**Summer 2016: Stochastic Processes and Orthogonal Polynomials**

**Key ideas**

2. For certain families of orthogonal polynomials, we can also have hypergeometric function representation, differential/difference equation, a Rodrigues-type formula, generating functions, and limit relations.
3. The polynomials corresponding to the “standard” distributions come from an equation, differential or finite difference, of a hypergeometric type. The rest of the Askey scheme is filled with certain terminating hypergeometric series.
4. A generator of a Markov process can lead to both an ordinary differential equation and a three-term-type recurrence, which, with some luck, can be studied using orthogonal polynomials. In particular, there are random walk polynomials and birth-death polynomials, as well as the Laguerre diffusion (known in finance as the CIR process of the Cox-Ingersoll-Ross interest rate model) and the Jacobi diffusion.
5. Lévy-Scheffer sequences of polynomials are a gateway to stochastic integration with respect to many “fancy” processes. The most valuable seem to be Lévy-Meixner polynomials, which are also orthogonal with respect to the distribution of the corresponding Lévy process.
6. CRP and PRP (chaotic and predictable representation properties) are good to have. It is even better when there are concrete numerical procedures for computing the corresponding representations. Malliavin calculus provides all the answers in the Gaussian case. Generalized polynomial chaos (gPC) is a popular engineering alternative to the multiple integral chaotic representation in a more general setting.

**Minor points**

1. Hypergeometric series can represent just about anything, and can be useful even with a zero radius of convergence.
2. Pochhammer symbol is used in the definition of the hypergeometric series; Pochhammer contour is often the key to the analytic continuation (as well as to the solution of some puzzles).
3. Strong Feller property and irreducibility usually imply existence and uniqueness of the invariant measure [and even more than that...]
4. Lévy process is a process with independent and stationary increments (plus start at zero, stochastic continuity, and cadlag [cadlag, or RCLL] trajectories). The distribution of the increments must be infinitely divisible and often lends the name to the corresponding process.
(5) The Faà di Bruno formula is an expression for the \(n\)-th derivative of a composition of two functions in terms of the derivatives of the functions involved.

(6) Sheffer sequences and umbral calculus are a systematic way to study generating functions [as well as numerous combinatorial identities], and a good way to understand many special families of polynomials.

(7) Stochastic integration is well-defined if and only if the integrator is a semi-martingale (Bichteler-Dellacherie Theorem). There is a general form of the Itô formula for semi-martingales.

(8) Teugels martingales [so named after a Belgian mathematician J. L. Teugels by one of his students] come from compensating the sum of the powers of jumps of a Lévy process and help to establish both CRP and PRP for a large class of those processes. They are also examples of pure jump normal martingales [Wiener process is the only continuous one...]. Incidentally, teugels in Dutch is the same as martingale [but now meaning a particular piece of the horse harness] in English.

(9) Malliavin calculus provides a probabilistic tool to study random processes; Stein’s method provides a probabilistic tool to study limit theorems. Both are somewhat well-developed in Gaussian and Poisson cases, but remain mostly open in the case of other distributions.

Problems

1. Come up with a convincing argument that the polynomials are always dense in \(L_2(\mathbb{R}; \mu)\) if \(\mu\) has compact support.

2. Investigate the connections among the following properties:

   (1) The polynomials are dense in \(L_2(\mathbb{R}; \mu)\).
   (2) The measure \(\mu\) is uniquely determined by its moments.
   (3) \(\int_\mathbb{R} e^{a|x|} \mu(dx) < \infty\) for some \(a > 0\).

   For each pair, either prove that one implies the other or construct a counterexample.

3. Consider the Stieltjes-Wigert weight function

\[ \varphi(x) = \frac{1}{\sqrt{\pi}} e^{-\ln^2 x}, \quad x > 0. \]

   (a) Confirm that

   \[ s_n = \int_0^{+\infty} x^n \varphi(x) dx = e^{(n+1)^2/4}, \quad n = 0, 1, 2, \ldots. \]

   (b) Confirm that

   \[ \int_0^{+\infty} x^n \sin(2\pi \ln x) \varphi(x) dx = 0 \]
for every $n = 0, 1, 2, \ldots$ and so the polynomials are not complete in $L_2((0, +\infty), \varphi(x)dx)$.

