Ph.D. Seminar Series in Advanced Mathematical Methods in Economics and Finance

UNIT ROOT DISTRIBUTION THEORY II

Roderick McCrorie

School of Economics and Finance
University of St Andrews

23 April 2009

Background

- Most economic time series in raw form appear exhibit trend behaviour

- We would like to construct a test that allows us to discriminate between deterministic and stochastic trending behaviour, because the underlying nature of the trend materially affects statistical inference

- As we’ve seen, even in the simplest framework of an AR(1) model, there are aspects of this problem that are non-standard

- The essential statistical problem is that estimates of the parameters of a model converge to the true parameter value at <u>different rates in different regions of the parameter space</u>, and some limiting distributions are non-standard

- While econometricians have invented ways around this problem, properly embedding the problem in a probabilistic framework remains open

- My contention is that the problem is intimately connected to aspects of planar Brownian motion

- See especially:-

Pitman & Yor (1986) Asymptotic laws of planar Brownian motion, *Annals of Probability* 14, 733-779.

- Furthermore, the problem can be generalized to a wider, possibly universal, context of problems that deal with Brownian functionals in all their aspects

Banderier et al (2000) Planar maps and Airy phenomena. *Lecture Notes in Computer Science*.

Flajolet & Louchard (2001) Analytic variations on the Airy distribution. *Algorithmica* 31(3), 361-377.

Matsumoto & Yor (2005) Exponential functionals of Brownian motion. *Probability Surveys* 2, 312-384.

Lyasoff (2007) Variational approach to the study of certain integral functionals of Brownian motion, preprint.

- In Econometrics, the “unit root problem” involves trying to characterize the (properties of) the density of the OLS estimator of the parameter in the AR(1) model and generalizations under the null hypothesis that it is unity

- This work involves a quadratic functional (the 2-norm) of Brownian motion

Lecture II plan

- Discussion of the unit root problem
- Outline of the standard approach of deriving laws as Wiener functionals
- Placing the unit root problem in the context of Levy's stochastic area
- Deriving moments of the distributions
- Giving a probabilistic interpretation to some integrals and densities that emerge
- Relations to the Riemann zeta function and various generalizations, and results in:-

  Pitman & Yor (2003) Infinitely divisible laws associated with hyperbolic functions. *Canadian Journal of Mathematics* 55, 292-330

REGRESSION WITH NON-STATIONARY TIME SERIES

Consider the simple $AR(1)$ model

$$x_t = \rho x_{t-1} + \varepsilon_t, \qquad (t = 1, \ldots, n)$$

where ρ is an unknown parameter, $x_0 = 0$, and

$\varepsilon_t \sim NID(0, 1)$.

Consider the joint density

$$f_\rho(x) = (\sqrt{2\pi})^{-n} \exp -\tfrac{1}{2}[(1+\rho^2)T_2 - 2\rho T_1 + x_n^2],$$

where

$$T_1 = \textstyle\sum_{t=1}^n x_t x_{t-1}, \qquad T_2 = \sum_{t=1}^n x_{t-1}^2 .$$

The maximum likelihood estimator (MLE) $\hat{\rho}_n$ of ρ is given by

$$\hat{\rho}_n = T_1 / T_2 .$$

It is worth taking a short digression here to discuss the possibility of obtaining exact results.

All Gaussian estimators in this context – not just the one above – can be expressed as a ratio of quadratic forms in normal variables $R = \dfrac{\varepsilon' A \varepsilon}{\varepsilon' B \varepsilon}$.

The distribution for *R* is given by

$$P(R \le r) = P\left(\frac{\varepsilon' A \varepsilon}{\varepsilon' B \varepsilon} \le r\right) = P(\varepsilon'(A - rB)\varepsilon \le 0)$$

$$= P(X_r \le 0)$$

Take the spectral decomposition

$$A - rB = P_r' \Lambda_r P_r .$$

Then with the eigenvalues ordered to be in ascending order, the distribution of X_r is given by

$$X_r = \sum_{i=1}^{n} \lambda_i \chi^2(1, \upsilon_i^2), \quad \varepsilon \equiv N(\mu, I), \upsilon = P_r \mu$$

- It is possible to evaluate this distribution using an inversion formula that involves the characteristic function of X_r.

IMHOF METHOD

Uses the Gil-Peleaz (1951) inversion theorem

$$F_{X_r}(x)=\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \frac{|z|\sin(\arg z)}{t}dt,\ \ z=e^{-itx}\phi_{X_r}(t),$$

where $\phi_{X_r}(t)$ is the characteristic function of X_r, to derive a more convenient formula

$$F_{X_r}(x)=\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \frac{\sin\beta(u,x)}{u\gamma(u)}du\,,$$

where $\beta(u,x)$ and $\gamma(u)$ are explicit functions of the eigenvalues λ_i.

Gil-Peleaz, J. (1951) Note on the inversion theorem. *Biometrika* 38, 481-482.

Imhof, J.P. (1961) Computing the distribution of quadrartic forms in normal variables. *Biometrika* 48, 419-426.

- In even the simplest AR model, this is complicated by two aspects:-

  - generically the distribution function may fail to be analytic at certain points of its domain and the inversion problem is highly oscillatory

  - the eigenvalues in the unit root case are only available implicitly (see, e.g., Tanaka, 1996, p. 16).

  Generally, finding the eigenvalues can be facilitated by the Fredholm approach described in Ch. 5 of Tanaka (1996).

References involving laws of quadratic forms

G. Hillier (2001) The density of a quadratic form uniformly distributed on the *n*-sphere, *Econometric Theory* 17, 1-28

G. Forchini (2002) The exact cumulative distribution function of a ratio of quadratic forms on normal variables, with application to the *AR*(1) model, *Econometric Theory*, 16, 823-852.

R.W. Butler & M.S. Paolella (2008) Uniform saddlepoint approximation for ratios of quadratic forms. *Bernoulli*, 14(1), 140-154.

Related literature

Pycke (2007) *U*-statistics based on the Green's function of the Laplacian on the circle and the sphere. *Statistics and Probability Letters* 77, 863-872.

Henze & Nikitin (2002) Watson-type goodness-of-fit tests based in the integrated empirical process. *Mathematical Methods of Statistics* 11, 183-202.

Early References involving determinants

J.S. White (1958) The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.*, 29, 1188-1197.

J.S. White (1959) The limiting distribution of the serial correlation coefficient in the explosive case II. *Ann. Math. Statist.*, 30, 831-834.

J.S. White (1961) Asymptotic expansions for the mean and variance of the serial correlation coefficient, *Biometrika* 48, 85-94.

