Moments of binomial random variables

Let X be a binomial random variable with parameters n and p. Our object is to determine moments $E\big[X^r\big]$ in terms of n and p. We first show that if r is a positive integer, then

$$E[X(X-1)\cdots(X-r+1)] = p^r n(n-1)\cdots(n-r+1)\,. \tag{1}$$

To show this, we first note that if $n < r$, then both sides above are 0. Hence we may assume that $n \ge r$. For such n, the left hand side above is

$$= \sum_{k=0}^{n} k(k-1)\cdots(k-r+1)\binom{n}{k}p^k(1-p)^{n-k}\,.$$

If $k < r$, then the summand above is 0, so we may restrict k to the interval $r \le k \le n$. Thus the above is

$$= \sum_{k=r}^{n} k(k-1)\cdots(k-r+1)\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=r}^{n} k(k-1)\cdots(k-r+1)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

$$= \sum_{k=r}^{n} \frac{n!}{(k-r)!(n-k)!}p^k(1-p)^{n-k} = p^r n(n-1)\cdots(n-r+1)\sum_{k=r}^{n}\frac{(n-r)!}{(k-r)!(n-k)!}p^{k-r}(1-p)^{n-k}$$

On setting $k-r=\ell$, we find that the last sum on the right above is

$$\sum_{\ell=0}^{n-r}\frac{(n-r)!}{\ell!(n-r-\ell)!}p^\ell(1-p)^{n-r-\ell} = \sum_{\ell=0}^{n-r}\binom{n-r}{\ell}p^\ell(1-p)^{n-r-\ell} = \big(p+(1-p)\big)^{n-r} = 1$$

by the binomial theorem. Thus we have (1).

Next we note that if f and g are real-valued functions and X is a discrete random variable, then

$$\begin{aligned} E[f(X)+g(X)] &= \sum_i \big(f(x_i)+g(x_i)\big)p(x_i) = \sum_i f(x_i)p(x_i) + \sum_i g(x_i)p(x_i) \\ &= E[f(X)] + E[g(X)]\,. \end{aligned} \tag{2}$$

This is a special case of a more general result: $E[X+Y] = E[X]+E[Y]$ for arbitrary random variables.

By taking $r=1$ in (1), we see that $E[X]=np$. By taking $r=2$, we find that $E[X(X-1)] = p^2n(n-1)$. By (2) it follows that

$$E\big[X^2\big] = E[X(X-1)+X] = E[X(X-1)] + E[X] = p^2n(n-1)+pn = n^2p^2 + np(1-p)\,.$$

Hence

$$\mathrm{Var}\,(X) = E\big[X^2\big] - E[X]^2 = n^2p^2 + np(1-p) - (np)^2 = np(1-p)\,.$$

For $r=3$ we find that $E[X(X-1)(X-2)] = p^3n(n-1)(n-2)$ and $X^3 = X(X-1)(X-2) + 3X(X-1)+X$, so

$$E\big[X^3\big] = p^3n(n-1)(n-2) + 3p^2n(n-1) + pn\,.$$

For general r the product $x(x-1)\cdots(x-r+1)$ can be expanded, so there are integers $\left[{r \atop j}\right]$, known as *Stirling numbers of the first kind* such that

$$x(x-1)\cdots(x-r+1)=\sum_{j=1}^{r}\begin{bmatrix} r \\ j \end{bmatrix} x^j\,.$$

The numbers $\left[{r \atop j}\right]$ are sometimes denoted $s(r,j)$. They can be generated by the Pascal-like recursion

$$\begin{bmatrix} r \\ j \end{bmatrix}=\begin{bmatrix} r-1 \\ j-1 \end{bmatrix}-(r-1)\begin{bmatrix} r-1 \\ j \end{bmatrix}.$$

In the reverse direction there exist numbers $\left\{{r \atop j}\right\}$, known as *Stirling numbers of the second kind* such that

$$x^r=\sum_{j=1}^{r}\begin{Bmatrix} r \\ j \end{Bmatrix} x(x-1)\cdots(x-j+1)\,.$$

The $\left\{{r \atop j}\right\}$ are sometimes denoted $S(r,j)$. They can be generated by the Pascal-like recursion

$$\begin{Bmatrix} r \\ j \end{Bmatrix}=j\begin{Bmatrix} r-1 \\ j \end{Bmatrix}+\begin{Bmatrix} r-1 \\ j-1 \end{Bmatrix}.$$

Stirling numbers arise in combinatorics: $(-1)^{r-j}\left[{r \atop j}\right]$ is the number of permutations of r objects that have exactly j cycles. The number $\left\{{r \atop j}\right\}$ is the number of ways of partitioning a set of r objects into exactly j nonempty subsets.

$n\backslash k$	1	2	3	4	5	6
1	1					
2	-1	1				
3	2	-3	1			
4	-6	11	-6	1		
5	24	-50	35	-10	1	
6	-120	274	-225	85	-15	1

TABLE 1. Stirling numbers $\left[{n \atop k}\right]$ of the first kind.

$n\backslash k$	1	2	3	4	5	6
1	1					
2	1	1				
3	1	3	1			
4	1	7	6	1		
5	1	15	25	10	1	
6	1	31	90	65	15	1

TABLE 2. Stirling numbers $\left\{{n \atop k}\right\}$ of the second kind.