(c) Confirm that the integral
\[ \int_0^\infty e^{ax} \varphi(x)dx \]
diverges for every $a > 0$, but the integral
\[ \int_0^\infty \frac{\ln \varphi(x)}{1 + x^2} dx \]
and the series
\[ \sum_{k \geq 1} \frac{1}{\sqrt{s_k}} \]
both converge.

(d) Comment on connections between the above results and Problem 2.

4. Confirm that the ordinary differential equation
\[ (\sigma_0 + \sigma_1 x + \sigma_2 x^2)y''(x) + (\tau_0 + \tau_1 x)y'(x) = (n\tau_1 + n(n-1)\sigma_2)y(x) \]
has a solution that is a polynomial of degree $n$. Work out the details in the following cases:

- $\sigma_0 = 1, \sigma_1 = 0, \sigma_2 = -1, \tau_0 = \beta - \alpha, \tau_1 = -(\alpha + \beta + 2)$ [this leads to Jacobi polynomials];
- $\sigma_0 = \sigma_1 = 0, \sigma_2 = 1, \tau_0 = \beta, \tau_1 = (2 - \alpha)$ [this leads to the Bessel polynomials];
- $\sigma_0 = \sigma_1 = 0, \sigma_2 = 1/n, \tau_0 = 0, \tau_1 = -(n-1)/n$ [this leads to the Romanovskii polynomials, coming from Student’s $t_n$ distribution];
- $\sigma_0 = \sigma_2 = 0, \sigma_1 = 1, \tau_0 = \alpha + 1, \tau_1 = -1$ [this leads to Laguerre polynomials];
- $\sigma_0 = 1, \sigma_1 = \sigma_2 = 0, \tau_0 = 0, \tau_1 = -2$ [this leads to Hermite polynomials].

5. Confirm that the finite difference equation
\[ (\sigma_0 + \sigma_1 x + \sigma_2 x^2)(y(x+1) - 2y(x) + y(x-1)) + (\tau_0 + \tau_1 x)(y(x+1) - y(x)) \]
\[ = (n\tau_1 + n(n-1)\sigma_2)y(x) \]
has a solution $y = y(x)$ that is a polynomial in $x$ of degree $n$.

6. If a random variable $X$ has cumulative distribution function $F = F(x)$, then, for $a > 0$, the random variable $Y = aX + b$ has cumulative distribution function $F_Y(x) = F_X((x - b)/a)$. Derive the corresponding relation between the orthogonal polynomials for $F_X$ and $F_Y$.

7. Confirm self-duality of the (suitably normalized) Kravchuk, Meixner, and Charlier polynomials. What (if anything) does self-duality mean for the corresponding distributions?
8. Derive an asymptotic expansion for the tail of the standard Gaussian distribution and write it as a suitable hypergeometric series.

9. Consider the Itô diffusion
\[ dX(t) = b(X(t))dt + \sqrt{2}\sigma(X(t))dW(t) \]
such that \( X(t) \in E \subset \mathbb{R} \) and
\[ \int_{E} \frac{1}{\sigma^2(x)} \exp \left( \int \frac{b(x)}{\sigma^2(x)} dx \right) dx = A < \infty. \]
Confirm that the function
\[ \rho(x) = \frac{1}{A\sigma^2(x)} \exp \left( \int \frac{b(x)}{\sigma^2(x)} dx \right) \]
is the density of an invariant measure for \( X \) and that the generator of \( X \) is symmetric on \( L_2(E; \rho(x)dx) \). Is this invariant measure unique?