There may also be a connection with:-

Keisake Hara (2004) Finite dimensional determinants as characteristic functions of quadratic Wiener functionals. *Elect. Comm. In Probab.*, 9, 26-35.

General references

Tanaka (1996) *Time Series Analysis*, Wiley.
Hamilton (1994) *Time Series Analysis*, Princeton.

The focus of today's talk is on a Brownian functional that arises as a limiting distribution in the $AR(1)$ model

$$x_t = \rho x_{t-1} + \varepsilon_t, \qquad (t = 1, \ldots, n)$$

where $\varepsilon_t \sim NID(0, 1)$.

The maximum likelihood estimator (MLE) $\hat{\rho}_n$ of ρ is given by

$$\hat{\rho}_n = \frac{\sum_{t=1}^{n} x_t x_{t-1}}{\sum_{t=1}^{n} x_{t-1}^2}.$$

If $|\rho| < 1$, the process is stationary and "the textbook treatment" goes something like this:

$$\sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{d} N(0, 1-\rho^2),$$

where $\xrightarrow{d}$ denotes convergence in distribution.

When $\rho = 1$, this isn't useful as a basis of testing for a unit root against stationarity; however it can be shown that

$$n^{-1}\sum_{t=1}^{n} x_{t-1}\varepsilon_t \xrightarrow{d} \tfrac{1}{2}(\chi_1^2 - 1),$$

which with

$$\sum_{t=1}^{n} x_{t-1}^2 = O(n^2),$$

suggests writing

$$n(\hat{\rho}_n - 1) = n^{-1}\sum_{t=1}^{n} x_{t-1}\varepsilon_t \Big/ n^{-2}\sum_{t=1}^{n} x_{t-1}^2 .$$

- There is no LLN with this normalization such that the denominator converges to a constant.

- What the expression does suggest is that if the statistic has a well-defined limit distribution, the ML estimator in this case converges to this distribution at a faster rate than it does to a normal distribution in the stationary case.

In fact, the asymptotic sampling properties of the ML estimator are determined by the <u>choice of initial condition</u>.

We could make the problem "circular" by putting say $x_{n+1} = x_1$, in which case we could (essentially) work with a distributional approximation derived by Leipnik, expressing the density as the sum of a *t*-density and another.

References (for the serial correlation coefficient)

R.B. Leipnik (1947) Distribution of the serial correlation coefficient in a circularly correlated universe, *Ann. Math. Stat.* 18, 80-87.

J.S. White (1957) A *t*-test for the serial correlation coefficient. *Ann. Math. Stat.* 28(4), 1046-1048 (and acknowledgement of priority: 1958, 29(3), p. 935).

This assumption, however, isn't appropriate for the trending series we see.

Let's try another initial condition:

Theorem ([Mann and Wald 43, White 58, Anderson 59, Dickey and Fuller 79]) *Let* $x_0 = 0$. *Then*

$$\sqrt{I_n(\rho)}\,(\hat{\rho}_n - \rho) \xrightarrow{d} \begin{cases} N, & |\rho| < 1 \\ \rho \dfrac{W^2(1) - 1}{2^{3/2}\int_0^1 W^2(r)\,dr}, & |\rho| = 1 \\ C, & |\rho| > 1 \end{cases}$$

where N *is a centred Gaussian random variable with variance* 1, C is a standard Cauchy random variable, $W = (W(r),\ r \geq 0)$ *is a standard Wiener process and* $I_n(\rho)$ *is the* (*expected*) *Fisher information contained in* $x_1, \ldots, x_n$ *about the parameter* ρ, *given by* $E(T_2)$. *As* $n \to \infty$,

$$E(T_2) \sim \begin{cases} \dfrac{n}{1-\rho^2}, & |\rho| < 1, \\ \dfrac{n^2}{2}, & |\rho| = 1, \\ \dfrac{\rho^{2n}}{(\rho^2 - 1)^2}, & |\rho| > 1. \end{cases}$$

- If $|\rho| \leq 1$, the result remains valid under the weaker assumptions that x_0 is an arbitrary constant or a random variable with a finite second moment not depending on the sequence $\varepsilon = (\varepsilon_t, t \geq 1)$ and ε is a arbitrary sequence of centred and normalized i.i.d. random variables.

- If $|\rho| > 1$, the limit distribution depends on the initial value, and in general on the particular distribution of each ε_t, even if ε forms a sequence of i.i.d. random variables [Koul & Pflug].

- Different results hold if ε forms a sequence of i.i.d. random variables with a stable distribution or in the domain of attraction of a stable law [Chan and Tran; Phillips 1990; Mijnheer 1997].

If we normalize by the observed rather than the expected Fisher information, the three limit distributions are reduced to two:

Theorem [Mann and Wald, Anderson, Dickey Fuller] *Let* $x_0 = 0$. *Then*

$$\sqrt{T_2}\,(\hat{\rho}_n - \rho) \xrightarrow{d} \begin{cases} N, & |\rho| \neq 1, \\ \rho \dfrac{W^2(1) - 1}{2\sqrt{\int_0^1 W^2(r)\,dr}}, & |\rho| = 1. \end{cases}$$

It is possible to use a *sequential maximum likelihood estimator* to obtain a unique limiting distribution but this would involve sampling up to a prescribed amount of Fisher information and isn't natural in the current context.

Shiryaev, A.N. & V.G. Spokoiny (1997) On sequential estimation of an autoregressive parameter. *Stochastics and Stochastic Reports* 60, 219-240.

In fact, there are <u>three</u> basic time series models that are used as building blocks in Econometrics.

This reflects

- the difficulty of treating the initial condition
- the desire to test the random walk hypothesis against a linear trend.

