13.1: Transform Methods

As pointed out in the units on Expectation and Variance, the mathematical expectation $E[X] = \mu_X$ of a random variable X locates the center of mass for the induced distribution, and the expectation

$$E[g(X)] = E[(X - E[X])^2] = \text{Var}[X] = \sigma_X^2$$

measures the spread of the distribution about its center of mass. These quantities are also known, respectively, as the mean (moment) of X and the second moment of X about the mean. Other moments give added information. For example, the third moment about the mean $E[(X - \mu_X)^3]$ gives information about the skew, or asymetry, of the distribution about the mean. We investigate further along these lines by examining the expectation of certain functions of X. Each of these functions involves a parameter, in a manner that completely determines the distribution. For reasons noted below, we refer to these as *transforms*. We consider three of the most useful of these.

Three basic transforms

We define each of three transforms, determine some key properties, and use them to study various probability distributions associated with random variables. In the section on integral transforms, we show their relationship to well known integral transforms. These have been studied extensively and used in many other applications, which makes it possible to utilize the considerable literature on these transforms.

Definition

The *moment generating function* M_X for random variable X (i.e., for its distribution) is the function

$$M_X(s) = E[e^{sX}] \ (s \text{ is a real or complex parameter})$$

The *characteristic function* ϕ_X for random variable X is

$$\varphi_X(u) = E[e^{iuX}] \ (i^2 = -1, u \text{ is a real parameter})$$

The *generating function* $g_X(s)$ for a nonnegative, integer-valued random variable X is

$$g_X(s) = E[s^X] = \sum_k s^k P(X = k)$$

The generating function $E[s^X]$ has meaning for more general random variables, but its usefulness is greatest for nonnegative, integer-valued variables, and we limit our consideration to that case.

The defining expressions display similarities which show useful relationships. We note two which are particularly useful.

$$M_X(s) = E[e^{sX}] = E[(e^s)^X] = g_X(e^s) \text{ and } \varphi_X(u) = E[e^{iuX}] = M_X(iu)$$

Because of the latter relationship, we ordinarily use the moment generating function instead of the characteristic function to avoid writing the complex unit *i*. When desirable, we convert easily by the change of variable.

The integral transform character of these entities implies that there is essentially a one-to-one relationship between the transform and the distribution.

Moments

The name and some of the importance of the moment generating function arise from the fact that the derivatives of M_X evaluateed at $s = 0$ are the moments about the origin. Specifically

$$M_X^{(k)}(0) = E[X^k], \text{ provided the } k\text{th moment exists}$$

Since expectation is an integral and because of the regularity of the integrand, we may differentiate inside the integral with respect to the parameter.

$$M_X'(s) = \frac{d}{ds}E[e^{sX}] = E[\frac{d}{ds}e^{sX}] = E[Xe^{sX}]$$

Upon setting $s = 0$, we have $M_X'(0) = E[X]$. Repeated differentiation gives the general result. The corresponding result for the characteristic function is $\varphi^{(k)}(0) = i^k E[X^k]$.

Example 13.1.1 The exponential distribution

The density function is $f_X(t) = \lambda e^{-\lambda t}$ for $t \ge 0$.

$$M_X(s) = E[e^{sX}] = \int_0^\infty \lambda e^{-(\lambda - s)t}\, dt = \frac{\lambda}{\lambda - s}$$

$$M'_X(s) = \frac{\lambda}{(\lambda - s)^2} \quad M''_X(s0 = \frac{2\lambda}{(\lambda - s)^3}$$

$$E[X] = M'_X(0) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda} \quad E[X^2] = M''_X(0) = \frac{2\lambda}{\lambda^3} = \frac{2}{\lambda^2}$$

From this we obtain $\text{Var}[X] = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$.

The generating function does not lend itself readily to computing moments, except that

$$g'_X(s) = \sum_{k=1}^\infty k s^{k-1} P(X = k) \quad \text{so that } g'_X(1) = \sum_{k=1}^\infty k P(X = k) = E[X]$$

For higher order moments, we may convert the generating function to the moment generating function by replacing s with e^s, then work with M_X and its derivatives.

