

多重積分ブラウン運動の統計学的応用について

田中 勝人（一橋大学名誉教授）

1. はじめに

$\{W(t)\}$ を $[0, 1]$ 上で定義されたブラウン運動 (Bm) とする. Bm は, 離散時間確率過程のランダム・ウォーク (=単純 I(1) 過程) $\{y_j^{(1)}\}$ と次の対応関係がある.

$$W(t) = \int_0^t dW(u) \Leftrightarrow y_j^{(1)} = (1-L)^{-1}\varepsilon_j = \sum_{i=1}^j \varepsilon_i, \quad \{\varepsilon_j\} \sim \text{i.i.d.}(0, 1). \quad (1)$$

ここで, L はラグ・オペレータであり, $(1-L)^{-1}$ は積和演算子である. さらに,

$$\int_0^t W(u) du = \int_0^t (t-u) dW(u) \Leftrightarrow y_j^{(2)} = (1-L)^{-2}\varepsilon_j \approx \sum_{i=1}^j (j-i)\varepsilon_i$$

$$\begin{aligned} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{g-1}} W(t_g) dt_g dt_{g-1} \cdots dt_1 &= \int_0^t \frac{(t-u)^g}{g!} dW(u) \\ &\Leftrightarrow y_j^{(g+1)} = (1-L)^{-(g+1)}\varepsilon_j \approx \frac{1}{g!} \sum_{i=1}^j (j-i)^g \varepsilon_i. \end{aligned}$$

定義 : g 重積分 Bm

$$F_g(t) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{g-1}} W(t_g) dt_g dt_{g-1} \cdots dt_1 \quad (2)$$

$$= \int_0^t F_{g-1}(u) du, \quad F_0(t) = W(t) \quad (3)$$

$$\Leftrightarrow y_j^{(g+1)} = (1-L)^{-(g+1)}\varepsilon_j \approx \frac{1}{g!} \sum_{i=1}^j (j-i)^g \varepsilon_i \quad (4)$$

を g 重積分 Bm という. $\{y_j^{(g+1)}\}$ は単純 I($g+1$) 過程である.

定理 1.

$$F_g(t) = \int_0^t \frac{(t-u)^g}{g!} dW(u) \sim N\left(0, \frac{t^{2g+1}}{(2g+1)(g!)^2}\right), \quad (5)$$

$$\int_0^1 F_g^2(t) dt = \int_0^1 \int_0^1 K_g(s, t) dW(s) dW(t) \quad (6)$$

$$\stackrel{\mathcal{D}}{=} \int_0^1 \int_0^1 L_g(s, t) dW(s) dW(t), \quad (7)$$

$$K_g(s, t) = \int_{\max(s, t)}^1 \frac{((u-s)(u-t))^g}{(g!)^2} du,$$

$$L_g(s, t) = \text{Cov}(F_g(s), F_g(t)) = \int_0^{\min(s, t)} \frac{((s-u)(t-u))^g}{(g!)^2} du.$$

ここで, (7) の関係式は Karhunen-Loève 展開 (後述) による.

定理 2. [Chan-Wei (1988), Tanaka (2017)] $C[0, 1]$ 上の部分和過程

$$Y_n^{(g)}(t) = \frac{1}{n^{g-1/2}} y_{[nt]}^{(g)} + (nt - [nt]) \frac{1}{n^{g-1/2}} y_{[nt]+1}^{(g-1)}$$

に対して、次の FCLT が成り立つ。

$$\{Y_n^{(g)}(t)\} \Rightarrow \{F_{g-1}(t)\}.$$

このことと CMT により、次のことが成り立つ。

$$\frac{1}{n^{2g}} \sum_{j=1}^n (y_j^{(g)})^2 = \frac{1}{n} \sum_{j=1}^n \left(Y_n^{(g)} \left(\frac{j}{n} \right) \right)^2 \Rightarrow \int_0^1 F_{g-1}^2(t) dt.$$

2. 多重単位根分布

次のモデルを考える。

$$y_j = \rho y_{j-1} + v_j, \quad (1-L)^g v_j = \varepsilon_j, \quad \{\varepsilon_j\} \sim \text{i.i.d.}(0, \sigma_\varepsilon^2). \quad (8)$$

