IEOR 6711: Stochastic Models I

Professor Whitt, Tuesday, September 11

Our Friends: Transforms

1. Our Friends

Identify the following:

- (a) generating function (of a sequence)
- (b) probability generating function (of a probability distribution or of a random variable)
- (c) exponential generating function (of a sequence)
- (d) z transform
- (e) moment generating function (of a probability distribution or of a random variable)
- (f) characteristic function (of a probability distribution or of a random variable)
- (g) Fourier transform (of a function)
- (h) Laplace transform (of a function)

Answers:

(a) generating function

The generating function of the sequence $\{a_n : n \ge 0\}$ is

$$\hat{a}(z) \equiv \sum_{n=0}^{\infty} a_n z^n \ ,$$

which is defined where it converges. There exists (theorem) a radius of convergence R such that the series converges absolutely for all z with |z| < R and diverges if |z| > R. The series is uniformly convergent on sets of the form $\{z : |z| < R'\}$, where 0 < R' < R.

References: (1) Wilf, Generatingfunctionology, (2) Feller, I, Chapter 11, (3) Grimmett and Stirzaker, Chapter 5. The book, An Introduction to Stochastic Processes, by E. P. C. Kao, 1997, Duxbury, provides introductory discussions of both generating functions and Laplace transforms, including numerical inversion algorithms using MATLAB (based on the 1992 and 1995 Abate and Whitt papers).

(b) probability generating function

Given a random variable X with a probability mass function (pmf)

$$p_n \equiv P(X=n) ,$$

the probability generating function of X (really of its probability distribution) is the generating function of the pmf, i.e.,

$$\hat{P}(z) \equiv E[z^X] \equiv \sum_{n=0}^{\infty} p_n z^n .$$

(c) exponential generating function

The exponential generating function of the sequence $\{a_n : n \geq 0\}$ is

$$\hat{a}_{exp}(z) \equiv \sum_{n=0}^{\infty} \frac{a_n z^n}{n!} ,$$

which is defined where it converges; see p. 36 of Wilf.

(d) z transform

A z transform is just another name for a generating function; e.g., see Section I.2, p. 327, of Appendix I of Kleinrock (1975), Queueing Systems.

(e) moment generating function

Given a random variable X the moment generating function of X (really of its probability distribution) is

$$\psi_X(t) \equiv E[e^{tX}] \ .$$

The random variable could have a continuous distribution or a discrete distribution; e.g., see p. 15 of Ross. It can be defined on the entire real line.

Discrete case: Given a random variable X with a probability mass function (pmf)

$$p_n \equiv P(X=n), \quad n > 0,$$

the moment generating function (mgf) of X (really of its probability distribution) is the generating function of the pmf, where e^t plays the role of z, i.e.,

$$\psi_X(t) \equiv E[e^{tX}] \equiv \hat{p}(e^t) \equiv \sum_{n=0}^{\infty} p_n e^{tn}$$
.

Continuous case: Given a random variable X with a probability density function (pdf) $f \equiv f_X$ on the entire real line, the moment generating function (mgf) of X (really of its probability distribution) is

$$\psi(t) \equiv \psi_X(t) \equiv E[e^{tX}] \equiv \int_{-\infty}^{\infty} f(x)e^{tx} dx$$
.

A major difficulty with the mgf is that it may be infinite or it may not be defined. For example, if X has a pdf $f(x) = A/(1+x)^p$, x > 0, then the mgf is infinite for all t > 0.

(f) characteristic function

The characteristic function (cf) is the mgf with an extra imaginary number $i \equiv \sqrt{-1}$:

$$\phi(t) \equiv \phi_X(t) \equiv E[e^{itX}]$$
.

where $i \equiv \sqrt{-1}$; see p. 17 of Ross, Chapter 6 of Chung, Chapter XV of Feller II and Lukacs. The random variable could have a continuous distribution or a discrete distribution.

