stichting mathematisch centrum

AFDELING MATHEMATISCHE STATISTIEK SW 56/78 APRIL
(DEPARTMENT OF MATHEMATICAL STATISTICS)

J.M. BUHRMAN

INEQUALITIES IN DISCRETE DISTRIBUTIONS

Preprint

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

AMS(MOS) subject classification scheme (1970): 60E05, 26A86

Inequalities in discrete distributions *)

by

J.M. Buhrman

SUMMARY

In this report sufficient conditions are derived for the implication

$$F_1(x) \leq F_2(x) \Rightarrow F_1(x-j) \leq F_2(x-j) \qquad \text{for all } j > 0$$

where F_1 and F_2 are discrete distribution functions with jumps on $\mathbb{Z}$. Results are given for pairs of - possibly shifted - distribution functions of binomial, Poisson, hypergeometric and negative binomial type.

KEY WORDS & PHRASES: *discrete distribution functions, inequalities*

---

*) This report will be submitted for publication elsewhere.

1. INTRODUCTION

In the search for a test concerning the ratio of two Poisson means, the question arose whether two binomial distribution functions F and G have the property (x and y integer)

$$F(x) \le G(x+y) \Rightarrow F(x-1) \le G(x+y-1). \tag{1.1}$$

This was a reason to investigate what conditions are sufficient for this kind of implication, not only when pairs of binomial distributions are involved, but in the case of other types of integer valued distributions as well. Notice, that, with $F_1(x) \overset{\text{def}}{=} F(x)$ and $F_2(x) \overset{\text{def}}{=} G(x+y)$, the above implication can be studied by considering

$$D(x) \overset{\text{def}}{=} F_1(x) - F_2(x) \tag{1.2}$$

$$p_i(x) \overset{\text{def}}{=} F_i(x) - F_i(x-1) \tag{1.3}$$

$$d(x) \overset{\text{def}}{=} p_1(x) - p_2(x) \tag{1.4}$$

A function f(x) is said to have at least one change of sign on an interval [a,b] (a<b) if $f(i)f(j) < 0$ for some i,j with $a \le i < j \le b$ *).

Some remarks will be made with respect to the relation between changes of sign of D(x) and changes of sign of d(x). We assume that a finite partition $\mathbb{Z} = T_1 \cup T_2 \cup \ldots \cup T_k$ exists, with

$$T_1 = \{x\colon d(h)d(j) \ge 0 \quad \text{for all } h,j \le x\} \tag{1.5}$$

$$T_i = \{x\colon x > \max T_{i-1},\ d(h)d(j) \ge 0 \quad \text{for } \max T_{i-1} + 1 \le h,j \le x\} \tag{1.6}$$

---

*) All variables indicated by x,y,a,b,c,i,j,h,k,m,n,r,t with or without subscripts, will be understood to be integers.

The assumption of such a finite partition is a restriction, but in all cases in which we are interested, it is fulfilled.

The maximum number of changes of sign of $D(x)$ is $k-2$, which can be seen as follows. Let

(1.7.a) $\quad b_i = \max T_i.$

Then

(1.7.b) $\quad T_1 = (-\infty\ ,\ b_1]$

(1.7.c) $\quad T_i = [b_{i-1}+1\ ,\ b_i] \quad i = 2,3,\ldots,k-1$

(1.7.d) $\quad T_k = [b_{k-1}+1\ ,\ \infty).$

Notice, that $D(x) = \sum_{j\le x} d(j)$ and that $D(-\infty) = d(-\infty) = 0$.
Of the two possibilities: $d(j) \ge 0$ for all $j \le b_1$ and $d(j) \le 0$ for all $j \le b_1$, the second one is chosen to be considered. Then $D(b_1) < 0$ and $D(x) \le 0$ for all $x \le b_1$. Now $d(j) \ge 0$ for $b_1 < j \le b_2$, so $D(x) = D(b_1) + \sum_{b_1<j\le x} d(j)$ is non-decreasing for $b_1 \le x \le b_2$. Similarly, $D(x)$ is non-decreasing for $b_3 \le x \le b_4$, $b_5 \le x \le b_6$, etc. and non-increasing for $b_2 \le x \le b_3$, $b_4 \le x \le b_5$, etc. So any of these intervals gives at most one change of sign of $D(x)$. Since no changes of sign occur for $x \le b_1$ and $x \ge b_{k-1}$, the maximum number of changes of sign of $D(x)$ is $k-2$. Moreover, it should be noticed, that if $D(x)$ has no change of sign for $b_1 \le x \le b_2$, then $D(b_2) \le 0$ and $D(x)$ has no change of sign for $b_2 \le x \le b_3$ either, etc. This means that the number of changes of sign of $D(x)$ equals $k-2i$ for some positive integer i. In the following applications $k = 3$ frequently occurs.

