Additional Comments

1 Computing Parameters of the Binomial Distribution Directly

We have that, for a $Bin\left(n,p\right)$ Random Variable X,

$$EX=\sum_{i=1}^{n}\binom{n}{i}ip^{i}\left(1-p\right)^{n-i}$$

$$EX^{2}=\sum_{i=1}^{n}\binom{n}{i}i^{2}p^{i}\left(1-p\right)^{n-i}$$

The book has a clever trick to compute, recursively, EX^{k}. The following method is an alternate way, recursive as well.

1.1 Expected Value

This is not so hard:

$$ip^{i}\left(1-p\right)^{n-i}=\left(1-p\right)^{n}i\left(\frac{p}{1-p}\right)^{i}=\frac{p}{1-p}\left(1-p\right)^{n}i\left(\frac{p}{1-p}\right)^{i-1}$$

Hence, setting, say, $u=\frac{p}{1-p}$, the expected value looks like

$$p\left(1-p\right)^{n-1}\sum_{i=1}^{n}\binom{n}{i}iu^{i-1}=p\left(1-p\right)^{n-1}\frac{d}{du}\left[\sum_{i=1}^{n}\binom{n}{i}u^{i}\right]=p\left(1-p\right)^{n-1}\frac{d}{du}\left(1+u\right)^{n}=$$

$$=p\left(1-p\right)^{n-1}n\left(1+\frac{p}{1-p}\right)^{n-1}=np$$

This is the well-known result, and it can be quickly proved in a number of other ways, including X being the sum of n Bernoulli Random Variables, or from its moment generating function, and connected functions:

$$\sum_{i=1}^{n}\binom{n}{i}p^{i}(1-p)^{n-i}e^{ti}=\sum_{i=1}^{n}\binom{n}{i}\left(pe^{t}\right)^{i}\left(1-p\right)^{n-i}=\left(1-p+pe^{t}\right)^{n}$$

and

$$\left[\frac{d}{dt}\left(1+p\left(e^{t}-1\right)\right)^{n}\right]_{t=0}=\left[n\left(1+p\left(e^{t}-1\right)\right)^{n-1}pe^{t}\right]_{t=0}=np$$

1.2 Expectation of the Square

We could go through the moment generating function, or through the sum of (and it's important here) *independent* Bernoulli variables. But here is a direct calculation. We have to be a little deceitful here: we note that

$$EX^2 = EX^2 - EX + EX$$

so that

$$\sum_{i=1}^{n} \binom{n}{i} i^2 p^i (1-p)^{n-i} = \sum_{i=1}^{n} \binom{n}{i} i^2 p^i (1-p)^{n-i} - \sum_{i=1}^{n} \binom{n}{i} i p^i (1-p)^{n-i} + EX =$$

$$= \sum_{i=1}^{n} \binom{n}{i} i (i-1) p^i (1-p)^{n-1} + EX$$

We are now ready to repeat the trick in section 1.1:

$$\sum_{i=1}^{n} \binom{n}{i} i (i-1) p^i (1-p)^{n-1} = (1-p)^n \left(\frac{p}{1-p}\right)^2 \sum_{i=1}^{n} \binom{n}{i} \left(\frac{p}{1-p}\right)^{i-2}$$

Hence, setting again $u = \frac{p}{1-p}$, we have

$$EX^2 - EX = p^2 (1-p)^{n-2} \frac{d^2}{du^2} \sum_{i=1}^{n} \binom{n}{i} u^i = p^2 (1-p)^{n-2} \frac{d^2}{du^2} (1+u)^n =$$

$$= p^2 (1-p)^{n-2} n(n-1) \left(1 + \frac{p}{1-p}\right)^{n-2} = n(n-1)p^2$$

Combining our results, we have that

$$EX^2 = n(n-1)p^2 + np$$

Going to the *variance*, we then find that

$$Var[X] = EX^2 - (EX)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p)$$

1.3 Higher Moments

Of course, the same trick works for any moment. After all

$$E[X^m] = \sum_{i=0}^{n} i^m \binom{n}{i} p^i (1-p)^{n-i}$$

So, for $m = 3$, we note that $i(i-1)(i-2) = i^3 - 3i^2 + 2i$, for $m = 4$, $i(i-1)(i-2)(i-3) = i^4 - 6i^3 + 11i^2 - 6i$, and so on, and to compute $E[X^n]$ we just need to add/subtract the corresponding combination of lower moments to have an expression that can be evaluated through

$$\frac{d^m}{du^m} \sum \binom{n}{i} u^i = \frac{d^m}{du^m} (u+1)^n \text{ for } u = \frac{p}{1-p}$$

2 Countable Random Variables

In general, if a Random Variable (abbreviated to RV from now on) takes a countable number of values, say $x_1, x_2, \ldots$, the associated probabilities must form a convergent series of positive numbers: $\sum_{k=1}^{\infty} P[X = x_i] = 1$. Thus, if we want to define distributions that take on a countable number of values we need to come up with absolutely convergent series. To make the distribution explicit, we should also be able to evaluate the sum of such series, which is a totally different game.

Restricting, for simplicity, to the case of integer values RV's, we will, setting $p_n = P[X = n]$, be able to evaluate $\sum_{k=1}^{\infty} p_n$. To be precise, for any known convergent series $\sum_{k=1}^{\infty} c_k = C$, we may define a probability distribution $P[X = k] = \frac{c_k}{C}$. For convenience, we may choose to consider variable with values in $\mathbb{N}$, or in $\mathbb{Z}$[1].

For example, we know that $\sum_{k=0}^{\infty} q^k = c = \frac{1}{1-q}$, hence we can come up with a distribution of the form

$$P[X = n] = cq^{n-1}$$

and it turns out that $c = (1 - q)$. Of course, this is the geometric distribution.