このモデルは誤差項 v_j が単純 $I(g)$ 過程に従う AR(1) モデルである。以下では、係数 ρ の真値を 1 とする。したがって、 $\{y_j\}$ は単純 $I(g+1)$ 過程となる。このとき、 ρ の LSE $\hat{\rho}$ は次のようになる。

$$\hat{\rho} = \frac{\sum_{j=2}^n y_{j-1} y_j}{\sum_{j=2}^n y_{j-1}^2}, \quad \hat{\rho} - 1 = \left(\frac{1}{2} (y_n^2 - y_1^2) - \frac{1}{2} \sum_{j=2}^n v_j^2 \right) / \sum_{j=2}^n y_{j-1}^2. \quad (9)$$

ここで、

$$y_n^2 = O_p(n^{1+2g}), \quad \sum_{j=2}^n v_j^2 = O_p(n^{2g}), \quad \sum_{j=2}^n y_{j-1}^2 = O_p(n^{2+2g}),$$

であるから、FCLT と CMT により、

$$n(\hat{\rho} - 1) \Rightarrow R_g = \frac{1}{2} F_g^2(1) / \int_0^1 F_g^2(t) dt = U_g / V_g \quad (g = 1, 2, \dots) \quad (10)$$

を得る。この結果は $g = 0$ の通常の単位根分布には適用されない。なお、誤差項 v_j が一般の $I(g)$ 過程、すなわち、 $\{\varepsilon_j\}$ が独立でない短期記憶定常過程の場合にも (10) は成り立つ。

3. フレッドホルム行列式の導入

多重単位根統計量 $R_g = U_g / V_g$ の分布を考える前に、 R_g の分母 V_g の分布について考察する。統計量 V_g は、

$$V_g = \int_0^1 F_g^2(t) dt = \int_0^1 \int_0^1 K_g(s, t) dW(s) dW(t)$$

と表される。密度を直接求めることは困難なので、 V_g の特性関数を導出する。そのための準備として、まず、フレッドホルム積分方程式とフレッドホルム行列式 (FD) について述べる。

同次第二種フレッドホルム積分方程式: $[0, 1] \times [0, 1]$ 上で定義された対称, 連続な核関数 $K(s, t)$ に対して, 方程式

$$f(t) = \lambda \int_0^1 K(s, t) f(s) ds$$

を同次第二種フレッドホルム積分方程式という. ここで, λ は固有値, $f(t)$ は λ に対応する自明でない固有関数である. これは, 行列の場合の固有値 $\mathbf{f} = \lambda K \mathbf{f}$ の連続バージョンである. ただし, 固有値は逆数となっている. 同一の固有値に対して線形独立な固有関数の最大数をその固有値の重複度という.

フレッドホルム行列式 (FD): $D(\lambda)$ を核関数 $K(s, t)$ の FD とするとき,

$$D(\lambda) = \lim_{N \rightarrow \infty} \left| I_N - \frac{\lambda}{N} K_N \right| \quad (11)$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n \lambda^n}{n!} \int_0^1 \cdots \int_0^1 K \begin{pmatrix} t_1 & \cdots & t_n \\ t_1 & \cdots & t_n \end{pmatrix} dt_1 \cdots dt_n \quad (12)$$

となる [Hochstadt (1973)]. ここで, K_N は (j, k) 要素が $K(j/N, k/N)$ である $N \times N$ の行列, $\lambda_1, \lambda_2, \dots$ は固有値であり, 重複度の分だけ繰り返される. また,

$$K \begin{pmatrix} s_1 & \cdots & s_n \\ t_1 & \cdots & t_n \end{pmatrix} = \begin{vmatrix} K(s_1, t_1) & K(s_1, t_2) & \cdots & K(s_1, t_n) \\ \vdots & \vdots & & \vdots \\ K(s_n, t_1) & K(s_n, t_2) & \cdots & K(s_n, t_n) \end{vmatrix}$$

である.

以上の準備から, V_g の特性関数を求めるための Anderson-Darling (1952) の定理が得られる.

