March 3, 2016

1 conditioning

Consider the original collection D consisting of m ones and $N-m$ zeros. Denote probability of number of successes for that collection as $P(S|D)$. The probability ratio at s is given by:

$$R_s = \frac{P(s|D)}{P(s-1|D)}$$

Denote expectation of s as μ:

$$\mu = mp + (N-m)q$$

It's known that for all $s < \mu$, the ratio R_s is greater than 1 and increasing:

$$R_{s-1} > R_s$$

Create two collections by adding to D one 1 and one 0. Call them D_1 and D_0 respectively. The probability of observing s from D_1 the is given by:

$$P(s|D_1) = pP(s-1|D) + qP(s|D)$$

Similarly for the second collection (with extra 0):

$$P(s|D_0) = qP(s-1|D) + pP(s|D)$$

Now consider the probability ratio for the first collections at some s:

$$R_s(D_1) = \frac{pP(s-1|D) + qP(s|D)}{pP(s-2|D) + qP(s|D)} = \frac{p + q\frac{P(s|D)}{P(s-1|D)}}{p\frac{P(s-2|D)}{P(s-1|D)} + q} = \frac{p + qR_s}{p\frac{1}{R_{s-1}} + q} \tag{1.1}$$

Similarly for the second collection, we have

$$R_s(D_0) = \frac{q + pR_s}{q\frac{1}{R_{s-1}} + p} \tag{1.2}$$

Lemma 1.

$$R_s(D_1) > R_s(D_0) \text{ for } s < \mu$$

proof:

$$\frac{p + qR_s}{p\frac{1}{R_{s-1}} + q} > \frac{q + pR_s}{q\frac{1}{R_{s-1}} + p} \tag{1.3}$$

$$(p + qR_s)(q\frac{1}{R_{s-1}} + p) - (q + pR_s)(p\frac{1}{R_{s-1}} + q) > 0 \tag{1.4}$$

$$(p^2 - q^2)(1 - \frac{R_s}{R_{s-1}}) > 0 \tag{1.5}$$

The above holds because $p > q$ and $R_{s-1} > R_s$.

Lemma 2.

$$R_{s-1}(D_0) > R_s(D_1) \text{ for } s < \mu$$

This lemma essentially says that if we step one point to the left of s, the ratio for the distribution with extra 0 is always greater.

proof:

$$\frac{q + pR_{s-1}}{q\frac{1}{R_{s-2}} + p} > \frac{p + qR_s}{p\frac{1}{R_{s-1}} + q} \tag{1.6}$$

$$(q + pR_{s-1})(p\frac{1}{R_{s-1}} + q) - (p + qR_s)(q\frac{1}{R_{s-2}} + p) > 0 \tag{1.7}$$

$$\frac{qp}{R_{s-1}} + q^2 + p^2 + pqR_{s-1} - \frac{qp}{R_{s-2}} - p^2 - q^2\frac{R_s}{R_{s-2}} - qpR_s > 0 \tag{1.8}$$

$$qp(\frac{1}{R_{s-1}} - \frac{1}{R_{s-2}}) + qp(R_{s-1} - R_s) + q^2(1 - \frac{R_s}{R_{s-2}}) > 0 \tag{1.9}$$

Each expression in the parenthesis is greater than 0, hence the above inequality holds.

Theorem 1

Consider different collections D_m where m represents number of ones. This collections correspond to different distributions of number of successes S with respective expectations μ_m. Consider values of s_m equidistant from each respective μ_m by same number of steps l.

$$s_m = \mu_m - l$$

Denote $R_{m,l}$ as probability ratio at $s_m = \mu_m - l$ for specific m.

We should be able to show that:

$$R_{0,l+1} > R_{m,l} \text{ for any } m \text{ and } l$$

Because $s_{0,l+1}$ is the smallest value of s for given m and l. But the discreetness is a problem here, because how do we round s_m? Need your advice.