Example 13.1.2 The Poisson (μ) distribution

$P(X = k) = e^{-\mu} \dfrac{\mu^k}{k!}$, $k \ge 0$, so that

$$g_X(s) = e^{-\mu} \sum_{k=0}^\infty s^k \frac{\mu^k}{k!} = e^{-\mu} \sum_{k=0}^\infty \frac{(s\mu)^k}{k!} = e^{-\mu} e^{\mu s} = e^{\mu(s-1)}$$

We convert to M_X by replacing s with e^s to get $M_X(s) = e^{u(e^s - 1)}$. Then

$$M'_X(s) = e^{u(e^s-1)} \mu e^s \quad M''_X(s) = e^{u(e^s-1)}[\mu^2 e^{2s} + \mu e^s]$$

so that

$$E[X] = M'_X(0) = \mu,\ E[X^2] = M''_X(0) = \mu^2 + \mu,\ \text{and } \text{Var}[X] = \mu^2 + \mu - \mu^2 = \mu$$

These results agree, of course, with those found by direct computation with the distribution.

Operational properties

We refer to the following as *operational properties*.

(T1): If $Z = aX + b$, then

$$M_Z(s) = e^{bs} M_X(as),\ \varphi_Z(u) = e^{iub} \varphi_X(au),\ g_Z(s) = s^b g_X(s^a)$$

For the moment generating function, this pattern follows from

$$E[e^{(aX+b)s}] = s^{bs} E[e^{(as)X}]$$

Similar arguments hold for the other two.

(T2): If the pair $\{X, Y\}$ is independent, then

$$M_{X+Y}(s) = M_X(s) M_Y(s),\ \varphi_{X+Y}(u) = \varphi_X(u) \varphi_Y(u),\ g_{X+Y}(s) = g_X(s) g_Y(s)$$

For the moment generating function, e^{sX} and e^{sY} form an independent pair for each value of the parameter s. By the product rule for expectation

$$E[e^{s(X+Y)}] = E[e^{sX} e^{sY}] = E[e^{sX}] E[e^{sY}]$$

Similar arguments are used for the other two transforms.

A partial converse for (T2) is as follows:

(T3): If $M_{X+Y}(s) = M_X(s) M_Y(s)$, then the pair $\{X + Y\}$ is uncorrelated. To show this, we obtain two expressions for $E[(X+Y)^2]$, one by direct expansion and use of linearity, and the other by taking the second derivative of the moment generating function.

$$E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[XY]$$

$$M''_{X+Y}(s) = [M_X(s)M_Y(s)]'' = M''_X(s)M_Y(s) + M_X(s)M''_Y(s) + 2M'_X(s)M'_Y(s)$$

On setting $s = 0$ and using the fact that $M_X(0) = M_Y(0) = 1$, we have

$$E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[X]E[Y]$$

which implies the equality $E[XY] = E[X]E[Y]$.

Note that we have *not* shown that being uncorrelated implies the product rule.

We utilize these properties in determining the moment generating and generating functions for several of our common distributions.

Some discrete distributions

Indicator function $X = I_E$ $P(E) = p$

$$g_X(s) = s^0 q + s^1 p = q + ps \quad M_X(s) = g_X(e^s) = q + pe^s$$

Simple random variable $X = \sum_{i=1}^n t_i I_{A_i}$ (primitive form) $P(A_i) = p_i$

$$M_X(s) = \sum_{i=1}^n e^{st_i} p_i$$

Binomial (n, p). $X = \sum_{i=1}^n I_{E_i}$ with $\{I_{E_i} : 1 \le i \le n\}$ iid $P(E_i) = p$

We use the product rule for sums of independent random variables and the generating function for the indicator function.

$$g_X(s) = \prod_{i=1}^n (q + ps) = (q + ps)^n \quad M_X(s) = (q + pe^s)^n$$

Geometric (p). $P(X = k) = pq^k$ $\forall k \ge 0$ $E[X] = q/p$ We use the formula for the geometric series to get

$$g_X(s) = \sum_{k=0}^\infty pq^k s^k = p\sum_{k=0}^\infty (qs)^k = \frac{p}{1-qs} \; M_X(s) = \frac{p}{1-qe^s}$$