MODEL 1: $x_t = \rho x_{t-1} + \varepsilon_t, \qquad (t = 1, \ldots, n)$

MODEL 2: $x_t = \alpha + \rho x_{t-1} + \varepsilon_t, \qquad (t = 1, \ldots, n)$

MODEL 3: $x_t = \alpha + \beta t + \rho x_{t-1} + \varepsilon_t, \quad (t = 1, \ldots, n)$

For Model 2,

$$n(\hat{\rho}_T - 1) \xrightarrow{D} U_2 / V_2$$

$$t_T \xrightarrow{D} U_2 / \sqrt{V_2}$$

where

$$U_2 = \begin{vmatrix} \int_0^1 W(r)dW(r) & \int_0^1 W(r)dr \\ \int_0^1 dW(r) & 1 \end{vmatrix}$$

and

$$V_2 = \begin{vmatrix} \int_0^1 W^2(r)dr & \int_0^1 W(r)dr \\ \int_0^1 W(r) & 1 \end{vmatrix}$$

For Model 3,

$$n(\hat{\rho}_T - 1) \xrightarrow{D} U_3 / V_3$$

$$t_T \xrightarrow{D} U_3 / \sqrt{V_3}$$

where

$$U_3 = 12 \begin{vmatrix} \int_0^1 W(r)dW(r) & \int_0^1 W(r)dr & \int_0^1 r\,W(r)dr \\ \int_0^1 dW(r) & 1 & \frac{1}{2} \\ \int_0^1 r dW(r) & \frac{1}{2} & \frac{1}{3} \end{vmatrix}$$

and

$$V_3 = 12 \begin{vmatrix} \int_0^1 W^2(r) & \int_0^1 W(r)dr & \int_0^1 r\,W(r)dr \\ \int_0^1 W(r) & 1 & \frac{1}{2} \\ \int_0^1 rW(r) & \frac{1}{2} & \frac{1}{3} \end{vmatrix}$$

The above formulae are related to results in Jandhyala, V.K. & I. B MacNeill (1991) *J. Statist. Plann., Inference* 27, 291-316.

As we saw this morning, these and other fomulae can be derived using a “two-stage” approach to examining limiting distribution theory pertaining to non-stationary regression, following work by P.C.B. Phillips (1987, *Econometrica*)

- It is extremely powerful. But it also has its drawbacks.

- It involves in the first stage using a FCLT to derive a limit distribution for the (normalized) integrated process:

$$n^{-1/2}x_{[rn]} \Rightarrow \sigma W(r)$$

  The advantage here is that this can be made to apply under assumptions that suit the data in hand (e.g. Donsker, Erdos and Kac, McLeish, Hernndorff, Domowitz and White, de Jong and Davidson, Doukhan and Louhichi, Beare).

- In the second stage, we derive the limit distribution of the sample statistic based explicitly on its construction as a functional of the integrated process:

$$n^{-1}\sum_{t=1}^{n} T(n^{-1/2}x_t) \xrightarrow{d} \int_0^1 T(\sigma W(r))dr .$$

EXAMPLE

Recall the MLE $\hat{\rho}_n$ of ρ in Model 1 is given by

$$\hat{\rho}_n = \frac{\sum_{t=1}^{n} x_t x_{t-1}}{\sum_{t=1}^{n} x_{t-1}^2}.$$

For the denominator above,

$$\begin{aligned} n^{-2}\sum_{t=1}^{n} x_{t-1}^2 &= n^{-1}\sum_{t=1}^{n}\left(n^{-1/2}\sum_{s=1}^{t-1}\varepsilon_s\right)^2 \\ &= n^{-1}\sum_{t=1}^{n}\left(n^{-1/2}\sum_{s=1}^{[rn]}\varepsilon_s\right)^2 \quad \left(\tfrac{t-1}{n} \le r \le \tfrac{t}{n}\right) \\ &\xrightarrow{d} \sigma^2\int_0^1 W(r)^2\, dr, \qquad (a) \end{aligned}$$

For the numerator,

$$\begin{aligned} n^{-1}\sum_{t=1}^{n} x_{t-1}\varepsilon_t &= \sum_{t=1}^{n}\left(n^{-1/2}\sum_{s=1}^{t-1}\varepsilon_s\right)n^{-1/2}\varepsilon_t \\ &\xrightarrow{d} \sigma^2\int_0^1 W(r)\,dW(r). \qquad (b) \end{aligned}$$

The integral in (b) is an Ito integral equal to $\frac{1}{2}(W^2(1)-1)$. The joint convergence of (a) and (b) gives the Dickey-Fuller distribution up to a normalization.

Note that the σ^2 cancels.

- Recent work has allowed this approach to be extended to the analysis of non-linear transformations of integrated processes and an asymptotic theory of inference that applies to non-linear regression.

  e.g. Park & Phillips, de Jong, Pötscher, de Jong & Wang, and Berkes & Horváth

- Major problems remain unsolved in spite of the importance of the unit root problem

  What is the natural probabilistic setting for this work?

  How do we characterize the properties of the densities?

Other properties of the ML estimator

Theorem [Mikulski and Monsour 91] *As* $n \to \infty$,

$$\sup_\rho E_\rho |\hat{\rho}_n - \rho| \to 0,$$

and in particular the estimators $\hat{\rho}_n$ *are* ***asymptotically uniformly consistent***, *i.e. as* $n \to \infty$,

$$\sup_\rho P_\rho \left(|\hat{\rho}_n - \rho| > \varepsilon \right) \to 0 \quad \textit{for all } \varepsilon > 0 .$$

Theorem [Mikulski and Monsour 91] *Consider a class of estimators such that, given* ρ, *the bias*

$$b_\rho(\rho_n) = E_\rho(\hat{\rho}_n - \rho),$$

is differentiable in ρ *and satisfies the conditions*

$$b_\rho(\rho_n) \to 0, \ \frac{d\, b_\rho(\rho_n)}{d\,\rho} \to 0, \text{ as } n \to \infty .$$

(Then the ML estimators belong to this class for each $|\rho| \neq 1$.)

1. *For each* $|\rho| \neq 1$, *the ML estimator* $\hat{\rho}_n$ *is asymptotically efficient in this class*: *if* ρ_n *is another member in the class, then*

$$\lim\sup_n \frac{E_\rho(T_2)(\hat{\rho}_n - \rho)^2}{E_\rho(T_2)(\rho_n - \rho)^2} \leq 1.$$

2. *In the case* $|\rho| < 1$, *the estimators* $\hat{\rho}_n$ *are efficient also in the* "*classical sense*", *i.e.*

$$\lim\sup_n \frac{Var_\rho(\hat{\rho}_n)}{Var_\rho(\rho_n)} \leq 1.$$

When $|\rho| = 1$, the ML estimators do NOT belong to this class.

- Here, we simply focus on the zero-mean AR(1) model, thinking of it as an <u>artificial</u> generating mechanism?

- As Abadir (1992) shows, if we know the densities of the relevant statistics for this model, we can generate the others using a transformation theorem.

- We still don’t have satisfactorily computable expressions (exact or approximate) for the densities and the usual approach is to derive Wiener functionals and simulate critical values.

