Moment Generating Functions

Let X be a simple random variable asssuming the distinct values $x_1, \ldots, x_l$ with respective probabilities $p_1, \ldots, p_l$. Its *moment generating function* is

$$(9.1) \qquad M(t) = E[e^{tX}] = \sum_{i=1}^{l} p_i e^{tx_i}.$$

(See (5.19) for expected values of functions of random variables.) This function, defined for all real t, can be regarded as associated with X itself or as associated with its distribution—that is, with the measure on the line having mass p_i at x_i (see (5.12)).

If $c = \max_i |x_i|$, the partial sums of the series $e^{tX} = \sum_{k=0}^{\infty} t^k X^k / k!$ are bounded by $e^{|t|c}$, and so the corollary to Theorem 5.4 applies:

$$(9.2) \qquad M(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E[X^k].$$

Thus $M(t)$ has a Taylor expansion, and as follows from the general theory [A29], the coefficient of t^k must be $M^{(k)}(0)/k!$ Thus

$$(9.3) \qquad E[X^k] = M^{(k)}(0).$$

Furthermore, term-by-term differentiation in (9.1) gives

$$M^{(k)}(t) = \sum_{i=1}^{l} p_i x_i^k e^{tx_i} = E[X^k e^{tX}];$$

taking $t = 0$ here gives (9.3) again. Thus the moments of X can be calculated by successive differentiation, whence $M(t)$ gets its name. Note that $M(0) = 1$.

Example 9.1. If X assumes the values 1 and 0 with probabilities p and $q = 1 - p$, as in Bernoulli trials, its moment generating function is $M(t) = pe^t + q$. The first two moments are $M'(0) = p$ and $M''(0) = p$, and the variance is $p - p^2 = pq$. ∎

If $X_1, \ldots, X_n$ are independent, then for each t (see the argument following (5.10)), $e^{tX_1}, \ldots, e^{tX_n}$ are also independent. Let M and $M_1, \ldots, M_n$ be the respective moment generating functions of $S = X_1 + \cdots + X_n$ and of $X_1, \ldots, X_n$; of course, $e^{tS} = \prod_i e^{tX_i}$. Since by (5.25) expected values multiply for independent random variables, there results the fundamental relation

$$(9.4) \qquad M(t) = M_1(t) \cdots M_n(t).$$

This is an effective way of calculating the moment generating function of the sum S. The real interest, however, centers on the distribution of S, and so it is important to know that distributions can in principle be recovered from their moment generating functions.

Consider along with (9.1) another finite exponential sum $N(t) = \sum_j q_j e^{ty_j}$, and suppose that $M(t) = N(t)$ for all t. If $x_{i_0} = \max x_i$ and $y_{j_0} = \max y_j$, then $M(t) \sim p_{i_0} e^{tx_{i_0}}$ and $N(t) \sim q_{j_0} e^{ty_{j_0}}$ as $t \to \infty$, and so $x_{i_0} = y_{j_0}$ and $p_{i_0} = q_{i_0}$. The same argument now applies to $\sum_{i \ne i_0} p_i e^{tx_i} = \sum_{j \ne j_0} q_j e^{ty_j}$, and it follows inductively that with appropriate relabeling, $x_i = y_i$ and $p_i = q_i$ for each i. Thus *the function* (9.1) *does uniquely determine the* x_i *and* p_i.

Example 9.2. If $X_1, \ldots, X_n$ are independent, each assuming values 1 and 0 with probabilities p and q, then $S = X_1 + \cdots + X_n$ is the number of successes in n Bernoulli trials. By (9.4) and Example 9.1, S has the moment generating function

$$E\left[e^{tS}\right] = \left(pe^t + q\right)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} e^{tk}.$$

The right-hand form shows this to be the moment generating function of a distribution with mass $\binom{n}{k} p^k q^{n-k}$ at the integer k, $0 \le k \le n$. The uniqueness just established therefore yields the standard fact that $P[S = k] = \binom{n}{k} p^k q^{n-k}$. ∎

The *cumulant generating function of* X (or of its distribution) is

$$C(t) = \log M(t) = \log E\left[e^{tX}\right]. \tag{9.5}$$

(Note that $M(t)$ is strictly positive.) Since $C' = M'/M$ and $C'' = (MM'' - (M')^2)/M^2$, and since $M(0) = 1$,

$$C(0) = 0, \qquad C'(0) = E[X], \qquad C''(0) = \operatorname{Var}[X]. \tag{9.6}$$

Let $m_k = E[X^k]$. The leading term in (9.2) is $m_0 = 1$, and so a formal expansion of the logarithm in (9.5) gives

$$C(t) = \sum_{v=1}^{\infty} \frac{(-1)^{v+1}}{v} \left(\sum_{k=1}^{\infty} \frac{m_k}{k!} t^k \right)^v. \tag{9.7}$$

Since $M(t) \to 1$ as $t \to 0$, this expression is valid for t in some neighborhood of 0. By the theory of series, the powers on the right can be expanded and