Negative binomial (m, p) If Y_m is the number of the trial in a Bernoulli sequence on which the mth success occurs, and $X_m = Y_m - m$ is the number of failures before the mth success, then

$$P(X_m = k) = P(Y_m - m = k) = C(-m, k)(-q)^k p^m$$

$$\text{where } C(-m, k) = \frac{-m(-m-1)(-m-2)\cdots(-m-k+1)}{k!}$$

The power series expansion about $t = 0$ shows that

$$(1+t)^{-m} = 1 + C(-m, 1)t + C(-m, 2)t^2 + \cdots \quad \text{for } -1 < t < 1$$

Hence,

$$M_{X_m}(s) = p^m \sum_{k=0}^\infty C(-m, k)(-q)^k e^{sk} = [\frac{p}{1-qe^s}]^m$$

Comparison with the moment generating function for the geometric distribution shows that $X_m = Y_m - m$ has the same distribution as the sum of m iid random variables, each geometric (p). This suggests that the sequence is characterized by independent, successive waiting times to success. This also shows that the expectation and variance of X_m are m times the expectation and variance for the geometric. Thus

$$E[X_m] = mq/p \text{ and } \text{Var}[X_m] = mq/p^2$$

Poisson (μ) $P(X = k) = e^{-\mu}\frac{\mu^k}{k!}$ $\forall k \ge 0$ In Example 13.1.2, above, we establish $g_X(s) = e^{\mu(s-1)}$ and $M_X(s) = e^{\mu(e^s-1)}$. If $\{X, Y\}$ is an independent pair, with X ~ Poisson (λ) and Y ~ Poission (μ), then $Z = X + Y$ ~ Poisson $(\lambda + \mu)$. Follows from (T1) and product of exponentials.

Some absolutely continuous distributions

Uniform on (a, b) $f_X(t) = \frac{1}{b-a}$ $a < t < b$

$$M_X(s) = \int e^{st} f_X(t)\, dt = \frac{1}{b-a} \int_a^b e^{st}\, dt = \frac{e^{sb} - e^{sa}}{s(b-a)}$$

Symmetric triangular $(-c, c)$

$$f_X(t) = I_{[-c,0)}(t) \frac{c+t}{c^2} + I_{[0,c]}(t) \frac{c-t}{c^2}$$

$$M_X(s) = \frac{1}{c^2} \int_{-c}^0 (c+t)e^{st}\, dt + \frac{1}{c^2} \int_0^c (c-t)e^{st}\, dt = \frac{e^{cs} + e^{-cs} - 2}{c^2 s^2}$$

$$= \frac{e^{cs} - 1}{cs} \cdot \frac{1 - e^{-cs}}{cs} = M_Y(s) M_Z(-s) = M_Y(s) M_{-Z}(s)$$

where M_Y is the moment generating function for Y ~ uniform $(0, c)$ and similarly for M_Z. Thus, X has the same distribution as the difference of two independent random variables, each uniform on $(0, c)$.

Exponential (λ) $f_X(t) = \lambda e^{-\lambda t}$, $t \ge 0$

In example 1, above, we show that $M_X(s) = \dfrac{\lambda}{\lambda - s}$.

Gamma(α, λ) $f_X(t) = \dfrac{1}{\Gamma(\alpha)} \lambda^\alpha t^{\alpha-1} e^{-\lambda t}$ $t \ge 0$

$$M_X(s) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-(\lambda - s)t}\, dt = [\frac{\lambda}{\lambda - s}]^\alpha$$

For $\alpha = n$, a positive integer,

$$M_X(s) = [\frac{\lambda}{\lambda - s}]^n$$

which shows that in this case X has the distribution of the sum of n independent random variables each exponential (λ).

Normal (μ, σ^2).

- The standardized normal, Z ~ $N(0, 1)$

$$M_Z(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{st} e^{-t^2/2}\, dt$$

  Now $st - \dfrac{t^2}{2} = \dfrac{s^2}{2} - \dfrac{1}{2}(t-s)^2$ so that

$$M_Z(s) = e^{s^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-(t-s)^2/2}\, dt = e^{s^2/2}$$

  since the integrand (including the constant $(1/\sqrt{2\pi})$ is the density for $N(s, 1)$.