10. By considering the process \( Y(t) = 1/X^2(t) \), find a closed-form solution of
\[ dX(t) = (aX(t) + bX^3(t))dt + \sigma X(t)dW(t), \quad X(0) = y, \]
with real numbers \( a, b, \sigma \), and determine the conditions on \( a, b, \sigma \) for the existence of an invariant measure. [With \( \sigma = \sqrt{2} \) (little loss of generality), it follows from problem 9 that invariant measure exists if \( a > 1 \) and \( b < 0 \).]

11. Recall that the Sheffer sequence of polynomials \( \{Q_k(x), \ k \geq 0\} \) is the coefficients of Taylor expansion at \( z = 0 \) of the function \( f(z) \exp(xu(z)) \), where \( f \) and \( u \) are analytic at \( z = 0 \) and \( f(0) \neq 0, u(0) = 0, u'(0) \neq 0 \):
\[ f(z)e^{xu(z)} = \sum_{k=0}^{\infty} \frac{Q_k(x)}{k!}z^k. \]
Confirm that
(a) If \( u(z) = z \), then we get the Appell sequence, that is,
\[ Q'_k(x) = kQ_{k-1}(x), \quad k \geq 1. \]
(b) If furthermore \( f(z) = 1/Ee^{\xi} \) for some random variable \( \xi \), then, in addition to (1), \( \mathbb{E}Q_k(\xi) = 0, k \geq 1 \), that is, \( Q_k \) are the Wick polynomials for \( \xi \).

11. One version of the Charlier polynomials \( c_n(x; t) \) can be defined from the relation
\[ e^{-t}(1+z)^x = \sum_{n=0}^{\infty} \frac{c_n(x; t)}{n!}z^n. \]
Let \( N = N_t \) be the Poisson process with unit intensity and let \( \tilde{N}_t = N_t - t \). For \( n = 1, 2, 3 \), confirm by direct computation that \( c_n(N_t; t)/n! \) is a martingale and is an
$n$ times iterated integral with respect to $\tilde{N}$. For example,
\[ c_2(N; t) = 2 \int_0^t \int_0^s d\tilde{N}_r d\tilde{N}_s. \]

Questions

1. What is the connection between the moment problem and continued fractions?
2. What happens to the conclusion of the Favard theorem if we omit the condition $A_{n-1}A_n C_n > 0$?
3. What can we say about uniqueness of the measure in the Favard theorem?
4. When does the three-term recurrence relation terminate?
5. We know that the roots of two consecutive orthogonal polynomials interlace. **Sturm’s separation theorem** also has interlacing of zeroes, this time for two linearly independent solutions of a linear second-order ODE. Can we make further connections between orthogonal polynomials and ODEs in this direction?
6. What is the difference between the moment problems of Hausdorff, Stieltjes, and Hamburger?
7. What is the Bieberbach conjecture and how does the proof use orthogonal polynomials?
8. What families of orthogonal polynomials are not in the Askey scheme and why? Do those families appear in any applications?
9. What other measures (beside Stieltjes-Wigert) lead to an $L_2$ space where polynomials are not complete?
10. What are Racah coefficients?
11. Every limit relation in the Askey scheme should correspond to a limit theorem in probability. For example, convergence of Krawtchouk to Hermite corresponds to CLT for Binomial. State the remaining limiting relations in probabilistic terms and provide probabilistic proofs.
12. Consider the ODE
\[ \sigma(x)y''(x) + \tau(x)y'(x) + c(x)y(x) = 0, \]
where $\sigma$ and $\tau$ are polynomials. Why is it natural to consider only the case when $\sigma$ has degree at most two and $\tau$ has degree at most one? [A possible lead: Fuchs-Frobenius theory].
13. What makes Bessel and Romanovski polynomials different from most other families of polynomials we consider? Can you include them in the Askey scheme?
14. In problem 4, we see that there is a natural classification of the polynomials coming out of the **differential equation**, depending on the number of real zeros of the leading polynomial coefficient. What would a similar classification be for the finite difference equation in problem 5?
15. We know that Bessel’s functions come from the solution of Bessel’s differential equation. Can you see other connections between Bessel’s functions, Bessel’s differential equation, Bessel’s polynomials, and Bessel’s process?
16. What can we say about the (random) orthogonal polynomials coming from the empirical distribution function? For example, is there a way to state the
Glivenko-Cantelli theorem in the language of the corresponding orthogonal polynomials? How about the Kolmogorov-Smirnov test?