定理 3 (Anderson-Darling). 統計量

$$S = \int_0^1 \int_0^1 K(s, t) dW(s) dW(t)$$

において, $K(s, t)$ は対称, 連続な関数とする. このとき,

$$\begin{aligned} E(e^{i\theta S}) &= E \left[\exp \left\{ i\theta \int_0^1 \int_0^1 K(s, t) dW(s) dW(t) \right\} \right] \\ &= (D(2i\theta))^{-1/2}. \end{aligned}$$

証明: S の近似として,

$$S_N = \frac{1}{N} \mathbf{Z}' K_N \mathbf{Z} = \frac{1}{N} \sum_{j=1}^N \sum_{k=1}^N K\left(\frac{j}{N}, \frac{k}{N}\right) Z_j Z_k, \quad \mathbf{Z} \sim N(\mathbf{0}, I_N),$$

を考える. ここで, K_N は (j, k) 要素が $K(j/N, k/N)$ の $N \times N$ 行列である. $N \rightarrow \infty$ のとき, $S_N \Rightarrow S$ である. また, S_N の特性関数は,

$$E(e^{i\theta S_N}) = \left| I_N - \frac{2i\theta}{N} K_N \right|^{-1/2}$$

であり, (11) の結果から,

$$\lim_{N \rightarrow \infty} E(e^{i\theta S_N}) = E(e^{i\theta S}) = (D(2i\theta))^{-1/2} = \prod_{n=1}^{\infty} \left(1 - \frac{2i\theta}{\lambda_n}\right)^{-1/2}$$

となることは明らかである.

同次第二種フレッドホルム積分方程式の固有値の性質:

- a) $D(\lambda) = 0$ の根は固有値である. 逆に, 積分方程式の固有値は $D(\lambda) = 0$ の根となる.
- b) $D(\lambda)$ は整関数であり, $D(0) = 1$ である. したがって, 固有値は 0 とはならない.
- c) 固有値の数は高々可算無限であり, 無限個ある場合は $\lambda = \infty$ が唯一の集積点である.
- d) 各固有値の重複度は有限である.
- e) 核関数 $K(s, t)$ が対称, 連続, ほぼ定符号 (固有値が有限個を除いて同符号) ならば, Mercer の定理

$$K(s, t) = \sum_{n=1}^{\infty} \frac{f_n(s)f_n(t)}{\lambda_n} \quad (\text{一様, 絶対収束})$$

が重複度を込めて成り立つ. ここで, $\{f_n(t)\}$ は正規直交固有関数である. このことから, 次の関係が得られる.

$$\int_0^1 \int_0^1 K(s, t) dW(s) dW(t) \stackrel{D}{=} \sum_{n=1}^{\infty} \frac{1}{\lambda_n} Z_n^2, \quad \{Z_n\} \sim \text{NID}(0, 1),$$

$$D(\lambda) = \prod_{n=1}^{\infty} \left(1 - \frac{\lambda}{\lambda_n}\right).$$

- f) 平均 0 の正規過程 $\{Y_t\}$ に対しては, Karhunen-Loève 展開

$$Y(t) = \sum_{n=1}^{\infty} \frac{f_n(t)}{\sqrt{\lambda_n}} Z_n, \quad \{Z_n\} \sim \text{NID}(0, 1)$$

が成り立つ. ここで, $\{\lambda_n\}$ は, 正値定符号の核関数 $K(s, t) = \text{Cov}(Y(s), Y(t))$ の固有値である. このことから, 次の関係が得られる.

$$\int_0^1 Y^2(t) dt = \sum_{n=1}^{\infty} \frac{1}{\lambda_n} Z_n^2 \stackrel{D}{=} \int_0^1 \int_0^1 \text{Cov}(Y(s), Y(t)) dW(s) dW(t).$$

4. フレッドホルム行列式の計算

積分方程式

$$f(t) = \lambda \int_0^1 K_g(s, t) f(s) ds, \quad K_g(s, t) = \int_{\max(s, t)}^1 \frac{((u-s)(u-t))^g}{(g!)^2} du,$$