- $X = \sigma Z + \mu$ implies by property (T1)

$$M_X(s) = e^{s\mu} e^{\sigma^2 s^2/2} = \exp(\frac{\sigma^2 s^2}{2} + s\mu)$$

Example 13.1.3 Affine combination of independent normal random variables

Suppose $\{X, Y\}$ is an independent pair with X ~ $N(\mu_X, \sigma_X^2)$ and Y ~ $N(\mu_Y, \sigma_Y^2)$. Let $Z = aX + bY + c$. The Z is normal, for by properties of expectation and variance

$$\mu_Z = a\mu_X + b\mu_Y + c \quad \text{and } \sigma_Z^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2$$

and by the operational properties for the moment generating function

$$M_Z(s) = e^{sc} M_X(as) M_Y(bs) = \exp(\frac{(a^2\sigma_X^2 + b^2\sigma_Y^2)s^2}{2} + s(a\mu_X + b\mu_Y + c))$$

$$= \exp(\frac{\sigma_Z^2 s^2}{2} + s\mu_Z)$$

This form of M_Z shows that Z is normally distributed.

Moment generating function and simple random variables

Suppose $X = \sum_{i=1}^{n} t_i I_{A_i}$ in canonical form. That is, A_i is the event $\{X = t_i\}$ for each of the distinct values in the range of X_i with $p_i = P(A_i) = P(X = t_i)$. Then the moment generating function for X is

$$M_X(s) = \sum_{i=1}^{n} p_i e^{st_i}$$

The moment generating function M_X is thus related directly and simply to the distribution for random variable X.

Consider the problem of determining the sum of an independent pair $\{X, Y\}$ of simple random variables. The moment generating function for the sum is the product of the moment generating functions. Now if $Y = \sum_{j=1}^{m} u_j I_{B_j}$, with $P(Y = u_j) = \pi_j$, we have

$$M_X(s)M_Y(s) = (\sum_{i=1}^{n} p_i e^{st_i})(\sum_{j=1}^{m} \pi_j e^{su_j}) = \sum_{i,j} p_i \pi_j e^{s(t_i+u_j)}$$

The various values are sums $t_i + u_j$ of pairs (t_i, u_j) of values. Each of these sums has probability $p_i \pi_j$ for the values corresponding to t_i, u_j. Since more than one pair sum may have the same value, we need to sort the values, consolidate like values and add the probabilties for like values to achieve the distribution for the sum. We have an m-function *mgsum* for achieving this directly. It produces the pair-products for the probabilities and the pair-sums for the values, then performs a csort operation. Although not directly dependent upon the moment generating function analysis, it produces the same result as that produced by multiplying moment generating functions.

Example 13.1.4 Distribution for a sum of independent simple random variables

Suppose the pair $\{X, Y\}$ is independent with distributions

$$X = [1\ 3\ 5\ 7]\ Y = [2\ 3\ 4]\ PX = [0.2\ 0.4\ 0.3\ 0.1]\ PY = [0.3\ 0.5\ 0.2]$$

Determine the distribution for $Z = X + Y$.

```
X = [1 3 5 7];
Y = 2:4;
PX = 0.1*[2 4 3 1];
PY = 0.1*[3 5 2];
[Z,PZ] = mgsum(X,Y,PX,PY);
disp([Z;PZ]')
    3.0000    0.0600
    4.0000    0.1000
    5.0000    0.1600
    6.0000    0.2000
    7.0000    0.1700
    8.0000    0.1500
    9.0000    0.0900
   10.0000    0.0500
   11.0000    0.0200
```

This could, of course, have been achieved by using icalc and csort, which has the advantage that other functions of X and Y may be handled. Also, since the random variables are nonnegative, integer-valued, the MATLAB convolution function may be used (see Example 13.1.7). By repeated use of the function mgsum, we may obtain the distribution for the sum of more than two simple random variables. The m-functions mgsum3 and mgsum4 utilize this strategy.

The techniques for simple random variables may be used with the simple approximations to absolutely continuous random variables.