Unlike mgf's, every probability distribution has a well-defined cf. To see why, recall that e^{it} is very different from e^t . In particular,

$$e^{itx} = \cos(tx) + i\sin(tx) .$$

(g) Fourier transform

A Fourier transform is just a minor variant of the characteristic function. Really, it should be said the other way around, because the Fourier transform is the more general notion. There are a few different versions, all differing from each other in minor unimportant ways. Under regularity conditions, a function f has Fourier transform

$$\tilde{f}(y) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy} dx .$$

Again under regularity conditions, the original function f can be recovered from the inversion integral

$$f(x) = \int_{-\infty}^{\infty} \tilde{f}(y)e^{2\pi ixy} dy .$$

For example, see D. C. Champeney, A Handbook of Fourier Theorems, Cambridge University Press, 1987.

(h) Laplace transform

Given a real-valued function f defined on the positive half line $\mathbb{R}^+ \equiv [0, \infty)$, its Laplace transform is

$$\hat{f}(s) \equiv \int_{0}^{\infty} e^{-sx} f(x) dx,$$

where s is a complex variable with positive real part, i.e., s = u + iv with $i = \sqrt{-1}$, u and v real numbers and u > 0. see Doetsch (1961, 1974) and Chapter XIII of Feller II.

The Laplace transform takes advantage of the fact that X is nonnegative (the density concentrates on the positive half line), so that it can damp the function. Recall that

$$e^{-sx} = e^{-(u+iv)x} = e^{-ux}e^{ivx} = e^{-ux}[\cos(vx) + i\sin(vx)], \text{ for } x \ge 0.$$

2. Generating Functions

A generating function is a clothesline on which we hang up a sequence of numbers for display (Wilf, Generating function ology).

Let $\{a_n : n \ge 0\}$ be a sequence of real numbers. Then

$$\hat{a}(z) \equiv \sum_{n=0}^{\infty} a_n z^n$$

is the generating function of the sequence $\{a_n : n \geq 0\}$.

Questions

- (a) At least in principle, how can we get the n^{th} term a_n back from the generating function $\hat{a}(z)$?
- (b) Let the sequence $\{c_n\} \equiv \{c_n : n \geq 0\}$ be the *convolution* of the two sequences $\{a_n\} \equiv \{a_n : n \geq 0\}$ and $\{b_n\} \equiv \{b_n : n \geq 0\}$; i.e.,

$$c_n \equiv \sum_{i=0}^n a_i b_{n-i}, \quad n \ge 0 .$$

Find the generating function of $\{c_n\}$ in terms of the generating functions of $\{a_n\}$ and $\{b_n\}$

- (c) Let X_i have a Poisson distribution with mean λ_i for i = 1, 2. Let X_1 and X_2 be independent. Use generating functions to find the distribution of $X_1 + X_2$.
- (d) Let $X_1, X_2, ...$ be IID random variables on the nonnegative integers each with probability generating function (pgf) $\hat{P}_X(z)$. Let N be a random variable on the nonnegative integers with pgf $\hat{P}_N(z)$ that is independent of $X_1, X_2, ...$ What is the generating function of $X_1 + ... + X_N$?
- (e) A hen lays N eggs, where N has a Poisson distribution with mean λ . Suppose that the eggs hatch independently. Suppose that each egg hatches with probability p. Use generating functions to find the distribution of the total number of chicks hatched from these eggs.
 - (f) Use generating functions to prove the identity

$$\left(\begin{array}{c} 2n\\ n \end{array}\right) = \sum_{j=0}^{j=n} \left(\begin{array}{c} n\\ j \end{array}\right)^2$$

(g) Consider the recurrence

$$a_{n+1} = 2a_n + n$$
, for $n \ge 0$ with $a_0 = 1$.

What is a_n ?

(h) Consider the recurrence

$$a_{n+1} = a_n + a_{n-1}$$
, for $n \ge 0$, with $a_0 = 1$ and $a_1 = 1$.

Find the generating function of the sequence $\{a_n : n \ge 0\}$. Use generating functions to find a good approximation for a_n for large n.

Answers

(a) At least in principle, how can we get the n^{th} term a_n back from the generating function $\hat{a}(z)$?

