Why the Multinomial, Geometric, and Poisson are Sound Probability Distributions

The definition of a probability distribution for any random variable has two parts. First, the distribution needs to reflect all possible values of the random variable, paired with the corresponding probabilities of those values. Second, the probabilities must sum to 1 (or 100%). Let's show you why the distributions from your last project satisfy these requirements pretty well.

Multinomial Distribution

The formula for the multinomial distribution is as follows...assuming you have an experiment with $\boldsymbol{n}$ disjoint outcomes 1,2,...$\boldsymbol{m}$, each with probability $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots \boldsymbol{p}_{\boldsymbol{m}}$, the probability of achieving $\boldsymbol{x}_1$ of type 1, $\boldsymbol{x}_2$ of type 2, up to $\boldsymbol{x}_{\boldsymbol{m}}$ of type $\boldsymbol{m}$, is

$$\frac{n!}{x_1!x_2!\ldots x_m!}p_1^{x_1}p_2^{x_2}\ldots p_m^{x_m}$$

It helps to see that this is a consequence of the binomial distribution; from there, the probability distribution is inevitable. For example, the coefficient $\frac{n!}{x_1!x_2!\ldots x_m!}$, although it looks terribly complex, can be broken down fairly nicely using combinations.

Since you have $\boldsymbol{n}$ things to assign, and each of those $\boldsymbol{n}$ things can be any of type 1, 2, up to $\boldsymbol{m}$, first, start by placing the $\boldsymbol{x}_1$ of type 1. You now have $\boldsymbol{n} - \boldsymbol{x}_1$ places left to assign the rest. Next up, place the type 2's, of which you have $\boldsymbol{x}_2$. After placing them, you have $\boldsymbol{n} - \boldsymbol{x}_1 - \boldsymbol{x}_2$ places left. Continuing until you get to type $\boldsymbol{m}$, you'll find you only have $\boldsymbol{x}_{\boldsymbol{m}}$ places left, so you have no choice but to place those in the remaining spots. Mathematically, that last paragraph looks like this:

$$\binom{n}{x_1}\binom{n-x_1}{x_2}\binom{n-x_1-x_2}{x_3}\ldots\binom{n-x_1-x_2-\ldots-x_{m+1}}{x_m}$$

However freakish that looks, don't fret...it breaks down nicely:

$$\left(\frac{n!}{(n-x_1)!x_1!}\right)\left(\frac{(n-x_1)!}{(n-x_1-x_2)!x_2!}\right)\left(\frac{(n-x_1-x_2)!}{(n-x_1-x_2-x_3)!x_3!}\right)\ldots\left(\frac{(n-x_1-x_2-\ldots-x_{m+1})!}{(n-x_1-x_2-\ldots-x_{m+1})!x_m!}\right)$$

What? You don't think that's ***nice***? OK, OK, OK...try this:

$$\left(\frac{n!}{\cancel{(n-x_1)!}x_1!}\right)\left(\frac{\cancel{(n-x_1)!}}{\cancel{(n-x_1-x_2)!}x_2!}\right)\left(\frac{\cancel{(n-x_1-x_2)!}}{\cancel{(n-x_1-x_2-x_3)!}x_3!}\right)\ldots\left(\frac{\cancel{(n-x_1-x_2-\ldots-x_{m+1})!}}{\cancel{(n-x_1-x_2-\ldots-x_{m+1})!}x_m!}\right)$$

Still don't like it? OK, let's remove all of those cancelled terms:

$$\frac{n!}{x_1!x_2!\ldots x_m!}$$

Look familiar? Heck yeah! It's the multinomial coefficient! That explains the first, somewhat awful looking factor. The probability factors must follow immediately from a neat extension of the binomial theorem:

$$(p+q)^n = \binom{n}{0}p^n + \binom{n}{1}p^{n-1}q + \binom{n}{2}p^{n-2}q^2 + \ldots + \binom{n}{n-1}pq^{n-1} + \binom{n}{n}q^n$$

In our case, however, we'll use the multinomial theorem to expand $\left(p_1 + p_2 + \ldots p_m\right)^n$. Now, if we expand that out, we're gonna get lots and lots and lots of terms; each one of those terms will be one possible arrangement of type 1, type 2, up to type ***m*** outcomes. Thus, the expansion of all of those terms represents the entire probability distribution of the multinomial. And, since the outcomes are disjoint, $p_1 = p_2 + = \ldots = p_m = 1$, which means that $\left(p_1 + p_2 + \ldots p_m\right)^n = 1$, which proves our probability requirement in our distribution.

A neat connection: the binomial coefficients gotten from the expansion of $(\boldsymbol{p} + \boldsymbol{q})^{\boldsymbol{n}}$ follow the entries ion the $\boldsymbol{n}^{\text{th}}$ row of Pascal's triangle. The multinomial coefficients of $(\boldsymbol{p}_1 + \boldsymbol{p}_2 + \ldots \boldsymbol{p}_{\text{m}})^{\boldsymbol{n}}$ follow the entries in the $\boldsymbol{n}^{\text{th}}$ row of Pascal's ***Pyramid***, with ***m*** faces, each face being a Pascal's Triangle. Then, any entry in the pyramid at any depth would be found by adding the numbers along intersecting lines on corresponding faces.