2. SOME LEMMAS

First we define

(2.1) $\quad V_i = \{x\colon p_i(x) > 0\}$ (support of F_i) $\quad i = 1,2$

(2.2) $a_i = \inf V_i;\ c_i = \sup V_i \qquad i = 1,2$

(2.3) $r(x) = p_1(x)/p_2(x) \qquad \text{for } x \in V_1 \cap V_2$

(2.4) $q(x) = r(x)/r(x-1) \qquad \text{for } x \text{ with } x-1,\ x \in V_1 \cap V_2$

V_1 and V_2 will be supposed coherent, that is such that $x,y \in V_i$ implies $k \in V_i$ if $x \le k \le y$ *). This condition on V_1 and V_2 is obviously fulfilled in the applications studied below. Now five cases are distinguished.

1. $a_2 < a_1 < c_1 < c_2$
2. $a_2 = a_1 < c_1 < c_2$
3. $a_2 < a_1 < c_1 = c_2$
4. $a_2 = a_1 < c_1 = c_2$
5. $a_1 < a_2 < c_1 < c_2$

Any of the remaining cases is reduced to one of the cases 1 to 5 after interchanging F_1 and F_2. Those cases will not be considered explicitly here, but obviously their treatment is similar to the treatment of the cases 1 to 5.

The implication in the summary can be studied by considering for all x

(2.5) $F_1(x) \le F_2(x) \quad \Rightarrow \quad F_1(x-1) \le F_2(x-1).$

LEMMA 2.1. (a) *The number of changes of sign of* D(x) *in case* 1 *is odd.*
(b) *The number of changes of sign of* D(x) *in case* 5 *is even.*

PROOF.
(a) $D(a_1-1) < 0,\ D(c_1+1) > 0$
(b) $D(a_2-1) > 0,\ D(c_1+1) > 0.$ □

---

*) See note on page 2.

LEMMA 2.2. *If implication* (2.5) *holds for all* x, *then the number of changes of sign of* D(x) *is* 0 *or* 1.

PROOF. If D(x) has two (or more) changes of sign, then there exist i,j and k (i<j<k) with either D(i), D(k) > 0 and D(j) < 0, or D(k), D(k) < 0 and D(j) > 0. This contradicts (2.5) for at least one x. □

COROLLARY 2.1. *Implication* (2.5) *holds for all* x *in case* 5 *if and only if* $F_1(x) > F_2(x)$ *for all* x.

For this reason case 5 will not be considered any furhter. Therefore we can assume $V_1 \subset V_2$ from now on.

REMARK 2.1. If $F_1(x) > F_2(x)$ for all x, the implication (2.5) holds trivially, the left part not being fulfilled for any x.

REMARK 2.2. In cases 2 to 4 the parity of the number of changes of sign of D(x) depends on the sign of $D(b_1)$ and $D(b_{k-1})$. In case 2 $D(b_{k-1}) > 0$, in case 3 $D(b_1) < 0$, in case 4 both $D(b_1)$ and $D(b_{k-1})$ can be positive or negative. Notice, that $D(b_1)$ and $D(b_{k-1})$ cannot be zero, since

$$D(b_1) = \sum_{j \le b_1} d(j)$$

$$D(b_{k-1}) = \sum_{j \le b_{k-1}} d(j) = - \sum_{j > b_{k-1}} d(j).$$

LEMMA 2.3. *If there exists an* m *such that* r(x) *is monotone non-decreasing for* $a_1 \le x \le m$ *and monotone non-increasing for* $m \le x \le c_1$, *then* (2.5) *holds vor all* x.

PROOF. Notice, that $r(m) \ge 1$, since $F_1(c_1) - F_1(a_1-1) = 1$ and $F_2(c_1) + - F_2(a_1-1) \le 1$, so $\sum_{j=a_1}^{c_1} d(j) \ge 0$, and $d(j) \ge 0$ for at least one j, say k. Then $r(m) \ge r(k) \ge 1$. The only case in which $r(m) = 1$, is the trivial case that $F_1(x) = F_2(x)$ for all $x \in V_1 = V_2$. Therefore $r(m) > 1$ will be assumed. Three possibilities remain (see definitions (1.5) and (1.6)).

