

March 2, 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 .

$$m + n = N \quad (1.1)$$

$$q + p = 1 \quad (1.2)$$

$$G(x) = (q + px)^m \cdot (p + qx)^n \quad (1.3)$$

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

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

On the other hand

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

Equating both expressions we get

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

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

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

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

$$a_{i-1}pq(N - i + 1) = a_{i+1}pq(i + 1) + a_i[(p^2 + q^2)i - (mp^2 + nq^2)] \quad (1.9)$$

$$a_{i-1}(N - i + 1) = a_{i+1}(i + 1) + a_i \frac{(p^2 + q^2)i - (mp^2 + nq^2)}{pq} \quad (1.10)$$

Denote the expectation of successes as μ , and note the obvious relationships:

$$mp^2 + nq^2 = mp^2 + (N - m)q^2 = m(p^2 - q^2) + Nq^2 = \quad (1.11)$$

$$m(p - q) + Nq - Nq + Nq^2 = \mu + Nq(q - 1) = \mu - Npq \quad (1.12)$$

$$p^2 - q^2 = (p + q)^2 - 2pq = 1 - 2pq \quad (1.13)$$

Replacing the expression above in 1.10, we have:

$$a_{i-1}(N - i + 1) = a_{i+1}(i + 1) + a_i \frac{(1 - 2pq)i - (\mu - Npq)}{pq} \quad (1.14)$$

$$a_{i-1}(N - i + 1) = a_{i+1}(i + 1) + a_i \left(N - 2i - \frac{\mu - i}{pq} \right) \quad (1.15)$$

$$\frac{a_{i-1}}{a_i}(N - i + 1) = \frac{a_{i+1}}{a_i}(i + 1) + N - 2i - \frac{\mu - i}{pq} \quad (1.16)$$

$$\frac{a_{i-1}}{a_i}(N - i + 1) = \frac{a_{i+1}}{a_i}(i + 1) + N - i + 1 - i - 1 - \frac{\mu - i}{pq} \quad (1.17)$$

$$\frac{a_{i-1}}{a_i}(N - i + 1) = N - i + 1 - \left(\frac{a_{i+1}}{a_i} - 1 \right)(i + 1) - \frac{\mu - i}{pq} \quad (1.18)$$

Rearranging the terms, we arrive to the recursive relationship between probabilistic ratios at i and $i - 1$:

$$\frac{N - i + 1}{N - i + 1 + \left(\frac{a_{i+1}}{a_i} - 1 \right)(i + 1) - \frac{\mu - i}{pq}} = \frac{a_i}{a_{i-1}} \quad (1.19)$$

NOTATION:

Denote the distance between μ and i as l . Also denote the probabilistic ratio at i as f_i . Then:

$$\mu - i = l \quad (1.20)$$

$$i = \mu - l \quad (1.21)$$

$$f_{i+1} = f_{l-1} = \frac{a_{i+1}}{a_i} \quad (1.22)$$

$$f_i = f_l = \frac{a_i}{a_{i-1}} \quad (1.23)$$

$$f_{i-1} = f_{l+1} = \frac{a_{i-1}}{a_{i-2}} \quad (1.24)$$

$$\frac{N - i + 1}{N - i + 1 + (f_{l-1} - 1)(i + 1) - \frac{l}{pq}} = \frac{1}{1 - \frac{1}{N-i+1} \left(\frac{l}{pq} - (f_{l-1} - 1)(i + 1) \right)} = f_l \quad (1.25)$$

$$\frac{N - i + 2}{N - i + 2 + (f_l - 1)i - \frac{l+1}{pq}} = \frac{1}{1 - \frac{1}{N-i+2} \left(\frac{l+1}{pq} - (f_l - 1)i \right)} = f_{l+1} \quad (1.26)$$