Example 13.1.5 Difference of uniform distribution

The moment generating functions for the uniform and the symmetric triangular show that the latter appears naturally as the difference of two uniformly distributed random variables. We consider X and Y iid, uniform on [0,1].

```
tappr
Enter matrix [a b] of x-range endpoints  [0 1]
Enter number of x approximation points  200
Enter density as a function of t  t<=1
Use row matrices X and PX as in the simple case
[Z,PZ] = mgsum(X,-X,PX,PX);
plot(Z,PZ/d)               % Divide by d to recover f(t)
%  plotting details   ---  see Figure 13.1.1
```

Figure one is a density graph. It is titled, Density for difference two variables, each uniform (0, 1). The horizontal axis of the graph is labeled, t, and the vertical graph is labeled fZ(t). The plot of the density is triangular, beginning at (-1, 0), and increasing at a constant slope to point (0, 1). The graph continues after this point downward with a constant negative slope to point (1, 0).

Figure 13.1.1. Density for the difference of an independent pair, uniform (0,1).

The generating function

The form of the generating function for a nonnegative, integer-valued random variable exhibits a number of important properties.

$$X = \sum_{k=0}^{\infty} kI_{A_i} \text{ (canonical form) } p_k = P(A_k) = P(X = k) \quad g_X(s) = \sum_{k=0}^{\infty} s^k p_k$$

As a power series in s with nonegative coefficients whose partial sums converge to one, the series converges at least for $|s| \le 1$.

The coefficients of the power series display the distribution: for value k the probability $p_k = P(X = k)$ is the coefficient of s^k.

The power series expansion about the origin of an analytic function is unique. If the generating function is known in closed form, the unique power series expansion about the origin determines the distribution. If the power series converges to a known closed form, that form characterizes the distribution.

For a simple random variable (i.e. $p_k = 0$ for $k > n$), g_X is a polynomial.

Example 13.1.6 The Poisson distribution

In Example 13.1.2, above, we establish the generating function for X ~ Poisson (μ) from the distribution. Suppose, however, we simply encounter the generating function

$$g_X(s) = e^{m(s-1)} = e^{-m}e^{ms}$$

From the known power series for the exponential, we get

$$g_X(s) = e^{-m} \sum_{k=0}^{\infty} \frac{(ms)^k}{k!} = e^{-m} \sum_{k=0}^{\infty} s^k \frac{m^k}{k!}$$

We conclude that

$$P(X = k) = e^{-m} \frac{m^k}{k!}, 0 \le k$$

which is the Poisson distribution with parameter $\mu = m$.

For simple, nonnegative, integer-valued random variables, the generating functions are polynomials. Because of the product rule (T2), the problem of determining the distribution for the sum of independent random variables may be handled by the process of multiplying polynomials. This may be done quickly and easily with the MATLAB *convolution* function.

Example 13.1.7 Sum of independent simple random variables

Suppose the pair $\{X, Y\}$ is independent, with

$$g_X(s) = \frac{1}{10}(2 + 3s + 3s^2 + 2s^5) \quad g_Y(s) = \frac{1}{10}(2s + 4s^2 + 4s^3)$$

In the MATLAB function convolution, all powers of *s* must be accounted for by including zeros for the missing powers.

```
gx = 0.1*[2 3 3 0 0 2];      % Zeros for missing powers 3, 4
gy = 0.1*[0 2 4 4];          % Zero  for missing power 0
gz = conv(gx,gy);
a = ['       Z         PZ'];
b = [0:8;gz]';
disp(a)
       Z         PZ          % Distribution for Z = X + Y
disp(b)
         0         0
    1.0000    0.0400
    2.0000    0.1400
    3.0000    0.2600
    4.0000    0.2400
    5.0000    0.1200
    6.0000    0.0400
    7.0000    0.0800
    8.0000    0.0800
```

If mgsum were used, it would not be necessary to be concerned about missing powers and the corresponding zero coefficients.

Integral transforms

We consider briefly the relationship of the moment generating function and the characteristic function with well known integral transforms (hence the name of this chapter).