(a) $V_2 = T_1 \cup T_2 \cup T_3$, $m \in T_2$
(b) $V_2 = T_1 \cup T_2$, $m \in T_1$
(c) $V_2 = T_1 \cup T_2$, $m \in T_2$

In case 1 the only possibility is (a). In case 2 both (a) and (b) are possible, in case 3 both (a) and (c) are possible, and in case 4 no possibility can be excluded in advance. If (a) is the case, we have $d(x) \le 0$ on T_1 and T_3, and $d(x) \ge 0$ on T_2. Thus $D(x) > 0$ on T_3, $D(x) < 0$ on T_1 and $D(x)$ has exactly one change of sign on T_2. If (b) or (c) is the case $D(x)$ has no change of sign, and the set $\{x: D(x) \neq 0\}$ is coherent. □

LEMMA 2.4. *If* $q(x)/q(x-1) \le 1$ *for all* x *with* $x-2, x \in V_1$, *then* (2.5) *holds for all* x.

PROOF.

$$q(x)/q(x-1) \le 1 \quad \text{for all } x$$

is equivalent to

$$\frac{r(x)}{r(x-1)} \le \frac{r(x-1)}{r(x-2)} \quad \text{for all } x$$

which implies

$$r(x-1) \le r(x-2) \;\Rightarrow\; r(x) \le r(x-1) \qquad \text{for all } x.$$

This means that $r(x)$ has the form as described in lemma 2.3, by which then (2.5) holds for all x. □

The following lemma will appear useful in several cases where $q(x)/q(x-1) \le 1$ must be proved.

LEMMA 2.5. *For* $a,b,c,d > 1$

(a) $\frac{a-1}{a}\frac{b}{b-1} \le 1 \iff a \le b$

(b) $\frac{a-1}{a}\frac{b}{b-1}\frac{c-1}{c}\frac{d}{d-1} \le 1 \iff \frac{ac}{bd} \le \frac{a+c-1}{b+d-1}$

(c) *Each of the following conditions* (2.6) *to* (2.8) *is sufficient for the inequalities in* (b).

(2.6) $a \le b$ and $c \le d$

(2.7) $ac \le bd$ and $b + d \le a + c$

(2.8) $\max(0,a-b) \le d - c$ and

$[(c \le a-1 \text{ and } d-1 \le b) \text{ or } (c \le b \text{ and } d \le a)]$.

PROOF. (a) and (b) are obvious. The sufficiency of (2.6) and (2.7) follows by application of (a) and (b) respectively. Further,

$$\frac{a-1}{a}\,\frac{b}{b-1}\,\frac{c-1}{c}\,\frac{d}{d-1} = \frac{1-\dfrac{d-c}{c(d-1)}}{1-\dfrac{a-b}{(a-1)b}}$$

which is ≤ 1 if $d - c \ge 0$ and $d - c \ge a - b$ and $[(c \le a-1 \text{ and } d-1 \le b) \text{ or } (c \le b \text{ and } d \le a)]$. □

3. RESULTS IN CLASSES OF EXPONENTIAL FAMILIES

Let a class, C, of distributions be given by their probability functions.

(3.1) $$p_{\eta,\theta}(x) = f(x,\eta)e^{xv(\eta,\theta)}g(\eta,\theta) \qquad \eta \in \Pi,\ \theta \in \Theta.$$

Any subset $C_\eta \subset C$ for fixed η (which may be a vector as may be θ) is an exponential family. Probability functions of binomial, Poisson, hypergeometric and negative binomial distributions have this form, as can be seen in table 3.1. The form (3.1) is useful, because the crucial quantity $q(x)/q(x-1)$ from lemma 2.4 will turn out to depend on $f(x,\eta)$ only. Let

(3.2) $$F_{\eta,\theta}(x) = \sum_{j\le x} p_{\eta,\theta}(j).$$

Consider the implication

(3.3) $\quad F_{\eta,\theta}(x) \le F_{\zeta,\xi}(x+y) \Rightarrow F_{\eta,\theta}(x-j) \le F_{\zeta,\xi}(x+y-j)$ for all $j \ge 0$.

Define

(3.4.a.) $\quad F_1(x) \overset{def}{=} F_{\eta,\theta}(x)$

(3.4.b.) $\quad F_2(x) \overset{def}{=} F_{\zeta,\xi}(x+y)$.