For some value of m denote corresponding expectation as μ_m . Denote the number of successes

equal as i_m , and a corresponding probability ratio at i_m as $f_{m,l}$:

$$\mu_m = mp + (N - m)q \quad (1.27)$$

$$i_m = \mu_m - l \quad (1.28)$$

$$f_{m,l} = \frac{a_{i_m+1}}{a_{i_m}} \quad (1.29)$$

Expressing $i_m = \mu_m - l$, allows to re-write formula 1.26 as follows:

$$\frac{1}{1 - \frac{1}{N - \mu_m + l + 2} \left(\frac{l+1}{pq} - (f_{m,l} - 1)(\mu_m - l) \right)} = f_{m,l+1} \quad (1.30)$$

Properties of $f_{0,l}$

Consider probability ratio for the case $m = 0$. Since all Bernoulli trials generate successes with probability q , the following relationships hold:

$$\mu_0 = Nq \quad (1.31)$$

$$f_{0,l} = \frac{a_i}{a_{i-1}} = \frac{\binom{N}{i} q^i p^{N-i}}{\binom{N}{i-1} q^{i-1} p^{N-i+1}} = \frac{N - i + 1}{i} \frac{q}{p} \quad (1.32)$$

Subtracting 1 from both sides of the equality above gives:

$$f_{0,l} - 1 = \frac{N - i + 1}{i} \frac{q}{p} - 1 = \frac{qN - qi + q - pi}{pi} = \frac{\mu_0 - i + q}{pi} = \frac{l + q}{p(\mu_0 - l)} \quad (1.33)$$

$$f_{0,l} - 1 = \frac{l + q}{p(\mu_0 - l)} = \frac{l + q}{pi} \quad (1.34)$$

Similarly:

$$f_{0,l-1} - 1 = \frac{l + p}{p(i + 1)} \quad (1.35)$$

We need few more expression relating differences between ratios of adjacent l :

$$f_{0,l} - f_{0,l+1} = \frac{q}{p} \left(\frac{N - i + 1}{i} - \frac{N - i}{i + 1} \right) = \frac{q}{p} \cdot \frac{N + 1}{(i + 1)i} \quad (1.36)$$

$$\frac{1}{f_{0,l+1}} - \frac{1}{f_{0,l}} = \frac{q}{p} \left(\frac{i + 1}{N - i} - \frac{i}{N - i + 1} \right) = \frac{p}{q} \cdot \frac{N + 1}{(N - i + 1)(N - i)} \quad (1.37)$$

We use this expressions in the theorem below.

THEOREM 1

For $l > \sigma$ probability ratio $f_{m,l}$ maximizes at $m = 0$.

PROOF:

Suppose that there's a particular value of m such that $f_{m,l+1} > f_{0,l+1}$ for some l and consider $f_{m,l+1}$ and $f_{0,l+1}$ for $m = m$ and $m = 0$:

$$\frac{1}{1 - \frac{1}{N - \mu_m + l + 2} \left(\frac{l+1}{pq} - (f_{m,l} - 1)(\mu_m - l) \right)} = f_{m,l+1} \quad (1.38)$$

$$\frac{1}{1 - \frac{1}{N - \mu_0 + l + 2} \left(\frac{l+1}{pq} - (f_{0,l} - 1)(\mu_0 - l) \right)} = f_{0,l+1} \quad (1.39)$$

For $f_{m,l+1} > f_{0,l+1}$ to hold the difference of denominators in the expressions for $f_{0,l+1}$ and $f_{m,l+1}$ must be positive, hence:

$$\frac{1}{N - \mu_m + l + 2} \left(\frac{l+1}{pq} - (f_{m,l} - 1)(\mu_m - l) \right) - \frac{1}{N - \mu_0 + l + 2} \left(\frac{l+1}{pq} - (f_{0,l} - 1)(\mu_0 - l) \right) > 0 \quad (1.40)$$

$$\frac{l+1}{pq} \left(\frac{1}{N - \mu_m + l + 2} - \frac{1}{N - \mu_0 + l + 2} \right) + \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l + 2} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l + 2} > 0 \quad (1.41)$$

$$\frac{l+1}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l + 2)(N - \mu_0 + l + 2)} + \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l + 2} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l + 2} > 0 \quad (1.42)$$

