe, S. (2014). Stochastic differential equations and diffusion processes. *Elsevier.*

[12] Karatzas, I., & Shreve, S. E. (1991). Brownian motion and stochastic calculus 2nd. *Springer Science & Business Media.*

[13] Lai, T. L., & Siegmund, D. (1983). Fixed accuracy estimation of an autoregressive parameter. *The Annals of Statistics, 478-485.*

[14] Liptser, R. S., & Shiriaev A. N. (1977). Statistics of random processes: General theory. *New York: Springer-verlag.*

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

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

[17] Revuz, D., & Yor, M. (1999). Representation of Martingales. In Continuous Martingales and Brownian Motion (pp. 179-220). *Springer, Berlin, Heidelberg.*

[18] Sriram, T. N., Basawa, I. V., & Huggins, R. M. (1991). Sequential Estimatin for Branching Processes with Immigration. *The Annals of Statistics, 2232-2243.*

[19] Tanaka, K. (2017). Time series analysis: nonstationary and noninvertible distribution theory 2nd. *John Wiley & Sons.*

[20] Vanyolos, A., Cho, M., & Glasgow, S. A. (2014). Probability density of the CIR model. *Available at SSRN 2508699.*