Comments on
Poisson Distribution Presentation Notes (first draft)
By Jeremy Mahoney

John McCuan

September 29, 2023

1 Mathematical limits

Overall this is a great first draft. There is a minor problem at the top of page 2 which I hope you are enthusiastic to think about further.

The first indication there is a problem is that when one writes

$$\lim_{n\to\infty} g(n) = L,$$

the limit L should not depend on n. The limit L should just be some number independent of n. In this case, it would be okay for L to depend on k, but if you look at what you have, namely,

$$\lim_{n\to\infty} \left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n^k}{k!}. \tag{1}$$

the limit you are claiming depends on both n and k. This is incorrect.

On the face of it, one might attempt to replace your limit assertion with some kind of assertion that the two quantities are **asymptotic**. It turns out that in your case, this is entirely unnecessary, but I'm going to go in that direction for two reasons. The first is that it can be done, and it was one of the first things I thought of as an appropriate modification of what you were trying to say. The other is that consideration of the asymptotics in general is interesting in this case and illustrates quite clearly I think why you should give some extra attention to understanding and using limits correctly.

There are different notions of asymptotic, and one of them is strongly suggested by the limit assertion you have written. Let

$$g(n) = \binom{n}{k}$$

with k fixed. Were this quantity to actually have a limit $L \in \mathbb{R}$, then the following would be possible:

For any $\epsilon > 0$, there would be some $N > 0$ for which

$$|g(n) - L| < \epsilon \qquad \text{for all natural numbers } n \text{ with } n > N.$$

This is the (mathematically precise) definition of a limit. Of course this doesn't happen here because $g(n) = C(n,k)$ tends to $+\infty$ as n tends to ∞.

The notion of $g(n)$ being **asymptotic** to a quantity $h(n)$, for example $h(n) = n^k/k!$, suggested by your limiting assertion is the following:

For any $\epsilon > 0$, there exists some $N > 0$ for which

$$|g(n) - h(n)| < \epsilon \qquad \text{for all natural numbers } n \text{ with } n > N.$$

It is interesting that this does not hold for your quantities

$$g(n) = \binom{n}{k} \qquad \text{and} \qquad h(n) = \frac{n^k}{k!}.$$

I have taken $k = 3$ and plotted $g(n)$ and $h(n)$ for two intervals $10 \leq n \leq 100$ and $100 \leq n \leq 1000$ in Figure 1. These initial plots might suggest my objection to the two quantities being asymptotic is unfounded. In fact, in the plot with $100 \leq n \leq 1000$, I made the points $(n, h(n))$ large and green and the points $(n, g(n))$ small and black in order to even see where they are both plotted; at this scale, the plots are almost indistinguishable. Looking at the plot on the left, involving the lower values of n and the shorter interval, it can be discerned that some of the values of $h(n)$ appear to be larger than the values of $g(n)$. In fact, this is verifiable. Since both $g(n)$ and $h(n)$ have a common factor of $k!$, after this is canceled, it is clear that the difference $h(n) - g(n)$ is proportional to

$$n^k - n(n-1)(n-2)\cdots(n-k+1).$$

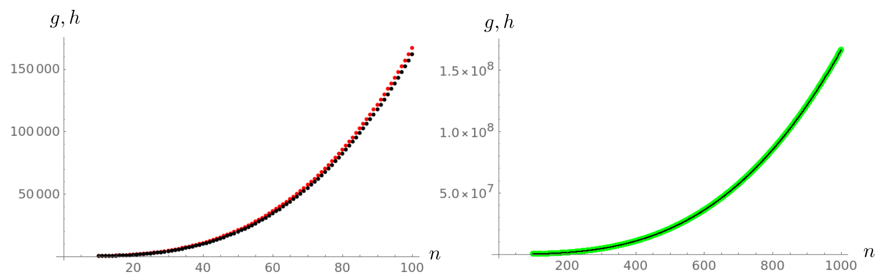

Figure 1: Plot of $g(n)$ (black) and $h(n)$ (red) for $10 \le n \le 100$ (left). Plot of $g(n)$ (black) and $h(n)$ (green) for $100 \le n \le 1000$ (right).

Since there are k factors in the second term and comparing factor by factor we see

$$\begin{aligned} n &= n \\ n &> n-1 \\ n &> n-2 \\ &\vdots \\ n &> n-k+1, \end{aligned}$$

it is clear that $h(n) > g(n)$.