The density of the Dickey-Fuller distribution is encoded in the following Laplace transform:

$$E\left[\exp\left(-\alpha\left(\tfrac{1}{2}B_1^2-\tfrac{1}{2}\right)-\tfrac{1}{2}\beta^2\int_0^1 B_s^2 ds\right)\right]$$

$$=\exp\left(\tfrac{1}{2}\alpha\right)\times E\left[\exp\left(-\tfrac{\alpha}{2}B_1^2-\tfrac{1}{2}\beta^2\int_0^1 B_s^2 ds\right)\right]$$

$$=\exp\left(\tfrac{1}{2}\alpha\right)\left(\cosh(\beta)+\frac{\alpha}{\beta}\sinh(\beta)\right)^{-1/2},$$

where the Brownian motion starts at 0.

At the heart of this formula is the following conditional Laplace transform:

$$E\left[\exp\left(-\tfrac{1}{2}\beta^2\int_0^1 B_s^2 ds \,\middle|\, B_1=a\right)\right]$$

$$=\left(\frac{\beta}{\sinh(\beta)}\right)^{1/2}\exp\left(-\frac{a^2}{2}\left(\beta\coth(\beta)-1\right)\right),$$

This is akin to Levy's formula for the area enclosed by the trajectory of a Wiener process and its chord.

- The above Laplace transforms are neatly expressed but are very difficult to invert into anything other than series that are oscillatory!

- We can see the source of the expressions involving parabolic cylinder functions that Karim Abadir derived in several papers in the 1990's:

- The density of $\int_0^1 B_s^2 ds$ given $B_1 = 0$ computed by Tolmatz (*Ann. Probab*., 2002) involves parabolic cylinder functions, the source of which is the Laplace inversion of $\left(\dfrac{\beta}{\sinh(\beta)}\right)^{1/2}$

  See:-

  Biane & Yor (1987) Valeurs principales associees aux temps locaux browniens *Bull. Sc. Math*., 111, 23-101.

Note also that

$$\left(\frac{\sqrt{2x}}{\sinh\sqrt{2x}}\right)^{1/2} = 2^{3/4}x^{1/4}\sum_{n=0}^{\infty}\frac{D_n^2(0)}{n!}\exp[-(n+\tfrac{1}{2})\sqrt{2x}] \ ,$$

where the $D_n(x)$ are parabolic cylinder functions, and

$$D_n(0) = \frac{2^{n/2}\sqrt{\pi}}{\Gamma((1-n)/2)} \, .$$

- This is the basis for the Karhunen-Loeve expansion of the density function (see again Pycke (2007), *Statist. Probab. Letters*) and depends on the zeros of the parabolic cylinder function (c.f. the work on limiting distributions based on the integral of the absolute value of Brownian motion, which is based around the zeros of the Airy function)

Related work

Kac (1949) On the distribution of certain Wiener functionals. *TAMS* 65, 1-13.

Kac (1951) On some connections between probability theory and differential and integral equations. *Proc. Second Berkeley Symp. Math. Statist. Probab.*, 189-215.

Abadir (1995) The joint density of two functionals of Brownian motion. *Mathematical Methods of Statistics* 4, 449-462.

Abadir (1995) The limiting distribution of the t ratio under a unit root. *Econometric Theory* 11, 75-793.

Ghomrasni (2004) On distributions associated with the generalized Levy's stochastic area formula. *Studia Scientiarum Mathematicarum Hungarica* 41, 93-100.

Two points are worth mentioning here:-

- A different, more traditional inversion of this Laplace transform undertaken by Anderson and Darling (1952, *Ann. Math. Statist.* 23, 193-212) involved a modified Bessel function $K_{1/4}$.

- As we've just said, in the case of Brownian areas involving the integral of the absolute value (1-norm) of Brownian motion, Airy functions emerge in place of the parabolic cylinder functions (pertaining to the 2-norm case).

The following expression (Biane and Yor, p.76) may suggest that in any general theory the Macdonald functions will be fundamental:

$$-\left(\frac{Ai'}{Ai}\right)(x) = \frac{x^{1/2}}{2}\left(\frac{K_{2/3}}{K_{1/3}}\right)\left(\tfrac{2}{3}x^{3/2}\right).$$

They also have a role in exponential functionals of Brownian motion (Lyasoff, 2007, preprint).

Note that there are more general formulae available that could facilitate a better treatment of the initial condition in the unit root problem. One example is a special case of formula (2.k) in Pitman and Yor (1982) which gives

$$E\left[\exp\left(-\alpha\left(\tfrac{1}{2}B_1^2-\tfrac{1}{2}\right)-\tfrac{1}{2}\beta^2\int_0^1 B_s^2 ds\right)\right]$$

$$=\exp\left(\tfrac{1}{2}\alpha\right)\left(\cosh(\beta)+\frac{\alpha}{\beta}\sinh(\beta)\right)^{-1/2}$$

$$\times\exp-\frac{a^2\beta}{2}\left\{\frac{(\alpha/\beta)+\tanh(\beta)}{1+(\alpha/\beta)\tanh(\beta)}\right\},$$

where now the Brownian motion is started at a.

<u>Reference</u>

Pitman & Yor (1982) A decomposition of Bessel bridges. *Zeit. Wahrsch. Geb.*, 59, 425-457

This formula can be conveniently derived by the “stochastic process approach” in Chapter 4 of Tanaka (1996). See also M. Yor (1992) *Some Aspects of Brownian Motion*, ETH Zürich, Ch. 2.

Another direction is to consider explicitly the testing problem

$H_0: \rho = 1$

not against

$H_1: |\rho| < 1$

but against

$$H_1': \rho = \rho(c) = \exp(c/n) = 1 + \frac{c}{n} + o(\frac{1}{n}).$$

This is useful when considering the power of unit root test statistics.

Peter C.B. Phillips (1987, *Econometrica*) showed that if $J_c(t)$ is an Ornstein-Uhlenbeck process, satisfying the differential equation

$$d\,J_c(t) = \beta J_c(t)dt + dW(t), \quad J_c(0) = 0,$$

then

$$\left(\frac{1}{n}\sum_{t=1}^{n} Y_{t-1}(Y_t - \rho(c)Y_{t-1}), \frac{1}{n^2}\sum_{t=1}^{n} Y_{t-1}^2\right)$$

$$\xrightarrow{D} \left(\int_0^1 J_c(t)dW(t), \int_0^1 J_c(t)^2\,dt\right)$$

and we have

$$n(\hat{\rho} - \rho(c)) \xrightarrow{D} \frac{\int_0^1 J_c(t)dW(t)}{\int_0^1 J_c(t)^2\,dt}$$

and

$$t_{\rho(c)} \xrightarrow{D} \frac{\int_0^1 J_c(t)dW(t)}{\sqrt{\int_0^1 J_c(t)^2\,dt}}.$$

Then if $U = \frac{1}{2} J_c(1)^2$ and $V = \int_0^1 J_c(t)^2\,dt$, we have

$$E\left[\exp\left(-\tfrac{1}{2}\alpha U - \tfrac{1}{2}\beta^2 V \,\middle|\, J_C(0)=0\right)\right]$$

$$= \exp(-\tfrac{1}{2}c)\left(\cosh(\gamma) + \frac{\alpha - c}{\gamma}\sinh(\gamma)\right)^{-1/2},$$

where $\gamma = \sqrt{\beta^2 + c^2}$.