Suppose $f_{m,l} \geq f_{0,l}$, then we can replace $f_{m,l} - 1$ with $f_{0,l} - 1$, and the inequality should still hold since we subtract a lesser value.

$$\frac{l+1}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l)(N - \mu_0 + l)} + \frac{(f_{0,l} - 1)(\mu_m - l)}{N - \mu_m + l} - \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l} > 0 \quad (1.43)$$

$$\frac{l+1}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l + 2)(N - \mu_0 + l + 2)} + (f_{0,l} - 1) \left[\frac{\mu_0 - l}{N - \mu_0 + l + 2} - \frac{\mu_m - l}{N - \mu_m + l + 2} \right] > 0 \quad (1.44)$$

The expression inside the brackets simplifies to:

$$\frac{\mu_0 - l}{N - \mu_0 + l + 2} - \frac{\mu_m - l}{N - \mu_m + l + 2} = - \frac{(N + 2)(\mu_m - \mu_0)}{(N - \mu_m + l)(N - \mu_0 + l)} \quad (1.45)$$

From here:

$$\frac{l+1}{pq} - (f_{0,l} - 1)(N+2) > 0 \quad (1.46)$$

$$\frac{l+1}{pq} > (f_{0,l} - 1)(N+2) \quad (1.47)$$

From properties of $f_{0,l}$ we have:

$$\frac{l+1}{pq} > (f_{0,l} - 1)(N+2) \quad (1.48)$$

$$\frac{l+1}{pq} > \frac{l+q}{p(\mu_0 - l)}(N+2) > \frac{l+q}{p(\mu_0 - l)}N \quad (1.49)$$

$$(l+1)(\mu_0 - l) > Nq(l+q) \quad (1.50)$$

$$(l+1)(\mu_0 - l) > \mu_0(l+q) \quad (1.51)$$

After simplification we have:

$$-l^2 - l + Npq > 0 \quad (1.52)$$

Since the standard deviation is $\sigma = \sqrt{Npq}$, this condition clearly does not hold for $l > \sigma$. Which proves an interesting point.

When l becomes sufficiently large, $f_{m,l+1}$ could only become greater than $f_{0,l+1}$ if at the previous iteration $f_{m,l}$ is less than $f_{0,l}$. Hence, there are only two outcomes: either $f_{m,l+1}$ is always less than $f_{0,l+1}$, or it oscillates. On one iteration $f_{m,l}$ is less than $f_{0,l}$ and on the next iteration it becomes greater than $f_{0,l+1}$. We prove by contradiction that the second outcome is impossible.

Assume that it's the case, then we must observe the following:

$$f_{m,l-1} > f_{0,l-1} \quad (1.53)$$

$$f_{m,l} < f_{0,l} \quad (1.54)$$

$$f_{m,l+1} > f_{0,l+1} \quad (1.55)$$

Note that last two conditions imply that:

$$\frac{1}{f_{m,l}} - \frac{1}{f_{m,l+1}} > \frac{1}{f_{0,l}} - \frac{1}{f_{0,l+1}} > \frac{p}{q} \cdot \frac{N+1}{(N-i_0+1)(N-i_0)} \quad (1.56)$$

Expressing the difference of quotients using formulas 1.26 and 1.25, we have:

$$\frac{1}{N - i_m + 2} \left(\frac{l+1}{pq} - (f_{m,l} - 1)i_m \right) - \frac{1}{N - i_m + 1} \left(\frac{l}{pq} - (f_{m,l-1} - 1)(i_m + 1) \right) = \frac{1}{f_{m,l}} - \frac{1}{f_{m,l+1}} \quad (1.57)$$

$$\frac{1}{N - i_m + 1} \left(\frac{l+1}{pq} - (f_{m,l} - 1)i_m - \frac{l}{pq} + (f_{m,l-1} - 1)(i_m + 1) \right) > \frac{p}{q} \cdot \frac{N+1}{(N - i_0 + 1)(N - i_0)} \quad (1.58)$$

$$\frac{1}{N - i_m + 1} \left(\frac{1}{pq} + (f_{m,l-1} - 1) - (f_{m,l} - f_{m,l-1})i_m \right) > \frac{p}{q} \cdot \frac{N+1}{(N - i_0 + 1)(N - i_0)} \quad (1.59)$$