The bad news becomes clearer, however, if we focus in on an even smaller interval and consider the actual magnitude of the difference $h(n) - g(n)$. In Figure 2 I have plotted $g(n)$ and $h(n)$ for $60 \le n \le 100$. The box in the center, which is not indicatated in the plot on the left and is 900 units tall, captures only one data point. The box on the right is the one indicated in the plot on the left and is 9900 units tall gives the bad news. In absolute terms $h(95)$ is approximately 5000 more than $g(95)$. The gap is large, and there is no reason to believe it is going to get smaller. In fact, the opposite is true: For $k = 2$

$$n^2 - n(n-1) = n \nearrow \infty \qquad \text{as } n \nearrow \infty;$$

for $k \ge 3$

$$\begin{aligned} n^k - n(n-1)(n-2)\cdots(n-k+1) &> n(n-1)\cdots(n-k+2)[n-(n-k+1)] \\ &> 2n \nearrow \infty \qquad \text{as } n \nearrow \infty. \end{aligned}$$

So the bad news is that this kind of asymptotic approximation of n^k does not work.

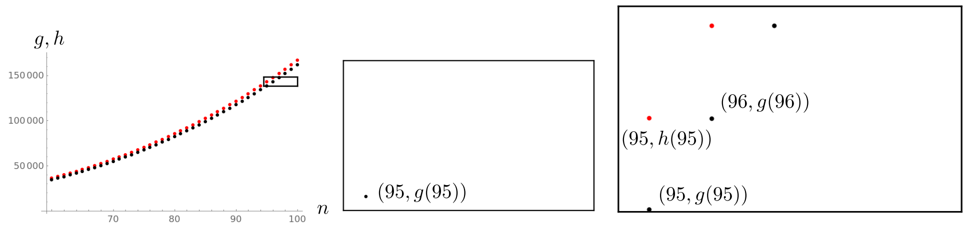

Figure 2: Plot of $g(n)$ (black) and $h(n)$ (red) for $60 \le n \le 100$ and $k = 3$ (left). A boxed region of the data points within $[94.5, 100] \times [g(95) - 100, g(95) + 1000]$ (center). A boxed region of the data points within $[94.5, 100] \times [g(95) - 100, g(95) + 10000]$(right).

The good news is that you do not need this asymptotic assertion. Here's my suggestion: When you have understood fully what I've explained above, forget about it and do two things:

1. Go back and look carefully at your original (incorrect) assertion (1). With a minor modification this can be corrected, will take the form of an acutal limit, and will define a different kind of asymptotic assertion—ultimately any asymptotic assertion does rely on some kind of limit.

2. Go back and look at the overall limit you are trying to compute and notice you do not really need (1) but rather precisely what you need is the corrected version.

Your treatment of the factor

$$\left(1 - \frac{\lambda}{n}\right)^{n-k}$$

suffers some of the same carelessness with limits. For example, the assertion that this "approaches"

$$\left(1 - \frac{\lambda}{n}\right)^{n}$$

also suggests a limiting assertion as $n \to \infty$ with an n still appearing in the limit:

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{n-k} = \left(1 - \frac{\lambda}{n}\right)^{n}.$$

This kind of assertion (with an n on the right) is never correct, but you can (almost) get away with it on this one.

2 Additional comments and suggestions

1. (typo) Your formula for f should have $e^{-\lambda}$ instead of e^{λ}.

2. For your binomial distribution, I believe you need to restrict to $\lambda < n$. I'm not sure you mentioned this. Of course, this condition will hold eventually for n large enough, but it can cause a problem plotting examples, so it might be worth mentioning.

3. It might be nice to mention that $p(k)$ for $k = 0, 1, 2, \ldots, n$ gives the values of a PMF with mean λ. You sort of flirt with this observation in your Problem 1, but it might be (also) nice to plot the points $(k, f(k)$ for some fixed λ and k on some fixed interval $0 \leq k \leq n_0$ and then superimpose the plots of $(k, p(k))$ for $k = 0, 1, 2, \ldots, n$ for several values of n (without the emphasis that n be large). See Figure 3.

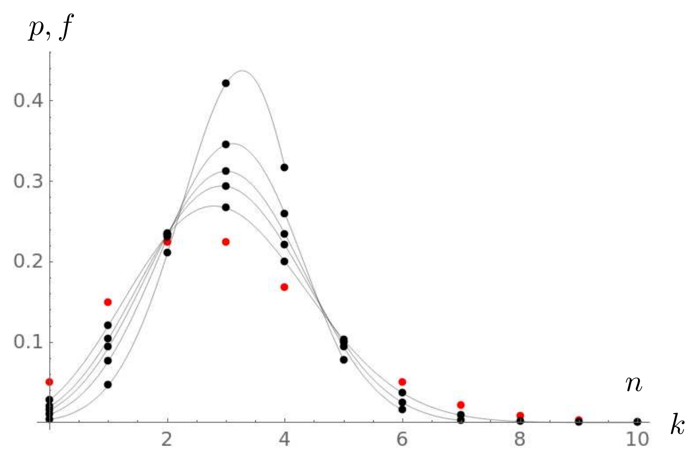

Figure 3: Plot of $f(k)$ for $\lambda = 3$ and $0 \leq k \leq 10$ (red). Superimposed plots of $p(k)$ for $\lambda = 3$, $n = 4, 5, 6, 7, 10$.