We have now expressed the Laplace transform in terms of a familiar problem in planar Brownian motion: the only differences are that the distributions are "shifted" by the term $\exp(-\frac{1}{2}c)$ and if we think in terms of $\left(\cosh(\gamma) + \frac{\alpha - c}{\gamma}\sinh(\gamma)\right)^{-\delta/2}$, we are "deconvolving" densities compared to the case $\delta = 2$ that pertains to Levy's stochastic area formula.

Some expressions for the density on part of its support in the case of a zero initial condition, and saddlepoint approximations to the density, are given by:-

Larsson (1995) The asymptotic distributions of some test statistics in near-integrated AR processes, *Econometric Theory* 11, 306-327.

This is probably the closest approach in the econometrics literature to one that would be based on planar Brownian motion.

An alternative approach I am working on is based on constructing parallels with the developments in distribution theory relating to the integral of the absolute value of Brownian motion via Mellin transform asymptotics.

Of course, this isn’t really an alternative approach: the ultimate objective is to marry both to construct a general theory that may be able even to encompass other functionals of Brownian motion. And work in representing densities in the following paper may be relevant in this general theory:-

Lyasoff (2007) Variational approach to the study of certain integral functionals of Brownian motion, preprint.

In econometrics we actually want more than this! What we really want is (a good approximation to) the finite-sample distribution of our test statistics. Note, however, that because the rate at which estimates approach their true parameter values is faster than in the stationary case, we may not require as many terms in any expansion as in the stationary case.

It is very likely that such an approximation will have its roots in the following paper:-

Götze (1979) Asymptotic expansions for bivariate von Mises functionals. *Z. Wahrsch. Ver. Geb.* 50, 333-355.

We now consider just the basic joint MGF (Laplace transform):

$$M_{X,Y}(-t_1,-t_2) = \exp(\tfrac{t_1}{2}) \times \left[\cosh(\sqrt{2t_2}) + \frac{t_1}{\sqrt{2t_2}}\sinh(\sqrt{2t_2})\right]^{-1/2}$$

Individual Laplace transforms are obtained by putting $t_1 = 0$ and $t_2 = 0$:

$M_X(-t_1) = \exp(\tfrac{t_1}{2}) \times (1+t_1)^{-1/2}$, and so X is $\frac{1}{2}(\chi^2(1) - 1)$

$M_Y(-t_2) = \dfrac{1}{\sqrt{\cosh(\sqrt{2t_2})}}$, the Laplace transform of $\int_0^1 B_s^2 ds$.

This gives the density of Y as

$$f(x) = \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \frac{(2n+\frac{1}{2})}{\sqrt{2\pi x^3}} \exp\left(-\frac{(2n+\frac{1}{2})^2}{2x}\right)$$

(see, e.g., Biane & Yor, 1987).

Remarks:

1. The above expression for the density of Y is not particularly tractable from the computational point of view!

2. The density can be given a probabilistic interpretation if it is written as

$$f(x) = \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \; f_{\tau(2n+\frac{1}{2})}(x)$$

where

$$f_{\tau(2n+\frac{1}{2})}(x) = \frac{(2n+\frac{1}{2})}{\sqrt{2\pi x^3}} \exp\left(-\frac{(2n+\frac{1}{2})^2}{2x} \right) 1_{\{x>0\}}$$

is the density of $\tau(c) = \inf\{s > 0 : B_s = c\}$, a first hitting time of Brownian motion.

3. The above derivation applies generally, meaning that the density is one in a class of conjugate densities. In particular, the density of the sum of Y and an independent copy of Y will have the same form and has density

$$f_{Y+\tilde{Y}}(x) = x^{-1}\sum_{n=1}^{\infty}\exp\left(-\pi^2(n-\tfrac{1}{2})^2 x/2\right),$$

which is a Jacobi theta function, whose Mellin transform is a Dirichlet beta series.

4. The moments of the Dickey-Fuller distribution are connected to a distribution that is a deconvolution of this theta function and so it isn't surprising that the moments involve functions like the Riemann zeta function and Dirichlet beta function. In fact, we'll see that the mean is the fractional derivative of a certain Hurwitz-Lerch zeta function.

For the purpose of deriving moments of the basic Dickey-Fuller distributions we use

$$E(\frac{X^k}{Y^b}) = \frac{1}{\Gamma(b)} \int_0^\infty M_{X,Y}^{(k,0)}(0,-t) t^{b-1} dt$$

where

$$M_{X,Y}^{(k,0)}(0,-t) = \frac{(-1)^k}{2^k \sqrt{\cosh\sqrt{2t}}} \sum_{i=0}^{k} (-1)^i (2i-1)!! \binom{k}{i} \left[\frac{\tanh(\sqrt{2t})}{\sqrt{2t}}\right]^i$$

and

$$(2i-1)!! = (2i-1)(2i-3)\ldots 1$$

Basic reference

Sawa (1972) Finite sample properties of the *k*-class estimators. *Econometrica* 40, 653-680.

The Ph.D. thesis of Bent Nielsen, currently Nuffield College and Oxford, contains a wealth of information that is extremely important.

The general expressions are

$$E\left(\frac{X}{Y}\right)^k = \frac{(-1)^k}{\Gamma(\frac{k}{2})2^{2k-1}} \sum_{j=1}^{k} (-1)^j (2j-1)!! \binom{k}{j} M_{j,k}$$

where

$$M_{j,k} = \int_0^\infty x^{2k} \left(\frac{\tanh x}{x}\right)^{k-j} \frac{1}{\sqrt{\cosh x}} dx,$$

and

$$E\left(\frac{X}{\sqrt{Y}}\right)^k = \frac{(-1)^k}{\Gamma(\frac{k}{2})2^{3k/2-1}} \sum_{j=1}^{k} (-1)^j (2j-1)!! \binom{k}{j} N_{j,k}$$

where

$$N_{j,k} = \int_0^\infty x^{k-1} \left(\frac{\tanh x}{x}\right)^{j} \frac{1}{\sqrt{\cosh x}} dx$$

The Mellin transforms satisfy a general quadratic recurrence relation that facilitates their evaluation, e.g. for $N_{j,k}$, we have

$$N_{j,k} = \frac{2(k-j-1)}{2j-1} N_{j-1,k-2} + \frac{2(j-1)}{2i-1} N_{j,k}$$

$$(0 < j < k,\ k > 2)$$

where $N_{j,k}$ is set equal to zero when $j < 0$.

