

March 8, 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  $\mu$ , 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  $\mu$  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  $\mu_m$ . 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:

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

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

### THEOREM 1

For  $l > \sigma$  probability ratio  $f_{0,l+1}$  at  $l + 1$  is higher then probability ratio  $f_{m,l}$  for any  $m$ .

#### 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  $m = 0$ :

$$\frac{1}{1 - \frac{1}{N - \mu_m + l + 2} \left( \frac{l+1}{pq} - (f_{m,l} - 1)(\mu_m - l) \right)} = f_{m,l+1} \quad (1.35)$$

$$\frac{1}{1 - \frac{1}{N - \mu_0 + l + 2} \left( \frac{l+1}{pq} - (f_{0,l} - 1)(\mu_0 - l) \right)} = f_{0,l+1} \quad (1.36)$$

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

$$\frac{1}{N - \mu_m + l + 2} \left( \frac{l+1}{pq} - (f_{m,l} - 1)(\mu_m - l) \right) - \frac{1}{N - \mu_0 + l + 2} \left( \frac{l+1}{pq} - (f_{0,l} - 1)(\mu_0 - l) \right) > 0 \quad (1.37)$$

$$\frac{l+1}{pq} \left( \frac{1}{N - \mu_m + l + 2} - \frac{1}{N - \mu_0 + l + 2} \right) + \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l + 2} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l + 2} > 0 \quad (1.38)$$

$$\frac{l+1}{pq} \frac{\mu_m - \mu_0}{(N - \mu_m + l + 2)(N - \mu_0 + l + 2)} + \frac{(f_{0,l} - 1)(\mu_0 - l)}{N - \mu_0 + l + 2} - \frac{(f_{m,l} - 1)(\mu_m - l)}{N - \mu_m + l + 2} > 0 \quad (1.39)$$

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.

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

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

The expression inside the brackets simplifies to:

$$\frac{\mu_0 - l}{N - \mu_0 + l + 2} - \frac{\mu_m - l}{N - \mu_m + l + 2} = -\frac{(N+2)(\mu_m - \mu_0)}{(N - \mu_m + l)(N - \mu_0 + l)} \quad (1.42)$$

From here:

$$\frac{l+1}{pq} - (f_{0,l} - 1)(N+2) > 0 \quad (1.43)$$

$$\frac{l+1}{pq} > (f_{0,l} - 1)(N+2) \quad (1.44)$$

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

$$\frac{l+1}{pq} > (f_{0,l} - 1)(N+2) \quad (1.45)$$

$$\frac{l+1}{pq} > \frac{l+q}{p(\mu_0 - l)}(N+2) > \frac{l+q}{p(\mu_0 - l)}N \quad (1.46)$$

$$(l+1)(\mu_0 - l) > Nq(l+q) \quad (1.47)$$

$$(l+1)(\mu_0 - l) > \mu_0(l+q) \quad (1.48)$$

After simplification we have:

$$-l^2 - l + Npq > 0 \tag{1.49}$$

Since the standard deviation is  $\sigma = \sqrt{Npq}$ , this condition clearly does not hold for  $l > \sigma$ . Which proves an important point. When  $l$  exceeds  $\sigma$ ,  $f_{m,l+1}$  could only become greater than  $f_{0,l+1}$  if at the previous iteration  $f_{m,l}$  is less than  $f_{0,l}$ . And conversely, if  $f_{m,l} > f_{0,l}$  at  $l$ , the next iteration  $f_{m,l+1} < f_{0,l+1}$ . We now can prove the main statement of the theorem.

Choose a particular value for  $l$  such that  $l > \sigma$ . For those values of  $m$  such that  $f_{0,l} > f_{m,l}$  the theorem holds already, since  $f_{0,l+1} > f_{0,l}$ . Now consider those  $m$  for which  $f_{0,l} < f_{m,l}$  at  $l$ . In the next iteration the probability ratio  $f_{m,l+1}$  for such  $m$  has to become less than  $f_{0,l+1}$ . Hence:

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

Which proves the theorem.