Look at the n^{th} derivative evaluated at zero:

$$a_n = n! \hat{a}^{(n)}(0)$$
, where $\hat{a}^{(n)}(0) \equiv \frac{d^n}{dz^n} \hat{a}(z)|_{z=0}$

(b) Let the sequence $\{c_n\} \equiv \{c_n : n \geq 0\}$ be the *convolution* of the two sequences $\{a_n\} \equiv \{a_n : n \geq 0\}$ and $\{b_n\} \equiv \{b_n : n \geq 0\}$; i.e.,

$$c_n \equiv \sum_{i=0}^n a_i b_{n-i}, \quad n \ge 0 .$$

Find the generating function of $\{c_n\}$ in terms of the generating functions of $\{a_n\}$ and $\{b_n\}$

 $\hat{c}(z) \equiv \sum_{n=0}^{\infty} c_n z^n$ $= \sum_{n=0}^{\infty} [\sum_{i=0}^n a_i b_{n-i}] z^n$ $= \sum_{i=0}^{\infty} [\sum_{n=i}^{\infty} a_i b_{n-i}] z^n \quad \text{changing the order of the sums}$ $= \sum_{i=0}^{\infty} a^i z^i \sum_{n=i}^{\infty} b_{n-i} z^{n-i}$ $= \sum_{i=0}^{\infty} a^i z^i \sum_{n=0}^{\infty} b_n z^n$ $= \hat{a}(z) \hat{b}(z) .$

Recall that we need conditions to justify the interchange. It suffices to have the two series $\sum_{i=0}^{\infty} |a^i| z^i$ and $\sum_{i=0}^{\infty} |a^i| z^i$ be finite for some z > 0, by Fubini.

This problem shows that the generating function of a convolution of two sequences is the product of the generating functions of the component sequences.

⁽c) Let X_i have a Poisson distribution with mean λ_i for i = 1, 2. Let X_1 and X_2 be independent. Use generating functions to find the distribution of $X_1 + X_2$.

First step: The distribution of $X_1 + X_2$ is a convolution of the distributions of X_1 and X_2 :

$$P(X_1 + X_2 = n) = \sum_{i=0}^{n} P(X_1 = i)P(X_2 = n - i), \quad n \ge 0.$$

Next we see that the generating function of X_i (actually of the probability distribution of X_i , say $p_n \equiv P(X_i = n)$) is

$$\hat{P}(z) \equiv \hat{P}_{X_i}(z) \equiv E[z^{X_i}] \equiv \sum_{n=0}^{\infty} p_n z^n$$

$$= \sum_{n=0}^{\infty} \frac{e^{-\lambda_i} \lambda_i^n}{n!} z^n$$

$$= e^{-\lambda_i} \sum_{n=0}^{\infty} \frac{\lambda_i^n z^n}{n!}$$

$$= e^{-\lambda_i} e^{\lambda_i z}$$

$$= e^{\lambda_i (z-1)}.$$

Hence,

$$\sum_{n=0}^{\infty} P(X_1 + X_2) z^n = \hat{P}_{X_1}(z) \hat{P}_{X_2}(z) = e^{(\lambda_1 + \lambda_2)(z-1)} ,$$

so that $X_1 + X_2$ is again Poisson distributed, but with mean $\lambda_1 + \lambda_2$.

(d) Let X_1, X_2, \ldots be IID random variables on the nonnegative integers each with probability generating function (pgf) $\hat{P}_X(z)$. Let N be a random variable on the nonnegative integers with pgf $\hat{P}_N(z)$. that is independent of X_1, X_2, \ldots What is the pgf of $Z \equiv X_1 + \ldots + X_N$?

$$\hat{P}_{Z}(z) \equiv \sum_{n=0}^{\infty} P(Z=n)z^{n}
= \sum_{n=0}^{\infty} P(X_{1} + \dots + X_{N} = n)z^{n}
= \sum_{n=0}^{\infty} [\sum_{k=0}^{\infty} P(X_{1} + \dots + X_{N} = n | N = k)P(N = k)]z^{n}
= \sum_{k=0}^{\infty} P(N = k) \sum_{n=0}^{\infty} P(X_{1} + \dots + X_{k} = n)z^{n}
= \sum_{k=0}^{\infty} P(N = k) \hat{P}_{X}(z)^{k}
= \hat{P}_{N}(\hat{P}_{X}(z)) .$$

Alternatively, in a more streamlined way,

$$E[z^{X_1 + \dots + X_N}] = E(E[z^{X_1 + \dots + X_N} | N])$$

$$= E(E[z^X]^N)$$

$$= \hat{P}_N(\hat{P}_X(z)) .$$

(e) A hen lays N eggs, where N has a Poisson distribution with mean λ . Suppose that the eggs hatch independently. Suppose that each egg hatches with probability p. Use probability generating functions (pgf's) to find the distribution of the number of chicks.