- Following the Mellin transform approach, we invert the Mellin transforms to support an analysis of the densities and moments of the distributions in a way that directly compares with the methods designed for 1-norm of Brownian motion.

- For example

$$N_{k+1,k} = \int_0^\infty x^{k-1}\left(\frac{\tanh x}{x}\right)^k \frac{1}{\sqrt{\cosh x}}\,dx$$

$$= \frac{1}{\sqrt{2}}\sum_{j=0}^{k-1}\binom{k-1}{j}(-1)^j \int_0^1 B\left(j+\tfrac{1}{4}+x,\, k-j+\tfrac{1}{4}-x\right)\,dx$$

(k odd)

$$= \sqrt{2}\sum_{j=0}^{k-1}\binom{k-1}{j}(-1)^j \int_0^1 B_{1/2}\left(j+\tfrac{1}{4}+x,\, k-j+\tfrac{1}{4}-x\right)\,dx$$

(k even)

in the standard textbook treatment but we use Mellin inversion to get a simpler form.

- But for the purpose of today's talk, we will concentrate on deriving the moments of the relevant densities and focus relating our results to Pitman & Yor (2003) *Canadian J. Math.* 55, 292-330.

The basis of this paper is that there are infinitely divisible distributions on the half-line C_t, S_t and T_t with Laplace transforms

$$\left(\frac{1}{\cosh\sqrt{2\lambda}}\right)^t, \left(\frac{\sqrt{2\lambda}}{\sinh\sqrt{2\lambda}}\right)^t \text{ and } \left(\frac{\tanh\sqrt{2\lambda}}{\sqrt{2\lambda}}\right)^t.$$

Now we are interested in the relations between the processes C, S and T and $\hat{C}, \hat{S}$ and $\hat{T}$ characterized by

$$E[\exp(i\theta\hat{C}_t) = E[\exp(-\tfrac{1}{2}\theta^2 C_t) = \left(\frac{1}{\cosh\theta}\right)^t$$

$$E[\exp(i\theta\hat{S}_t) = E[\exp(-\tfrac{1}{2}\theta^2 S_t) = \left(\frac{\theta}{\sinh\theta}\right)^t$$

$$E[\exp(i\theta\hat{T}_t) = E[\exp(-\tfrac{1}{2}\theta^2 T_t) = \left(\frac{\tanh\theta}{\theta}\right)^t$$

The point is that since $\exp(-\frac{1}{2}\theta^2)$ is the characteristic function of a centred Gaussian r.v. N with variance 1, we see that $\hat{X}$ and X are related in the following way:

$$\hat{X}_t = N\sqrt{X_t}\,.$$

It transpires that for each moment of the Dickey-Fuller distribution, we only need to be able to evaluate integrals of the form:

$$\int_0^\infty \frac{x^{2n-1}}{\sqrt{\cosh x}}dx;$$

and for the distribution pertaining to the t-statistic, we additionally need to evaluate integrals of the form

$$\int_0^\infty \frac{x^{2n}}{\sqrt{\cosh x}}dx; \quad \text{and} \quad \int_0^\infty x^{n-1}\left(\frac{\tanh x}{x}\right)^n \frac{1}{\sqrt{\cosh x}}dx.$$

Dickey-Fuller distribution

$$m_1 = -\frac{1}{2}\int_0^\infty \left(\frac{x}{\sqrt{\cosh x}} \right) dx + 1$$

$$m_2 = \frac{1}{2}\int_0^\infty \left(\frac{x^3}{4\sqrt{\cosh x}} - \frac{3x}{2\sqrt{\cosh x}} \right) dx + \frac{1}{2}$$

$$m_3 = -\frac{1}{4}\int_0^\infty \left(\frac{x^5}{16\sqrt{\cosh x}} - \frac{9x^3}{8\sqrt{\cosh x}} + \frac{x}{\sqrt{\cosh x}} \right) dx + \frac{1}{4}$$

$$m_4 = \frac{1}{8}\int_0^\infty \left(\begin{array}{l} \dfrac{x^7}{96\sqrt{\cosh x}} - \dfrac{3x^5}{8\sqrt{\cosh x}} \\ \qquad\qquad + \dfrac{23x^3}{24\sqrt{\cosh x}} + \dfrac{3x}{\sqrt{\cosh x}} \end{array} \right) dx + \frac{1}{8}$$

t-type distribution

$$n_1 = -\frac{1}{\sqrt{2\pi}} \int_0^\infty \left(\frac{1}{\sqrt{\cosh x}} \right) \left(1 - \frac{\tanh x}{x} \right) dx$$

$$n_2 = \int_0^\infty \left(\frac{x}{4\sqrt{\cosh x}} \right) \left(1 + 3 \left(\frac{\tanh x}{x} \right)^2 \right) dx - 1$$

$$n_3 = -\frac{1}{\sqrt{2\pi}} \int_0^\infty \left(\frac{x^2}{4\sqrt{\cosh x}} \right) \left(1 - 15 \left(\frac{\tanh x}{x} \right)^3 \right) dx$$

$$n_4 = \int_0^\infty \left(\frac{x^3}{32\sqrt{\cosh x}} \right) \left(1 + 105 \left(\frac{\tanh x}{x} \right)^4 \right) dx - \int_0^\infty \left(\frac{x}{8\sqrt{\cosh x}} \right) dx - \frac{9}{4}$$

Concentrating on the Dickey-Fuller distribution, Gonzalo & Pitarakis (*International Economic Review*, 1998) showed that the mean (case n = 1) could be written in terms of the following series:

$$\sum_{n=0}^{\infty}(-1)^n \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)} \frac{1}{(n+\frac{1}{4})^2}$$

This is a slightly more complicated function than appeared in Pitman and Yor (2003) and we will need what Goyal and Laddha (1997) called the <u>unified zeta function</u>:

$$\Phi_{\mu}^{*}(z,s,a) = \sum_{n=0}^{\infty}(a+n)^{-s}(\mu)_m \frac{z^n}{n!}.$$

Well-known special cases include the Hurwitz-Lerch zeta function

$$\Phi(z,s,a)=\Phi_1^*(z,s,a)=\sum_{n=0}^{\infty}\frac{z^n}{(n+a)^s} \ ,$$

the Hurwitz zeta function

$$\zeta(s,a)=\Phi_1^*(1,s,a)=\sum_{n=0}^{\infty}\frac{1}{(n+a)^s} \ ,$$

and the Riemann zeta function

$$\zeta(s)=\Phi_1^*(1,s,1)=\sum_{n=0}^{\infty}\frac{1}{n^s} \ .$$

- Historically, there have been many approaches to summing these series including (or beginning with) the famous (Basler) problem solved by Euler of evaluating $\zeta(s)$ when $s = 2$.