から固有値を求めることは一般に困難である. そこで, 積分方程式と同値な微分方程式および境界条件を考える. この積分方程式は, 微分方程式

$$f^{(2g+2)}(t) + (-1)^g \lambda f(t) = 0$$

と $2g + 2$ 個の境界条件

$$f^{(j)}(1) = 0, \quad f^{(g+1+j)}(0) = 0 \quad (j = 0, 1, \dots, g)$$

と同値になる．証明は Tanaka (2023)．この微分方程式の一般解は，

$$f(t) = c_1 \cos x_1 t + c_2 \sin x_1 t + \dots + c_{2g+1} \cos x_{g+1} + c_{2g+2} \sin x_{g+1} \quad (13)$$

と表すことができる．ここで， $\mathbf{c} = (c_1, c_2, \dots, c_{2g+2})'$ は任意定数である．また， x_1, x_2, \dots, x_{g+1} は特性方程式 $x^{2g+2} = (-1)^{g-1} \lambda$ の根である．根は $2g + 2$ 個あるが，このうちの $g + 1$ 個は残りの $g + 1$ 個にマイナスの符号を付けたものであるので，このようなペアとならない $g + 1$ 個の根を使えばよい． $2g + 2$ 個の境界条件を (13) に代入することにより， \mathbf{c} に関する $2g + 2$ 本の方程式 $M(\lambda) \mathbf{c} = \mathbf{0}$ が得られる．このとき， $|M(\lambda)| = 0$ となることが $K(s, t)$ の固有値であることの必要十分条件となる．ただし， $\lambda \neq 0$ ．また，

$$\text{固有値 } \lambda_n(g) \text{ の重複度} = 2g + 2 - \text{rank}(M(\lambda_n(g))) = \text{零化空間の次元}$$

となる．

$K_g(s, t)$ の FD を $D_g(\lambda)$ とする． D_g を任意の g に対して求めることはできない．また，逐次的に求めることも不可能である．以下， $g = 0, 1, 2, 3$ の結果を示す． $K(s, t)$ は対称なので $s \leq t$ の表現のみを記してある．なお， $g = 4, 5$ の場合の FD も求めたが，最小固有値（表 1 に記載）以外はここでは省略する．

$$K_0(s, t) = 1 - t,$$

$$D_0(\lambda) = \cos \sqrt{\lambda},$$

$$K_1(s, t) = \frac{1}{6} (1 - t)^2 (t + 2 - 3s),$$

$$D_1(\lambda) = \frac{1}{2} (1 + \cos \lambda^{1/4} \cosh \lambda^{1/4}),$$

$$K_2(s, t) = \frac{1}{120} (1 - t)^2 (t^2 + 3t + 6 - 5st + 10s^2 - 15s),$$

$$D_2(\lambda) = \frac{1}{9} [2(1 + \cos \lambda^{1/6} + \cos \lambda^{1/6} \omega + \cos \lambda^{1/6} \omega^2) + \cos \lambda^{1/6} \cos \lambda^{1/6} \omega \cos \lambda^{1/6} \omega^2],$$

$$\omega = \frac{1 + \sqrt{3}i}{2},$$

$$K_3(s, t) = \frac{1}{5040} (t - 1)^4 (t^3 + 4t^2 + 10t + 20 - 35s^3 + 21s^2t + 84s^2 - 7st^2 - 28st - 70s),$$

$$D_3(\lambda) = \frac{1}{16} [3 \cos a \cos b \cos c \cos d + 2(\cos a \cos b + \cos b \cos c + \cos c \cos d + \cos d \cos a) + \cos a \cos c + \cos b \cos d + 3]$$

$$+ \frac{\sqrt{2}}{16} [\sin a \sin b (1 + \cos c \cos d) + \sin b \sin c (1 + \cos d \cos a) + \sin c \sin d (1 + \cos a \cos b) - \sin d \sin a (1 + \cos b \cos c)],$$

$$a = \lambda^{1/8}, \quad b = a\omega, \quad c = a\omega^2, \quad d = a\omega^3, \quad \omega = \frac{1 + i}{\sqrt{2}}.$$

5. ブラウン運動の正値 2 次汎関数の分布の計算

V_g は正値確率変数であるから、分布関数は Lévy の反転公式より、

$$P(V_g < x) = \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{1 - e^{-i\theta x}}{i\theta} (D_g(2i\theta))^{-1/2} \right] d\theta$$