We can apply part (d), because the total number of chicks is the random sum $Z = X_1 + \cdots + X_N$. The random variables X_i are the outcome of the i^{th} egg, i.e., $X_i = 1 = 1 - P(X_i = 0) = p$. Then the pgf of X_i is

$$\hat{P}_X(z) = pz + (1-p) .$$

Since N is the number of eggs laid, the pgf of N is

$$\hat{P}_N(z) = e^{\lambda(z-1)} ;$$

see Part (c) above. Thus

$$\hat{P}_Z(z) = \hat{P}_N(\hat{P}_X(z)) = e^{\lambda([pz+(1-p)]-1)} = e^{\lambda p(z-1)}$$

which is the pgf of the Poisson distribution with mean λp .

(f) Use generating functions to prove the identity

$$\binom{2n}{n} = \sum_{i=0}^{j=n} \binom{n}{j}^2$$

Note that the righthand side can be written as a convolution, because

$$a_j = a_{n-j}$$

where

$$a_j = \left(\begin{array}{c} n \\ j \end{array}\right) .$$

Moreover, using the binomial form,

$$\hat{a}(z) \equiv \sum_{j=0}^{j=n} a_j z^j = (1+z)^n .$$

Multiplying the two generating functions, we get $(1+z)^{2n}$ as the generating function, which is the generating function of

$$\begin{pmatrix} 2n \\ k \end{pmatrix}$$
;

i.e.,

$$(1+z)^{2n} = \sum_{k=0}^{k=2n} z^k \binom{2n}{k}$$

Then c_n is the coefficient of z^n in $(1+z)^{2n}$. So we see that the relationship is just convolution for the specific sequences.

A few references on combinatorics:

- J. Riordan, Introduction to Combinatorial Analysis, 1958, Wiley.
- J. Riordan, Combinatorial Identities, 1968, Wiley.

Chapter II in W. Feller, volume I.

Richard A. Brualdi, Introductory Combinatorics, 3rd Edition

(g) Consider the recurrence

$$a_{n+1} = 2a_n + n$$
, for $n \ge 0$ with $a_0 = 1$.

What is a_n ?

This comes from Section 1.2 of Wilf. If we start iterating, then we see that the successive

This comes from Section 1.2 of Will. If we start iterating, then we see that the successive terms are $1, 2, 5, 12, 27, 58, 121, \ldots$ It becomes hard to see the general form. We can multiply both sides by z^n and add to get expressions involving the generating function $\hat{a}(z)$. In particular, for the left side we get

$$a_1 + a_2 z + a_3 z^2 + a_4 z^3 + \dots = \frac{(\hat{a}(z) - a_0)}{z} = \frac{\hat{a}(z) - 1}{z}$$
.

On the other hand, for the right side we get

$$2\hat{a}(z) + \sum_{n=0}^{\infty} nz^n .$$

We identify the second term by relating it to the geometric series

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z} \quad \text{for} \quad |z| < 1 .$$

In particular,

$$\sum_{n=0}^{\infty} nz^n = \sum_{n=0}^{\infty} z \frac{d}{dz} z^n$$

$$= z \frac{d}{dz} \sum_{n=0}^{\infty} z^n$$

$$= z \frac{d}{dz} (1-z)^{-1}$$

$$= \frac{z}{(1-z)^2}.$$

We now solve the equation

$$\frac{(\hat{a}(z) - 1}{z} = 2\hat{a}(z) + \frac{z}{(1 - z)^2} .$$

Solving for the function $\hat{a}(z)$, we get

$$\hat{a}(z) = \frac{1 - 2z + 2z^2}{(1 - z)^2 (1 - 2z)} \ .$$

Now we use a partial fraction expansion to simplify this expression. Partial fraction expansions are discussed in Chapter 11 of Feller I, for example. You can look it up online too; e.g.,

http://cnx.rice.edu/content/m2111/latest/

You could even do the partial fraction with an online tool: http://mss.math.vanderbilt.edu/cgi-

bin/MSSAgent/pscrooke/MSS/partialfract.def

To do the partial fraction expansion, we write

$$\hat{a}(z) = \frac{1 - 2z + 2z^2}{(1 - z)^2(1 - 2z)} = \frac{A}{(1 - z)^2} + \frac{B}{(1 - z)} + \frac{C}{(1 - 2z)}$$
.