Note the the oscillating conditions mean the following holds:

$$f_{m,l-1} - 1 > f_{0,l-1} - 1 \quad (1.60)$$

$$f_{m,l} - f_{m,l-1} < f_{0,l} - f_{0,l-1} \quad (1.61)$$

Hence, we can replace $f_{m,l-1}$ with $f_{0,l-1}$ and that will decrease the left part of the inequality, because we add a lesser value. And by the same argument we can replace $f_{m,l} - f_{m,l-1}$ with $f_{0,l} - f_{0,l-1}$, and that will further decrease the inequality since we subtract a greater value. From here:

$$\frac{1}{pq} + (f_{m,l-1} - 1) - (f_{m,l} - f_{m,l-1})i_m > \frac{1}{pq} + (f_{0,l-1} - 1) - (f_{0,l} - f_{0,l-1})i_m \quad (1.62)$$

Using formulas 1.35 and 1.36, we arrive:

$$\frac{1}{pq} + (f_{0,l-1} - 1) - (f_{0,l} - f_{0,l-1})i_m = \frac{1}{pq} + \frac{l+p}{p(i_0+1)} - \frac{q}{p} \cdot \frac{N+1}{(i_0+1)i_0} \cdot i_m = \quad (1.63)$$

$$\frac{1}{pq} + \frac{l-p-qN-q}{p(i+1)} = \frac{1}{pq} - \frac{\mu_0 - l - p + q}{p(i+1)} \quad (1.64)$$

The last inequality contradicts LEMMA 2, which means that $f_{m,l} < f_{0,l}$.

On the other hand, $f_{m,l} > f_{N,l}$. Since

$$f_{m,l+1} > f_{0,l+1} > f_{N,l+1}$$

We have:

$$\frac{1}{N - \mu_m + l} \left(\frac{l}{pq} - (f_{m,l} - 1)(\mu_m - l) \right) - \frac{1}{N - \mu_N + l} \left(\frac{l}{pq} - (f_{N,l} - 1)(\mu_N - l) \right) > 0 \quad (1.65)$$

$$\frac{l}{pq} \frac{\mu_m - \mu_N}{(N - \mu_m + l)(N - \mu_N + l)} + \frac{(f_{N,l} - 1)(\mu_N - l)}{N - \mu_N + l} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l} > 0 \quad (1.66)$$

$$\frac{(f_{N,l} - 1)(\mu_N - l)}{N - \mu_N + l} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l} - \frac{l}{pq} \frac{\mu_N - \mu_m}{(N - \mu_m + l)(N - \mu_N + l)} > 0 \quad (1.67)$$

Suppose $f_{m,l} \leq f_{N,l}$, then we can replace $f_{m,l} - 1$ with $f_{N,l} - 1$, and the inequality should still hold since we subtract a lesser value. Then

$$\frac{l}{pq} \frac{\mu_m - \mu_N}{(N - \mu_m + l)(N - \mu_N + l)} - (f_{N,l} - 1) \frac{N(\mu_N - \mu_m)}{(N - \mu_m + l)(N - \mu_N + l)} > 0 \quad (1.68)$$

$$\frac{l}{pq} - (f_{N,l} - 1)N > 0 \quad (1.69)$$

$$(f_{N,l} - 1) < \frac{l}{pqN} \quad (1.70)$$

Which again contradicts LEMMA 2, and we must conclude that:

$$f_{0,l} > f_{m,l} > f_{N,l}$$

Suppose $f_l^x \geq f_l^0$, then we can replace $f_l^x - 1$ with $f_l^0 - 1$ and the inequality should still hold since we reduce the left part.