により計算できる。 V_g は、Mercer の定理より、

$$\begin{aligned} V_g &= \int_0^1 F_g^2(t) dt = \int_0^1 \int_0^1 K_g(s, t) dW(s) dW(t) \\ &\stackrel{D}{=} \sum_{n=1}^\infty \frac{1}{\lambda_n(g)} Z_n^2, \quad \{Z_n\} \sim \text{NID}(0, 1) \end{aligned}$$

と表されるので、 V_g の分布は小さい固有値に依存して決まる。固有値がすべて正の場合には、

$$\left(\sum_{n=1}^\infty \frac{1}{\lambda_n(g)} \right)^2 = \left(\int_0^1 K_g(t, t) dt \right)^2 > \sum_{n=1}^\infty \frac{1}{\lambda_n^2(g)} = \int_0^1 \int_0^1 K_g^2(s, t) ds dt > \frac{1}{\lambda_1^2(g)}.$$

が成り立つ。したがって、最小固有値の逆数の 1 つの上界として、

$$B_g = \sqrt{\int_0^1 \int_0^1 K_g^2(s, t) ds dt} > \frac{1}{\lambda_1(g)}$$

を考えることができる。表 1 には、 $g = 0, 1, \dots, 10$ の場合の $E(V_g)$ (=固有値の逆数の総和)、最小固有値の逆数 $1/\lambda_1(g)$ 、その上界 B_g 、最小固有値の逆数と固有値の逆数の総和の比 $r_1(g)$ 、および最小固有値の逆数の上界と固有値の逆数の総和の比 $r_2(g)$ が示されている。

表 1. 最小固有値と上界

g	0	1	2	3	4	5
$E(V_g)$	1/2	1/12	1/120	1/2016	1/51840	1/1900800
$1/\lambda_1(g)$	4.053/10	8.0891/10 ²	8.24681/10 ³	4.934432/10 ⁴	1.922972/10 ⁵	5.249965/10 ⁷
B_g	4.082/10	8.0917/10 ²	8.24721/10 ³	4.934497/10 ⁴	1.922981/10 ⁵	5.249976/10 ⁷
$r_1(g)$	0.8106	0.9707	0.9896	0.9948	0.9969	0.9979
$r_2(g)$	0.8165	0.9710	0.9897	0.9948	0.9969	0.9979

g	6	7	8	9	10
$E(V_g)$	1.05990/10 ⁸	1.64032/10 ¹⁰	2.01019/10 ¹²	1.99844/10 ¹⁴	1.64374/10 ¹⁶
B_g	1.05831/10 ⁸	1.63849/10 ¹⁰	2.00845/10 ¹²	1.99705/10 ¹⁴	1.64280/10 ¹⁶
$r_2(g)$	0.9985	0.9989	0.9991	0.9993	0.9994

6. 多重単位根分布の計算

多重単位根統計量 R_g の分布は, Imhof (1961) の公式から,

$$\begin{aligned} P(R_g < x) &= P\left(x \int_0^1 F_g^2(t) dt - \frac{1}{2} F_g^2(1) > 0\right) \\ &= P\left(\int_0^1 \int_0^1 K_g(s, t; x) dW(s) dW(t) > 0\right) \end{aligned} \quad (14)$$

$$= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{1}{\theta} \operatorname{Im} \left[\left(D_g(2i\theta; x) \right)^{-1/2} \right] d\theta \quad (15)$$

により計算することができる. ここで, $D_g(\lambda; x)$ は,

$$K_g(s, t; x) = \frac{1}{(g!)^2} \left[x \int_{\max(s, t)}^1 ((u-s)(u-t))^g du - \frac{1}{2} ((1-s)(1-t))^g \right]$$

の FD である. $K_g(s, t; x)$ は対称, 連続, ほぼ定符号の核関数である.
積分方程式

$$f(t) = \lambda \int_0^1 K_g(s, t; x) f(s) ds$$

は, 微分方程式

$$f^{(2g+2)}(t) + (-1)^g \lambda x f(t) = 0$$

および $2g + 2$ 個の境界条件

$$\begin{aligned} f(1) &= f'(1) = \dots = f^{(g-1)}(1) = 0, \quad f^{(g)}(1) = \frac{(-1)^{g-1} \lambda}{2g!} \int_0^1 (1-s)^g f(s) ds, \\ f^{(g+1)}(0) &= f^{(g+2)}(0) = \dots = f^{(2g+1)}(0) = 0. \end{aligned}$$

と同値である. 証明は Tanaka (2023).