- Many ingenious arguments have been offered although there are very few *constructive* proofs that systematically deliver sums of zeta series.

- One *systematic* approach is to write the zeta series as a generalized hypergeometric series and use (high-level) transformations to try to reduce the series to something that has a known sum.

GENERALIZED HYPERGEOMETRIC FUNCTIONS

The generalized hypergeometric function ${}_pF_q$ with p numerator and q denominator parameters is defined by

$$ {}_pF_q(\alpha_1,\ldots\alpha_p;\beta_1,\ldots,\beta_q;z) = {}_pF_q\left[\begin{matrix}\alpha_1,\ldots,\alpha_p \\ \beta_1,\ldots,\beta_q\end{matrix};z\right] $$

$$ = \sum_{n=0}^{\infty}\frac{(\alpha_1)_n\ldots(\alpha_p)_n}{(\beta_1)_n\ldots(\beta_p)_n}\frac{z^n}{n!} $$

+ **conditions**

Here, $(\lambda)_\kappa$ denotes the Pochhammer symbol representing the function given (in terms of the gamma function)

$$ (\lambda)_\kappa = \frac{\Gamma(\lambda+\kappa)}{\Gamma(\lambda)} = \begin{cases} 1 \\ \lambda(\lambda+1)\ldots(\lambda+n-1) \end{cases} $$

+ **conditions**

In the case of $\zeta(2)$,

$$\zeta(2) = \frac{4}{3}\,{}_3F_2(1,\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{2};\tfrac{3}{2};1)$$

and we can reduce this series using a summation theorem due to Dixon (or another due to Whipple) to obtain

$$\zeta(2) = \frac{4}{3}\left[\Gamma(\frac{3}{2})^3\right]\Gamma(\frac{1}{2}) = \frac{\pi^2}{6}$$

Another type of series, the series whose sum is called Catalan's constant, is

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2} = \frac{1}{4}\,{}_3F_2(1,\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{2};\tfrac{3}{2};-1)$$

This is an *alternating version* of the above series. Here, ${}_3F_2(1)$ transformation theorems are not applicable and indeed there is no known method of summing this series to express it in a closed form like $\zeta(2)$.

For the (integral involved in the) mean of the Dickey-Fuller distribution, the following result holds:

$$\int_0^\infty \frac{x\,dx}{\sqrt{\cosh x}} = 4\sqrt{2}\;{}_3F_2(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2};\tfrac{5}{4},\tfrac{5}{4};-1),$$

where ${}_3F_2(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2};\tfrac{5}{4},\tfrac{5}{4};-1) \approx 0.9833\ 84067\ 75370\ 96.$

This alternating series shares the same generic form as Catalan's constant:

$${}_3F_2(a,a,b;a+1,a+1;-1)$$

whose reduction (and those of higher-order than p = 2) were considered by Gottschalk and Maslen (1988).

- Unfortunately, the only reductions were derived in terms of integral *a*, although formulae were also offered for general *z* rather than z =1. Krupnikov and Kölbig (1997) show that closed forms *are available* for the functions for all orders *p* at *z* = 1 but their techniques do not extent to *z* = –1.

This points towards the *possibility* that the moments of the Dickey-Fuller distribution are themselves generic constants.

The next result shows that it is possible to transform the particular ${}_3F_2(-1)$ to a ${}_3F_2(1)$ hypergeometric series (although this transformation does not generalize to higher p):

LEMMA

$$ {}_3F_2(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2};\tfrac{5}{4},\tfrac{5}{4};-1)=\frac{1}{\sqrt{2}}\,{}_3F_2(\tfrac{1}{4},\tfrac{3}{4},1;\tfrac{5}{4},\tfrac{5}{4};1). $$

This opens up the possibility of using the established ${}_3F_2$ transformations. We want to use (high-level) transformation theorems to re-express the above hypergeometric series in a different and hopefully recognizable form.

The following identity, due to Thomae, when applied iteratively, captures contiguous relationships between ${}_3F_2(1)$ series and represents possible transformations of the series:

$$ {}_3F_2\left[\begin{matrix} a, b, c \\ d, e \end{matrix}; 1\right] = \frac{\Gamma(e)\Gamma(d+e-a-b-c)}{\Gamma(e-a)\Gamma(d+e-b-c)} \, {}_3F_2\left[\begin{matrix} a, d-b, d-c \\ d, d+e-b-c \end{matrix}; 1\right] $$

There are exactly 120 formal relations, although only 10 independent relations (once trivial symmetries are ignores), but fortunately a paper by Whipple allows the independent relations to be found in a reasonably straightforward manner.

THEOREM. *The independent Thomae relations corresponding to the mean of the Dickey-Fuller distribution are*

$$
\begin{aligned}
{}_3F_2\left(\tfrac{1}{4},\tfrac{3}{4},1;\tfrac{5}{4},\tfrac{5}{4};1\right) &= \frac{\Gamma^2(\frac{1}{4})}{4\sqrt{2\pi}}\,{}_3F_2\left(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2};\tfrac{3}{4},\tfrac{5}{4};1\right)\\
&= \frac{\Gamma^4(\frac{1}{4})}{16\pi^2}\,{}_3F_2\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{4},\tfrac{3}{2};1\right)\\
&= \frac{1}{2}\,{}_3F_2\left(\tfrac{1}{2},1,1;\tfrac{3}{4},\tfrac{5}{4};1\right)
\end{aligned}
$$

$${}_3F_2\left(\tfrac{1}{4},\tfrac{1}{4},\tfrac{1}{2};\tfrac{3}{4},\tfrac{5}{4};1\right) \approx 1.390725085638919$$

$${}_3F_2\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2};\tfrac{3}{4},\tfrac{3}{2};1\right) \approx 1.060781509820346$$

- None of these generalized hypergeometric series is summed in either Gradshteyn & Ryzhik (2007) or Prudnikov et al (1990), and the decimal expansions neither appear in Sloane's Online Encyclopaedia of Integer Sequences or are contained in Plouffe's Inverter.