$$\frac{(f_0^x - 1)(\mu_x - l)}{N - \mu_x + l} - \frac{(f_l^0 - 1)(\mu_0 - l)}{N - \mu_0 + l} < \frac{l}{pq} \frac{\mu_x - \mu_0}{(N - \mu_x + l)(N - \mu_0 + l)} \quad (1.71)$$

$$(f_l^0 - 1) \left[\frac{\mu_x - l}{N - \mu_x + l} - \frac{\mu_0 - l}{N - \mu_0 + l} \right] < \frac{l}{pq} \frac{\mu_x - \mu_0}{(N - \mu_x + l)(N - \mu_0 + l)} \quad (1.72)$$

The expression inside the brackets simplifies to:

$$\frac{\mu_x - l}{N - \mu_x + l} - \frac{\mu_0 - l}{N - \mu_0 + l} = \frac{N(\mu_x - \mu_0)}{(N - \mu_x + l)(N - \mu_0 + l)} \quad (1.73)$$

From here:

$$(f_l^0 - 1) \frac{N(\mu_x - \mu_0)}{(N - \mu_x + l)(N - \mu_0 + l)} < \frac{l}{pq} \frac{\mu_x - \mu_0}{(N - \mu_x + l)(N - \mu_0 + l)} \quad (1.74)$$

$$(f_l^0 - 1)Npq < l \quad (1.75)$$

$$(f_l^0 - 1) < \frac{l}{Npq} \quad (1.76)$$

Since this inequality holds for all l , we arrived at contradiction. Then f_x must be strictly less than f_0 . This may be sufficient result on its own.

$$\frac{1}{N - \mu_m + l} \left(\frac{l}{pq} - (f_{m,l} - 1)(\mu_m - l) \right) - \frac{1}{N - \mu_0 + l} \left(\frac{l}{pq} - (f_{0,l} - 1)(\mu_0 - l) \right) > 0 \quad (1.77)$$

$$\frac{1}{N - \mu_m + l} \frac{l}{pq} - \frac{1}{N - \mu_0 + l} \frac{l}{pq} + \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l} > 0 \quad (1.78)$$

$$\frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l} > - \frac{l}{pq} \frac{\mu_m - \mu_0}{((N - \mu_m + l)(N - \mu_0 + l))} \quad (1.79)$$

$$\frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l} - \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l} < \frac{l}{pq} \frac{\mu_m - \mu_0}{((N - \mu_m + l)(N - \mu_0 + l))} \quad (1.80)$$

Suppose $f_{m,l} \geq f_{0,l}$, then we can replace $f_{m,l} - 1$ with $f_{0,l} - 1$ and the inequality should still hold since we reduce the left part.

$$\frac{(f_0^x - 1)(\mu_m - l)}{N - \mu_m + l} - \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l} < \frac{l}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l)(N - \mu_0 + l)} \quad (1.81)$$

$$(f_{0,l} - 1) \left[\frac{\mu_m - l}{N - \mu_m + l} - \frac{\mu_0 - l}{N - \mu_0 + l} \right] < \frac{l}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l)(N - \mu_0 + l)} \quad (1.82)$$

The expression inside the brackets simplifies to:

$$\frac{\mu_m - l}{N - \mu_m + l} - \frac{\mu_0 - l}{N - \mu_0 + l} = \frac{N(\mu_m - \mu_0)}{(N - \mu_m + l)(N - \mu_0 + l)} \quad (1.83)$$

From here:

$$(f_{0,l} - 1) \frac{N(\mu_m - \mu_0)}{(N - \mu_m + l)(N - \mu_0 + l)} < \frac{l}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l)(N - \mu_0 + l)} \quad (1.84)$$

$$(f_{0,l} - 1)Npq < l \quad (1.85)$$

$$(f_{0,l} - 1) < \frac{l}{Npq} \quad (1.86)$$