(17) Charlier polynomials can be expressed in terms of Laguerre polynomials. On the other hand, we know that the corresponding distributions (Poisson and Gamma) are related as conjugate priors. Is it a coincidence or is there more to it? Namely, if two distributions are related as conjugate priors, is there a particular relation between the corresponding orthogonal polynomials, and conversely, can we use the Askey scheme to discover more conjugate priors?

(18) What kind of orthogonal polynomials correspond to singular distributions, for example, Cantor’s?

(19) How will the forward Kolmogorov equation look if the process has a unique invariant measure and we use this measure to compute the transition density and the corresponding adjoint of the generator?
Glossary of terms

**Hypergeometric function** is an analytic extension of the corresponding hypergeometric series.

**Hypergeometric series** is a power series such that the ratio of two consecutive coefficients is (the same) rational function of \( n \); as a result, the coefficients are represented using the Pochhammer symbol.

**Pochhammer symbol** is \((x)_n = x(x+1)\cdots(x+n-1)\).

**CRP**, or **chaotic representation property**, is the possibility to write a random variable as a sum of iterated stochastic integrals of deterministic functions.

**PPR**, or **predictable representation property**, is the possibility to write a random variable as a single stochastic integral.

**Stroock’s formula** is a way to compute the chaotic representation in the Gaussian case.

**Clark-Ocone formula** is a way to compute the predictable representation in the Gaussian case.

**Cameron-Martin expansion** is an alternative form of chaotic representation in the Gaussian case.

**gPC**, or **generalized polynomial chaos**, is an extension of the Cameron-Martin expansion to non-Gaussian case.

**Askey scheme** is a particular way to organize several families of orthogonal polynomials indicating various connections between the families via suitable limiting procedures.

**Sheffer sequence of polynomials** is the coefficients of a certain Taylor expansion.

**Appel sequence of polynomials** is a particular case of the Sheffer sequence, when the derivative of every polynomial gives the previous one.

**Wick polynomials** for a given random variable are a particular case of Appel polynomials, when the expected value of all non-constant polynomials of that random variable is zero.

**Wick product** is a way to multiply Wick polynomials so that the product of \( n \)-th polynomial and \( k \)-th polynomial is \( n + k \)-th polynomial.

**Lévy-Sheffer sequence of polynomials** is a particular case of the Sheffer sequence when the corresponding generating function involves the moment generating function of an infinitely divisible distribution.

**Lévy-Meixner polynomials** are orthogonal polynomials that also form a Lévy-Sheffer sequence.

---

\(^1\)A random selection in random order
**Faà di Bruno formula** is an expression for the \( n \)-th derivative of a composition of two functions in terms of the derivatives of the functions involved.

**Leibnitz formula** is an expression for the \( n \)-th derivative of the product of two functions in terms of the derivatives of the functions involved.

**Doléans-Dade exponential** is the generalization of the geometric Brownian motion to semi-martingales, named after the French-American mathematician Catherine Doléans-Dade.

**Bichteler-Dellacherie Theorem** states that stochastic integration is well-defined only with respect to a semi-martingale.

**Normal martingale** is a (square-integrable local) martingale \( M \) such that \( M^2(t) - ct, c > 0 \), is also a (local) martingale.

**Lévy characterization** of the Brownian motion is the statement that the Brownian motion is the only continuous normal martingale (\( c = 1 \) for the standard Brownian motion).

**Lévy process** is a process with independent and stationary increments.

**Sturm-Liouville theory** is about eigenvalue problem for linear second-order ODEs.

**Fuchs-Frobenius theory** is about power series method for linear second-order ODEs.

**Favard theorem** is characterization of orthogonal polynomials by a three term recursion.