Multiply both sides by $(1-z)^2$ and then let z=1 to get A=-1; then multiply both sides by (1-2z) and let z=1/2 to get C=2; substitute z=0 with known A and C to get B=0. Thus the final form is

$$\hat{a}(z) = \frac{-1}{(1-z)^2} + \frac{2}{(1-2z)}$$
.

. In this simple form, we recognize that

$$a_n = 2^{n+1} - n - 1, \quad n \ge 0.$$

(h) Consider the recurrence

$$a_{n+1} = a_n + a_{n-1}$$
, for $n \ge 0$, with $a_0 = 1$ and $a_1 = 1$.

Find the generating function of the sequence $\{a_n : n \ge 0\}$. Use generating functions to find a good approximation for a_n for large n.

This is the recurrence for the famous *Fibonacci numbers*. Proceeding as in the previous part (multiplying both sides by z^n and adding), we get

$$\hat{a}(z) = \frac{1}{1 - z - z^2} \ .$$

Recall that

$$1 - z - z^2 = (1 - zr_+)(1 - zr_-) ,$$

where

$$r_{+} = \frac{(1+\sqrt{5})}{2}$$
 and $r_{-} = \frac{(1-\sqrt{5})}{2}$.

Now doing the partial-fraction expansion, we get

$$\frac{1}{1-z-z^2} = \frac{1}{(r_+ - r_-)} \left(\frac{r_+}{1-zr_+} - \frac{r_-}{1-zr_-} \right)$$

and

$$a_n = \frac{1}{\sqrt{5}}(r_+^{n+1} - r_-^{n+1}), \text{ for } n \ge 0.$$

That is Binet's Formula for the n^{th} Fibonacci number; e.g.,

http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fib.html

Note that r_+ is $\phi \equiv 1.618 \cdots$.

However, since $r_+ > 1$ and $|r_-| < 1$, the second term becomes negligible for large n. Hence we have

 $a_n \approx \frac{1}{\sqrt{5}} r_+^{n+1}$ for large n.

Moreover, the formula produces the *exact* result if we take the nearest integer to the computed value.

3. What Can We Do With Transforms?

- (a) Characterize the distribution of a sum of independent random variables.
- (b) Calculate moments of a random variable.
- (c) Find limits of sequences.
- (d) Establish probability limits, such as the LLN and CLT.

A good reference for this is Chapter 6 of Chung, A Course in Probability Theory. That is a great book on measure-theoretic probability. The key result behind these proofs is the continuity theorem for cf's.

Theorem 0.1 (continuity theorem) Suppose that X_n and X are real-valued random variables, $n \geq 1$. Let ϕ_n and ϕ be their characteristic functions (cf's), which necessarily are well defined. Then

$$X_n \Rightarrow X \quad as \quad n \to \infty$$

if and only if

$$\phi_n(t) \to \phi(t)$$
 as $n \to \infty$ for all t .

Now to prove the WLLN (convergence in probability, which is equivalent to convergence in distribution here, because the limit is deterministic) and the CLT, we exploit the continuity theorem for cf's and the following two lemmas:

Lemma 0.1 (convergence to an exponential) If $\{c_n : n \ge 1\}$ is a sequence of complex numbers such that $c_n \to c$ as $n \to \infty$, then

$$(1+(c_n/n))^n \to e^c \quad as \quad n \to \infty$$
.

Lemma 0.2 (Taylor's theorem) If $E[|X^k|] < \infty$, then the following version of Taylor's theorem is valid for the characteristic function $\phi(t) \equiv E[e^{itX}]$

$$\phi(t) = \sum_{i=0}^{j=k} \frac{E[X^j](it)^j}{j!} + o(t^k) \quad as \quad t \to 0$$

where o(t) is understood to be a quantity (function of t) such that

$$\frac{o(t)}{t} \to 0$$
 as $t \to 0$.