Moment generating function and the Laplace transform

When we examine the integral forms of the moment generating function, we see that they represent forms of the Laplace transform, widely used in engineering and applied mathematics. Suppose F_X is a probability distribution function with $F_X(-\infty) = 0$. The bilateral Laplace transform for F_X is given by

$$\int_{-\infty}^{\infty} e^{-st} F_X(t)\, dt$$

The Laplace-Stieltjes transform for F_X is

$$\int_{-\infty}^{\infty} e^{-st} F_X(dt)$$

Thus, if M_X is the moment generating function for X, then $M_X(-s)$ is the Laplace-Stieltjes transform for X (or, equivalently, for F_X).

The theory of Laplace-Stieltjes transforms shows that under conditions sufficiently general to include all practical distribution functions

$$M_X(-s) = \int_{-\infty}^{\infty} e^{-st} F_X(dt) = s \int_{-\infty}^{\infty} e^{-st} F_X(t)\, dt$$

Hence

$$\frac{1}{s} M_X(-s) = \int_{-\infty}^{\infty} e^{-st} F_X(t)\, dt$$

The right hand expression is the bilateral Laplace transform of F_X. We may use tables of Laplace transforms to recover F_X when M_X is known. This is particularly useful when the random variable X is nonnegative, so that $F_X(t) = 0$ for $t < 0$.

If X is absolutely continuous, then

$$M_X(-s) = \int_{-\infty}^{\infty} e^{-st} f_X(t)\, dt$$

In this case, $M_X(-s)$ is the bilateral Laplace transform of f_X. For nonnegative random variable X, we may use ordinary tables of the Laplace transform to recover f_X.

Example 13.1.8 Use of Laplace transform

Suppose nonnegative X has moment generating function

$$M_X(s) = \frac{1}{(1-s)}$$

We know that this is the moment generating function for the exponential (1) distribution. Now,

$$\frac{1}{s} M_X(-s) = \frac{1}{s(1+s)} = \frac{1}{s} - \frac{1}{1+s}$$

From a table of Laplace transforms, we find $1/s$ is the transform for the constant 1 (for $t \ge 0$) and $1/(1+s)$ is the transform for $e^{-t}, t \ge 0$, so that $F_X(t) = 1 - e^{-t} t \ge 0$, as expected.

Example 13.1.9 Laplace transform and the density

Suppose the moment generating function for a nonnegative random variable is

$$M_X(s) = [\frac{\lambda}{\lambda - s}]^{\alpha}$$

From a table of Laplace transforms, we find that for $\alpha > 0$.

$$\frac{\Gamma(\alpha)}{(s-a)^{\alpha}} \text{ is the Laplace transform of } t^{\alpha-1} e^{at}\ t \ge 0$$

If we put $a = -\lambda$, we find after some algebraic manipulations

$$f_X(t) = \frac{\lambda^{\alpha} t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)}, t \ge 0$$

Thus, X ~ gamma (α, λ), in keeping with the determination, above, of the moment generating function for that distribution.

The characteristic function

Since this function differs from the moment generating function by the interchange of parameter s and iu, where i is the imaginary unit, $i^2 = -1$, the integral expressions make that change of parameter. The result is that Laplace transforms become Fourier transforms. The theoretical and applied literature is even more extensive for the characteristic function.

Not only do we have the operational properties (T1) and (T2) and the result on moments as derivatives at the origin, but there is an important expansion for the characteristic function.

An expansion theorem

If $E[[X]^n] < \infty$, then

$$\varphi^{(k)}(0) = i^k E[X^k], \text{ for } 0 \le k \le n \text{ and } \varphi(u) = \sum_{k=0}^{n} \frac{(iu)^k}{k!} E[X^k] + o(u^n) \text{ as } u \to 0$$

We note one limit theorem which has very important consequences.

A fundamental limit theorem

Suppose $\{F_n : 1 \le n\}$ is a sequence of probability distribution functions and $\{\varphi_n : 1 \le n\}$ is the corresponding sequence of characteristic functions.

If F is a distribution function such that $F_n(t) \to F(t)$ at every point continuity for F, and ϕ is the characteristic function for F, then

$$\varphi_n(u) \to \varphi(u) \ \forall u$$

If $\varphi_n(u) \to \varphi(u)$ for all u and ϕ is continuous at 0, then ϕ is the characteristic function for distribution function F such that

$$F_n(t) \to F(t) \text{ at each point of continuity of } F$$

— □