7. 多重単位根分布のモーメント

多重単位根統計量

$$R_g = \frac{1}{2} F_g^2(1) / \int_0^1 F_g^2(t) dt = \frac{U_g}{V_g}$$

の k 次のモーメントは

$$E \left[\left(\frac{U_g}{V_g} \right)^k \right] = \frac{1}{(k-1)!} \int_0^\infty \theta_2^{k-1} \frac{\partial^k \psi(\theta_1, -\theta_2)}{\partial \theta_1^k} \Big|_{\theta_1=0} d\theta_2$$

から求めることができる. ここで, $\psi(\theta_1, \theta_2)$ は U_g と V_g の同時積率母関数 (m.g.f.) である. $g = 0, 1, 2$ の場合について計算した結果から, 任意の g に対して, 私は

$$E(R_g) = g + 1$$

となることを予想したが, Jin-Li (2015) により, 正しいことが証明された.

8. 分布のグラフ表示

図 1. $\int_0^1 F_g^2(t) dt \stackrel{\mathcal{D}}{=} \sum_{n=1}^{\infty} \frac{1}{\lambda_n(g)} Z_n^2$ における最小固有値の対数值

横軸 : $g = 0, 1, \dots, 10$, 縦軸 : $\log(\lambda_1(g))$

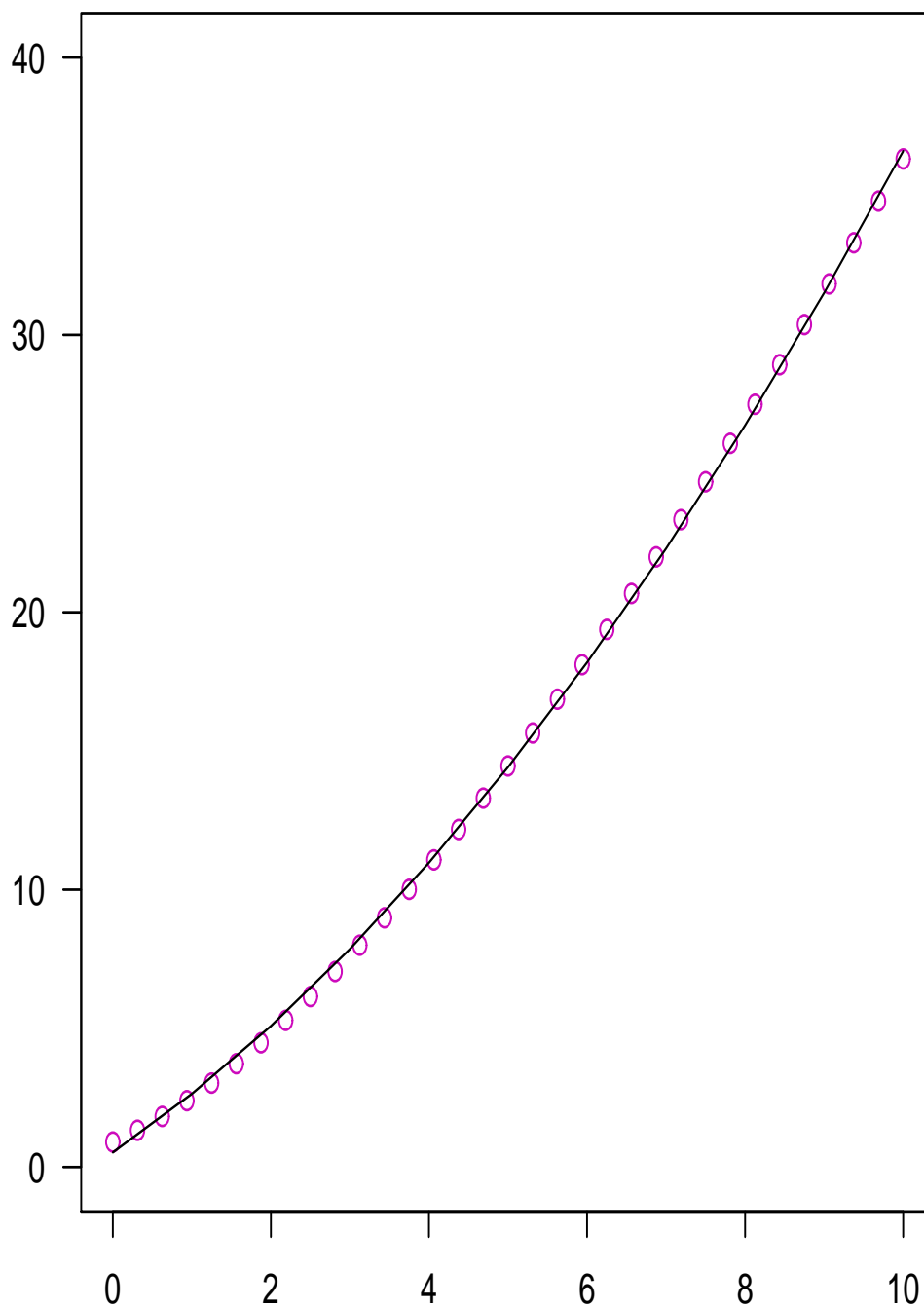
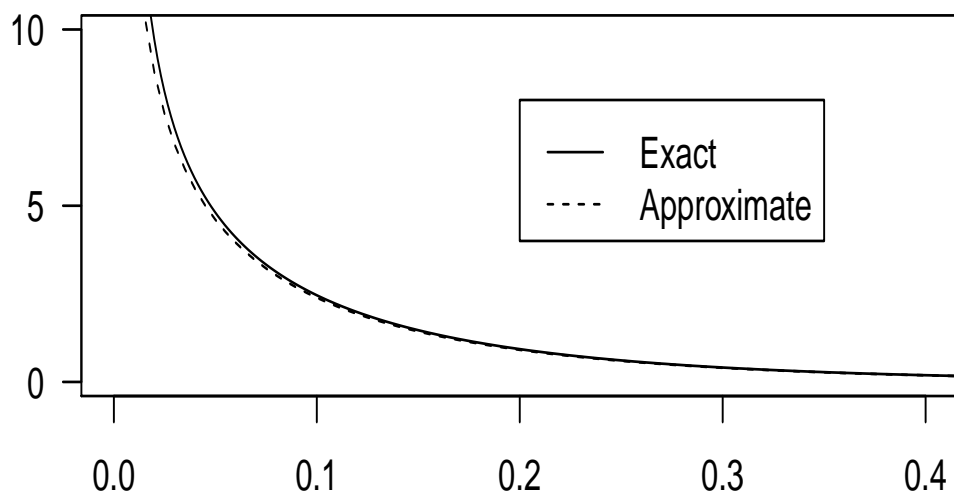