Similarly, we are familiar with the convergent series

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$$

We may thus define a distribution via

$$P[X = k] = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, \ldots$$

which, of course, turns out to be the Poisson distribution.

The problem here is that it is generally hard to express the sum of a convergent series explicitly. Thus, it is not so easy to come up with clever and diverse explicit countable distributions.

Thus, for example, it turns out that $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$, so

$$P[X = k] = \frac{6}{\pi^2 k^2}$$

is a probability distribution (for variables such that no moments exist). In general,

$$\sum_{k=1}^{\infty} \frac{1}{k^n} = \zeta(n)$$

[1] As you most likely know, $\mathbb{N}$ stands for the set of *natural numbers*, that is positive integers, while $\mathbb{Z}$ is the set of non negative integers.

where ζ is Riemann's zeta function (a function with amazing ramifications in many diverse parts of mathematics), so we can say that

$$P\left[X=k\right]=\frac{1}{\zeta(n)k^n}$$

is a probability distribution (listed in the book as "Zeta" or "Zipf"). If you are curious, here is a snapshot from http://mathworld.wolfram.com/RiemannZetaFunction.html listing the first few positive integer values for ζ :

The values of $\zeta(n)$ for small positive integer values of n are

$$\zeta(1)=\infty$$
$$\zeta(2)=\frac{\pi^2}{6}$$
$$\zeta(3)=1.2020569032\ldots$$
$$\zeta(4)=\frac{\pi^4}{90}$$
$$\zeta(5)=1.0369277551\ldots$$
$$\zeta(6)=\frac{\pi^6}{945}$$
$$\zeta(7)=1.0083492774\ldots$$
$$\zeta(8)=\frac{\pi^8}{9450}$$
$$\zeta(9)=1.0020083928\ldots$$
$$\zeta(10)=\frac{\pi^{10}}{93\,555}.$$

These are distributions "with fat tails", in that they have moments only up to $n-2$. This shows up also in the fact that they *do not admit a moment generating function*, and *the corresponding characteristic function only has* $n-2$ *derivatives.*

3 Characteristic Functions of Distributions

For completeness, let's compute (as far as feasible) the moment generating function (or, when that's not available, the characteristic function: if the moment generating function exists, you get the characteristic function by substituting it in place of the real variable t, and, for integer valued distribution, the generating function by substituting z for e^t) of the distributions we have listed:

3.1 Bernoulli (parameter p)

We have

$$M_B(t)=pe^{1\cdot t}+(1-p)e^{0\cdot t}=1+p\left(e^t-1\right)$$

3.2 Binomial (parameter n,p)

We have

$$M(t)=\left[M_B(t)\right]^n=\left[1+p\left(e^t-1\right)\right]^n$$

3.3 Geometric (parameter q)

We have

$$M(t) = \sum_{n=1}^{\infty} e^{nt} q (1-q)^{n-1} = e^t q \sum_{n=1}^{\infty} \left[e^t (1-q) \right]^{n-1} = \frac{e^t q}{1 - e^t (1-q)}$$

3.4 Poisson (parameter λ)

We have

$$M_P = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda (e^t - 1)}$$

3.4.1 Another way to view the Poisson distribution as the limit of binomials

We have to rely on a theorem that is beyond the scope of our course (and is based, essentially, on the possibility of inverting the Laplace and/or Fourier transform), which states, roughly, that if a sequence of distributions corresponds to a sequence of MGF's (or characteristic functions) that converge to a MGF (or a characteristic function), then, in a natural sense, so do the distributions themselves[2].

Suppose we have a sequence of Binomial distributions with

$$M_n(t) = \left[1 + p \left(e^t - 1 \right) \right]^n$$

and $n \to \infty, p \to 0, np = \lambda > 0$ constant. Then,

$$M_n(t) = \sum_{k=0}^{n} \binom{n}{k} p^k \left(e^t - 1 \right)^k = \sum_{k=0}^{\infty} \frac{n!}{k!(n-k)!} \left(e^t - 1 \right)^k p^k$$

Now, just as in the usual proof, we may use the fact that

$$\frac{n!}{(n-k)!} \approx n^k$$

so that

$$M_n(t) \approx \sum_{k=0}^{\infty} \frac{\left(np \left(e^t - 1 \right) \right)^k}{k!} = e^{\lambda (e^t - 1)}$$

having set $\lambda = np$.

Note that, taking derivatives, the sequences of derived functions, $\{M_n'\}, \{M_n''\}$, etc. converge nicely, so that it is legitimate to exchange derivatives and limits. It follows that we can indeed evaluate the moments fo the Poisson distirbution by taking limits of the moments of the approximating Binomials.

[2] The technical definition is as follows. Suppose you have a sequence of RV's X_n, whose characteristic functions are such that $\lim_{n\to\infty} C_{X_n}(t) = C(t)$. Then for any continuous bounded function f, $E[f(X_n)] \to E[f(X)]$, where X is a random variable whose distribution has characteristic function $C(t)$.

3.5 Zeta Distributions

Here we are not going to get a closed form expression. However, we can say that

$$C_n(t) = \sum_{k=1}^{\infty} e^{ikt} \frac{K_n}{k^n}$$

(where $K_n = \frac{1}{\zeta(n)}$). As long as the series of derivatives Are uniformly convergent, we can derive term by term, and find

$$\frac{d^m}{dt^m} C_n(t) = \sum_{k=1}^{\infty} K_n e^{ikt} \frac{k!}{(k-m)!k^n}$$

which, of course, is valid, as long as $\frac{k!}{(k-m)!} \leq k^{n-2}$, or $m \leq n-2$