I discovered this in grad school, and was going to publish it (even had my advisor on board), but, like many good, useful, somewhat tricky ideas...it had already been published (by Martin Gardner, so I don't feel so badly). Still feels pretty cool to have "discovered" it, though.

Poisson Distribution

The Poisson distribution actually has its own little justification elsewhere on the enrichment page, but I thought it might be worthwhile for you to see why the probabilities sum to 1. For a Poisson distribution, the random variables have values 0, 1, 2,... so there is no "top", so to speak. In order for us to sum the probabilities, therefore, we have to use the idea of an infinite sum, which is discussed in calculus classes. Here's that that sum would look like:

$$p(0) + p(1) + p(2) + \ldots + p(n) = \frac{e^{-\mu}\mu^0}{0!} + \frac{e^{-\mu}\mu^1}{1!} + \frac{e^{-\mu}\mu^2}{2!} + \ldots + \frac{e^{-\mu}\mu^n}{n!}$$

Let's apply a little algebra first: factor out the $\boldsymbol{e}^{-\boldsymbol{\mu}}$ factor from each term:

$$\frac{e^{-\mu}\mu^0}{0!} + \frac{e^{-\mu}\mu^1}{1!} + \frac{e^{-\mu}\mu^2}{2!} + \ldots + \frac{e^{-\mu}\mu^n}{n!} = e^{-\mu}\left(\frac{\mu^0}{0!} + \frac{\mu^1}{1!} + \frac{\mu^2}{2!} + \ldots + \frac{\mu^n}{n!}\right)$$

Thinking forward, if we can get the quantity in parentheses to equal $\boldsymbol{e}^{\boldsymbol{\mu}}$, the distribution will sum to 1. Here comes the calculus: we're going to use something called a power series here; specifically, we're going to use the Taylor Series expansion of $\boldsymbol{e}^{\boldsymbol{x}}$:

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}, \text{ as } n \text{ goes to infinity}$$

Letting $x = \mu$ in the Taylor expansion above, we have that

$$e^{\mu} = 1 + \mu + \frac{\mu^2}{2!} + \ldots + \frac{\mu^n}{n!}, \text{ as } n \text{ goes to infinity}$$

And there we have it...that right hand side is exactly what's in parentheses below:

$$\frac{e^{-\mu}\mu^0}{0!} + \frac{e^{-\mu}\mu^1}{1!} + \frac{e^{-\mu}\mu^2}{2!} + \ldots + \frac{e^{-\mu}\mu^n}{n!} = e^{-\mu}\left(\frac{\mu^0}{0!} + \frac{\mu^1}{1!} + \frac{\mu^2}{2!} + \ldots + \frac{\mu^n}{n!}\right)$$

So, let's substitute the left hand side in there, and see what we get:

$$\begin{aligned}\frac{e^{-\mu}\mu^0}{0!} + \frac{e^{-\mu}\mu^1}{1!} + \frac{e^{-\mu}\mu^2}{2!} + \ldots + \frac{e^{-\mu}\mu^n}{n!} &= e^{-\mu}\left(\frac{\mu^0}{0!} + \frac{\mu^1}{1!} + \frac{\mu^2}{2!} + \ldots + \frac{\mu^n}{n!}\right) \\ &= e^{-\mu}(e^{\mu}) \\ &= 1\end{aligned}$$

And so, we've happily proved that the sum of the probabilities heads to 1, which makes the Poisson a bombproof probability distribution.

Geometric Distribution

To round out our little discussion, let's prove that the geometric distribution $p(x) = q^{n-1}p$ is a valid distribution for a discrete random variable. The values of this random variable, like the Poisson, are 0,1,2..., and that makes it more interesting to prove that $p + qp + q^2p + q^3p + \ldots + q^{n-1}p$ sums to 1 as n heads off to infinity. Here we go!

Assume that $p + qp + q^2p + q^3p + \ldots + q^{n-1}p$ does indeed sum to a number as n grows without bound (this might seem like either a silly assumption or a brazen one, depending on your viewpoint. Don't worry; whatever your view, it does converge to a number):

$$S = p + qp + q^2p + q^3p + \ldots + q^{n-1}p$$

Here's the slick (and sometimes contentious) part: multiply that entire equation by q:

$$qS = qp + q^2p + q^3p + q^4p + \ldots + q^np$$

Now, subtract this from the the original $S = p + qp + q^2p + q^3p + \ldots + q^{n-1}p$. Many, many terms will cancel:

$$S - qS = p + qp + q^2p + q^3p + \ldots + q^{n-1}p - (qp + q^2p + q^3p + q^4p + \ldots + q^np), \text{ so}$$

$$(1 - q)S = p - q^np$$

At this point, remember that n is very, very large. Also remember that q is a fraction between 0 and 1. That means that q^n is a number very, very close to zero...in fact, as n approaches infinity, q^n approaches 0. That means that the term q^np effectively equals 0:

$$(1 - q)S = p - q^np = p$$

Now, divide both sides by $(1 - q)$:

$$S = \frac{p}{1 - q}$$

At this point, remember that $p + q = 1$. Thus, $1 - q = p$. That means that

$$S = \frac{p}{1 - q} = \frac{p}{p} = 1$$

And this shows that the sum of the probabilities in the geometric distribution satisfies the "= 1" requirement.