- The series were computed with a high degree of accuracy

Higher-order moments:

THEOREM. For $n \in \mathbb{N}$,

$$\int_0^\infty \frac{x^n\,dx}{\sqrt{\cosh x}} = 2^{n+3/2}\, n!\; {}_{n+2}F_{n+1}\left(\frac{1}{4},\ldots,\frac{1}{4},\frac{1}{2};\frac{5}{4},\ldots,\frac{5}{4};-1\right)$$

But in fact we can show for $r \in \mathbb{R}$,

$$\int_0^\infty \frac{x^r\,dx}{\sqrt{\cosh x}} = 2^{-(r+1/2)}\frac{\Gamma(r+1)}{\sqrt{\pi}} \times \Phi^*_{1/2}(-1, r+1, \tfrac{1}{4})$$

This result has important implications from the point of view of analytic number theory because of the way that it will generalize Pitman and Yor's results for general t, e.g.

$$\int_0^\infty \frac{x^r\,dx}{(\cosh x)^2} = 2^{(1-r)}\left(1-2^{(1-r)}\right)\Gamma(r+1)\times\Phi^*_1(1,r,1)$$

$$= 2^{(1-r)}\left(1-2^{(1-r)}\right)\Gamma(r+1)\times\zeta(r)$$

and its interpolation at integral values:

$$\int_0^\infty \frac{x^n\,dx}{(\cosh x)^2} = 2^{(1-n)}\left(\frac{2^n-2}{2^n-1}\right) n!\; {}_{n+1}F_n\left(\frac{1}{2},\ldots,\frac{1}{2},1;\ \frac{3}{2},\ldots,\frac{3}{2};1\right)$$

Further representations:

- *Integral representation*

$$\Phi^{*}_{1/2}(-1, r+1, \frac{1}{4}) = \frac{1}{\Gamma(r+1)} \int_0^\infty t^{r+1} e^{-\frac{1}{4}t} (1+e^{-t})^{-1/4} dt .$$

- *Fractional derivative representation* – follows simply from the fact that the unified zeta function is a fractional derivative of the Hurwitz zeta function:

$$\Phi^{*}_{1/2}(-1, r+1, \frac{1}{4}) = \frac{1}{\sqrt{\pi}} D^{1/2}_{-1} \left\{ z^{-1/2} \Phi(z, r+1, \frac{1}{4}) \right\}.$$

- *Beuker's integrals via Zudilin's transformation* (Zudilin, 2004, J. London Math. Soc.)

Can we explain this?

We use Pitman and Yor's stochastic subordination approach:

Note that what we are really interested in is the Mellin transform $\int_0^\infty \frac{x^r\,dx}{\sqrt{\cosh x}}$.

1. We have

$$E[\exp(i\theta\hat{C}_t)] = E[\exp(-\tfrac{1}{2}\theta^2 C_t)] = \left(\frac{1}{\cosh\theta}\right)^t$$

and so

$$\hat{C}_t = N\sqrt{C_t}\,.$$

2. Now we think that for a <u>positive</u> random variable, its <u>negative</u> moments are given by

$$E[X^{-p}] = \frac{2^{1-p}}{\Gamma(p)}\int_0^\infty \theta^{2p-1}\phi_X(\tfrac{1}{2}\theta^2)\,d\theta,$$

where ϕ_X is the characteristic function of X.

3. In our Dickey-Fuller problem, the moments are a series in essentially absolute moments of a (positive) random variable whose <u>reciprocal</u> has characteristic function

$$\left(\frac{1}{\cosh\theta}\right)^{1/2}.$$

4. The random variable with this characteristic function is the GHS(1/2) distribution.

5. What we are saying is that:

$$\frac{1}{\sqrt{\int_0^1 W^2(r)dr}} = \frac{1}{\sqrt{C_{1/2}}} \equiv \frac{N}{\hat{C}_{1/2}} = \frac{N}{GHS(1/2)},$$

i.e. the reciprocal of the L^2-norm of Brownian motion is a normal scale mixture with mixing distribution given by the GHS(1/2) distribution.

6. We can find the characteristic function of the following function:

$$\int_{-\infty}^{\infty} e^{i\lambda x}\left(\frac{1}{\cosh x}\right)^{1/2} dx = \frac{1}{\sqrt{2\pi}}\left|\Gamma\left(\frac{1}{4}+\frac{i\lambda}{2}\right)\right|^2 .$$

but the Mellin transfom is more complicated.

7. This Fourier transform generalizes to many L-functions and other functions, *via*

$$\int_{-\infty}^{\infty} e^{i\lambda x}\left(\frac{1}{\cosh x}\right)^{t} dx = \frac{2^{t-1}}{\Gamma(t)}\left|\Gamma\left(\frac{t+i\lambda}{2}\right)\right|^2 .$$

for example the Dirichlet beta function (case $t=1$) and Riemann zeta function (case $t=2$).

One final result, again derived from work by Marc Yor, reveals the extreme-value nature of the Dickey-Fuller distribution.

If X_1 and X_2 are i.i.d. Gompertz-type r.v.'s with density function

$$f(x) = \frac{2}{\Gamma(\frac{1}{4})} \exp\left(\frac{1}{2}x - \exp(2x)\right),$$

then

$$\phi_{X_1 - X_2}(t) = \phi_{X_1}(t)\phi_{X_2}(-t) = \frac{\left|\Gamma(\frac{1}{4} + i\frac{t}{2})\right|^2}{\Gamma(\frac{1}{4})^2}.$$

Then:

$$\phi_{X_1 - X_2}(t) = const \times E\left[\frac{1}{\sqrt{\int_0^1 B_s^2}} \exp -\frac{t^2}{2}\left(\frac{1}{\sqrt{\int_0^1 B_s^2}}\right)^2\right].$$

Concluding remarks

- We have presented the unit root distribution problem in a wider context than usual, that has potential application across Mathematical Finance, Econometrics and Combinatorics

- In Mathematical Finance – the pricing of Asian options, which can naturally be related to problems involving exponential functionals of Brownian motion (Yor, Dufresne, Schröder, Lyasoff)

- In Econometrics – the unit root problem, which involves trying to characterize the (properties of) the density of the OLS estimator of the parameter in the AR(1) model under the null hypothesis that it is unity

- This work involves a quadratic functional (the 2-norm) of Brownian motion

- In combinatorics – work on limiting distributions based on the integral of the absolute value (the 1-norm) of Brownian motion, e.g. the area below a lattice path, the path length in trees

Can we construct a universal theory?