Suppose that $\{X_n : n \ge 1\}$ is a sequence of IID random variables. Let

$$S_n \equiv X_1 + \dots + X_n, \quad n \ge 1$$
.

Theorem 0.2 (WLLN) If $E[|X|] < \infty$, then

$$\frac{S_n}{n} \Rightarrow EX \quad as \quad n \to \infty .$$

Proof. Look at the cf of S_n/n :

$$\phi_{S_n/n}(t) \equiv E[e^{itS_n/n}] = \phi_X(t/n)^n = (1 + \frac{itEX}{n} + o(t/n))^n$$

by the second lemma above. Hence, we can apply the first lemma to deduce that

$$\phi_{S_n/n}(t) \to e^{itEX}$$
 as $n \to \infty$.

By the continuity theorem for cf's (convergence in distribution is equivalent to convergence of cf's), the WLLN is proved.

Theorem 0.3 (CLT) If $E[X^2] < \infty$, then

$$\frac{S_n - nEX}{\sqrt{n\sigma^2}} \Rightarrow N(0,1) \quad as \quad n \to \infty ,$$

where $\sigma^2 = Var(X)$.

Proof. For simplicity, consider the case of EX = 0. We get that case after subtracting the mean. Look at the cf of $S_n/\sqrt{n\sigma^2}$:

$$\begin{split} \phi_{S_n/\sqrt{n\sigma^2}}(t) & \equiv E[e^{it[S_n/\sqrt{n\sigma^2}]}] \\ & = \phi_X(t/\sqrt{n\sigma^2})^n \\ & = (1 + (\frac{it}{\sqrt{n\sigma^2}})EX + (\frac{it}{\sqrt{n\sigma^2}})^2 \frac{EX^2}{2} + o(t/n))^n \\ & = (1 + \frac{-t^2}{2n} + o(t/n))^n \\ & \to e^{-t^2/2} = \phi_{N(0,1)}(t) \end{split}$$

by the two lemmas above. Thus, by the continuity theorem, the CLT is proved.

- (e) Determine the asymptotic form of sequences and functions.
- (f) Help solve differential equations.

Let f' and f'' be the first and second derivative of the function f, which is a real-valued function on the positive half line. Using integration by parts, we see that

$$\hat{f}'(s) \equiv \int_0^\infty e^{-sx} f'(x) dx = -f(0) + s\hat{f}(s)$$

and

$$\hat{f}''(s) \equiv \int_0^\infty e^{-sx} f''(x) \, dx = -f'(0) - sf(0) + s^2 \hat{f}(s) \ .$$

Example 0.1 (first example)

Now, to illustrate how Laplace transforms can help solve differential equations, suppose that we want to solve

$$f'(t) - f(t) = -2, \quad f(0) = 1.$$

By taking Laplace transforms, we can replace the differential equation by an algebraic equation. Specifically, we get

$$-1 + s\hat{f}(s) - \hat{f}(s) = \frac{-2}{s}$$
,

which implies that

$$\hat{f}(s) = \frac{s-2}{s(s-1)} = \frac{2}{s} - \frac{1}{s-1} \ ,$$

using a partial-fraction expansion in the last step. But then we can directly recognize the form of the two terms, so that we see that

$$f(t) = 2 - e^t, \quad t \ge 0$$
.

Example 0.2 (second example)

If, instead, we have the slightly more complicated example

$$f''(t) + f(t) = \sin t$$
, $f(0) = f'(0) = 0$,

then we get

$$s^{2}\hat{f}(s) + \hat{f}(s) = \frac{1}{s^{2} + 1}$$

or

$$\hat{f}(s) = \frac{1}{(s^2 + 1)^2} \ .$$

We can use a book of Laplace transform tables to find that

$$f(t) = \frac{(\sin t - t \cos t)}{2}, \quad t \ge 0.$$

(g) Calculate cumulative distributions functions by numerical inversion.

See J. Abate and W. Whitt, "Numerical inversion of Laplace transforms of probability distributions," *ORSA Journal of Computing*, vol. 7, 1995, pages 36–43. The longer-term numerical-inversion homework assignment covers this.