図 2. $\int_0^1 F_g^2(t) dt$ の密度関数

g=1



g=2

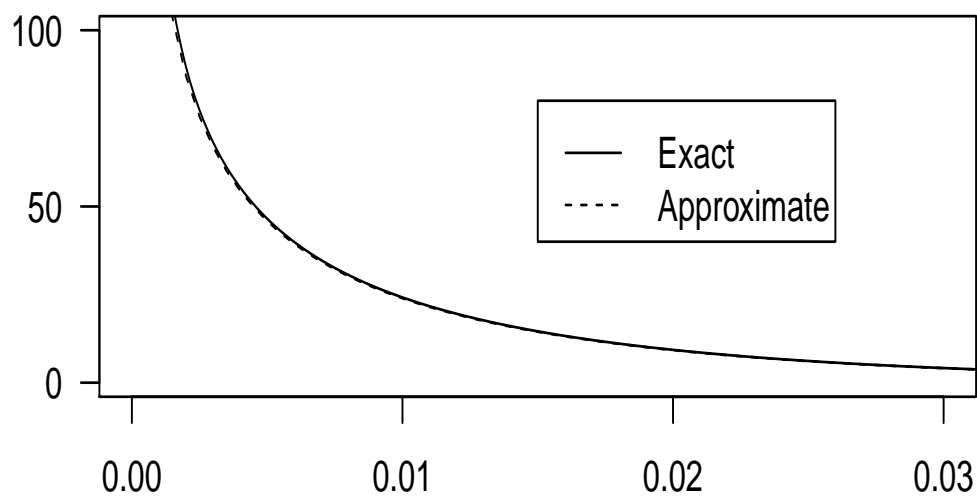
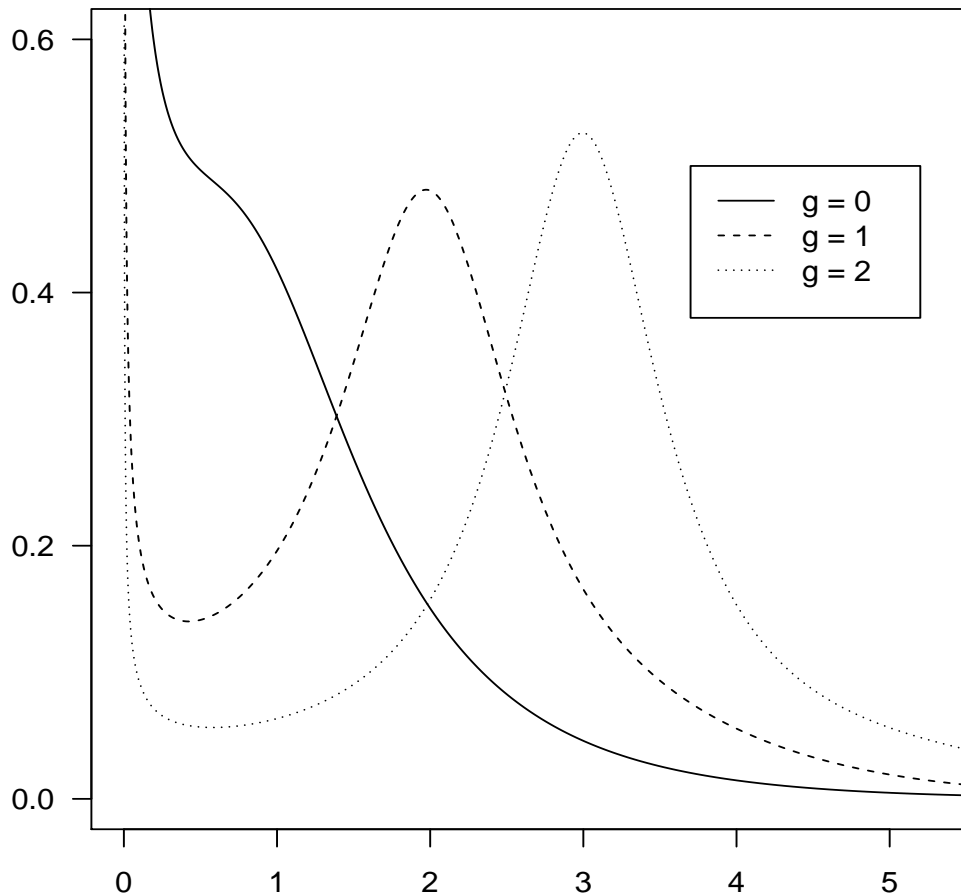


図 3. 多重単位根分布: $\frac{1}{2} F_g^2(1) / \int_0^1 F_g^2(t) dt$ の密度関数



参考文献

- Anderson, T.W. and Darling, D.A. (1952). "Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes," *Annals of Mathematical Statistics*, **23**, 193-212.
- Chan, N.H. and Wei, C.Z. (1988). "Limiting distributions of least squares estimates of unstable autoregressive processes," *Annals of Statistics*, **16**, 367-401.
- Hochstadt, H. (1973). *Integral Equations*, Wiley, New York.
- Imhof, J.P. (1961). "Computing the distribution of quadratic forms in normal variables," *Biometrika*, **48**, 419-426.
- Jin, S. and Li, W.V. (2015). "Expectation of the limiting distribution of the LSE of a unit root process," *Statistica Sinica*, **25**, 529-536.
- Tanaka, K. (2017). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, 2nd ed., Wiley, New York.
- Tanaka, K. (2023). *Brownian Motion and the Fredholm Determinant*, to be published.