Maxim Zhilyaev David Zeber

February 23, 2016

1 recursion formular - simple version

Consider generating function for Poisson-Binomial of you case. m Bernoulli trials with success p and n Bernoulli trials with success q.

$$m + n = N \tag{1.1}$$
$$q + p = 1 \tag{1.2}$$
$$G(x) = (q + px)^m \cdot (p + qx)^n \tag{1.3}$$

The derivative of $ln(f(x)$ is given by:

$$[ln(g(x))]' = \frac{g^{`}(x)}{f(x)} = \frac{\left(\sum_{i=0}^{N} a_i x^i\right)^{`}}{\sum_{i=0}^{N} a_i x^i} = \frac{\sum_{i=1}^{N} i \cdot a_i x^{i-1}}{\sum_{i=0}^{N} a_i x^i} \tag{1.4}$$

On the other hand

$$[ln(g(x_)]^{`} = (m(q + px) + n(p + qx))^{`} = \frac{mp}{q + px} + \frac{nq}{p + qx} = \frac{xpq(m + n) + mp^2 + nq^2}{x^2pq + x(p^2 + q^2) + pq} \tag{1.5}$$

Equating both expressions we get

$$\frac{xpqN + mp^2 + nq^2}{x^2pq + x(p^2 + q^2) + pq} = \frac{\sum_{i=1}^{N} i \cdot a_i x^{i-1}}{\sum_{i=0}^{N} a_i x^i} \tag{1.6}$$
$$(xpqN + mp^2 + nq^2)(\sum_{i=0}^{N} a_i x^i) = (x^2pq + x(p^2 + q^2) + pq)(\sum_{i=1}^{N} i \cdot a_i x^{i-1}) \tag{1.7}$$

Multiplying and equating terms with same power of x we get:

$$a_i(mp^2 + nq^2) + a_{i-1}pqN = a_{i+1}pq(i + 1) + a_i(p^2 + q^2)i + a_{i-1}pq(i - 1) \tag{1.8}$$
$$a_i(mp^2 + nq^2) + a_{i-1}pqN = a_{i+1}pqi + a_i(p^2 + q^2)i + a_{i-1}pqi + (a_{i+1} - a_{i-1}) \tag{1.9}$$

Ignore the difference of $a_{i+1} - a_{i-1}$, and denote the expectation of successes as μ. Then the expression simplifies to:

$$\frac{N-i}{\frac{a_{i+1}}{a_i}i - (\frac{\mu - Npq}{pq} - \frac{p^2+q^2}{pq}i)} = \frac{a_i}{a_{i-1}} \tag{1.10}$$

$$\frac{N-i}{N-i+(\frac{a_{i+1}}{a_i}-1)i - \frac{\mu-i}{pq}} = \frac{a_i}{a_{i-1}} \tag{1.11}$$

$$\tag{1.12}$$

Denote the distance between i and μ as l. Then:

$$\mu - i = l \tag{1.13}$$

$$i = \mu - l \tag{1.14}$$

$$\frac{a_{i+1}}{a_i} = f_l \tag{1.15}$$

$$\frac{a_i}{a_{i-1}} = f_{l+1} \tag{1.16}$$

$$\frac{N-\mu+l}{N-\mu+l+(f_l-1)(\mu-l)-\frac{l}{pq}} = f_{l+1} \tag{1.17}$$

$$f_{l+1} = \frac{1}{1-\frac{1}{N-\mu+l}(\frac{l}{pq}-(f_l-1)(\mu-l))} \tag{1.18}$$

Differentiating by μ, we have:

$$F'(\mu) = -\frac{pq\left(\left(f_l-1\right)Npq-l\right)}{\left(\left(f_l-2\right)pqx+\left(n+\left(2-f_l\right)l\right)pq-l\right)^2} \tag{1.19}$$

Since the denominator is always positive, the derivative is always negative as long as for all l the inequality below holds:

$$(f_l-1)Npq > l \tag{1.20}$$

This is a magic formula. No matter what I do, and how I twist the proof, it always comes down to that inequality. I tested it empirically and it indeed works, and in fact $(f_l-1)Npq$ grows much much faster than l. But i had difficulties proving that analytically. A very interesting observation that Npq is the variance, which would suggest using Chebusheff inequality, but this, again, didn't get me anywhere. I know the inequality holds, but not sure how to prove it.

Note that when $l=0$ and $f_0=1$, $f_1=1$ for all μ. When $l=1$, f_1 is given by

$$f_1 = \frac{1}{1-\frac{1}{N-\mu+1}\frac{1}{pq}} \tag{1.21}$$

Clearly, f_1 is largest for the smallest μ which is reached when $m=0$ and the smallest when $m=N$. And we also know that f_l when $m=0$ is always greater than f_l for $m=N$, perhaps there's some way to utilize monotonicity of the f_{l+1} when f_l is fixed.