Now all definitions (1.2) to (1.7) and (2.1) to (2.4) are meaningful here, and

(3.5) $$r(x) = \frac{p_1(x)}{p_2(x)} = \frac{f(x,\eta)\ e^{xv(\eta,\theta)}\ g(\eta,\theta)}{f(x+y,\xi)\ e^{(x+y)v(\zeta,\xi)}\ g(\zeta,\xi)}$$

(3.6) $$\frac{r(x)}{r(x-1)} = \frac{f(x,\eta)\ f(x+y-1,\xi)\ e^{v(\eta,\theta)}}{f(x+y,\xi)\ f(x-1,\eta)\ e^{v(\zeta,\xi)}}$$

(3.7) $$\frac{q(x)}{q(x-1)} = \frac{f(x,\eta)\ f(x+y-1,\xi)^2\ f(x-2,\eta)}{f(x+y,\xi)\ f(x-1,\eta)^2\ f(x+y-2,\xi)}$$

In table 3.2 the ratio (3.7) is given for the four classes of table 3.1.

TABLE 3.1. The probability function of the binomial, Poisson, hypergeometric and negative binomial distribution in form (3.1).

$p_{\eta,\theta}(x)$	$f(x,\eta)$	$e^{x\,v(\eta,\theta)}$	$g(\eta,\theta)$	η	θ
$\binom{n}{x}p^x(1-p)^{n-x}$	$\binom{n}{x}$	$e^{x \ln \frac{p}{1-p}}$	$(1-p)^n$	n	p
$e^{-\mu}\frac{\mu^x}{x!}$	$\frac{1}{x!}$	$e^{x \ln\mu}$	$e^{-\mu}$	-	μ
$\frac{\binom{n}{x}\binom{t-n}{r-x}}{\binom{t}{r}}$	$\frac{\binom{n}{x}\binom{t-n}{r-x}}{\binom{t}{r}}$	1	1	(t,n,r)	-
$\binom{x-1}{n-1}p^n(1-p)^{x-n}$	$\binom{x-1}{n-1}$	$e^{x \ln(1-p)}$	$\left(\frac{p}{1-p}\right)^n$	n	p

TABLE 3.2. $q(x)/q(x-1)$ in binomial, Poisson, hypergeometric and negative binomial classes.

class	parameters	$\frac{q(x)}{q(x-1)}$
binomial	n_1, p_1 n_2, p_2	$\left(\frac{x-1}{x}\frac{x+y}{x+y-1}\right)\left(\frac{n_1-x+1}{n_1-x+2}\frac{n_2-x-y+2}{n_2-x-y+1}\right)$
Poisson	μ_1 μ_2	$\frac{x-1}{x}\frac{x+y}{x+y-1}$
Hyper-geometric	t_1, n_1, r_1 t_2, n_2, r_2	$\left(\frac{x-1}{x}\frac{x+y}{x+y-1}\right)\left(\frac{n_1-x+1}{n_1-x+2}\frac{n_2-x-y+2}{n_2-x-y+1}\right)\times$ $\times\left(\frac{r_1-x+1}{r_1-x+2}\frac{r_2-x-y+2}{r_2-x-y+1}\right)\left(\frac{t_1-n_1-r_1+x-1}{t_1-n_1-r_1+x}\frac{t_2-n_2-r_2+x+y}{t_2-n_2-r_2+x+y-1}\right)$
Negative binomial	n_1, p_1 n_2, p_2	$\frac{x+y-2}{x+y-1}\frac{x-1}{x-2}\frac{x-n_1-1}{x-n_1}\frac{x+y-n_2}{x+y-n_2-1}$

THEOREM 3.1. *Implication* (3.3) *holds*

a. *in the binomial case if* $0 \le y \le n_2-n_1$

b. *in the Poisson case if* $y \ge 0$

c. *in the hypergeometric case if* $0 \le y \le n_2-n_1$ *and*
$t_1-n_1-r_1-(t_2-n_2-r_2) \le y \le r_2-r_1$

d1. *in the negative binomial case if* $n_2-n_1 \le y$ *and* $n_2 - n_1 \le 0$

d2. *in the negative binomial case if* $p_1 = p_2$ *and* $n_2 \ge n_1$

PROOF. Cases a to c are dealt with by showing that $q(x)/q(x-1) \le 1$ for all x by lemma 2.5(a). In case d1 we have $q(x)/q(x-1) \le 1$ for all $x \ge n_1+2$ by (2.8) in lemma 2.5 with $a = x + y - 1$, $b = x - 1$, $c = x - n_1$, $d = x + y - n_2$. Lemma 2.4 completes the proof in the cases a to d1. In case d2 the shape of $r(x)$ is considered.

$$q(x) = 1 + \frac{(n_1-n_2)(x-1) + y(n_1-1)}{(x-n_1)(x+y-1)}$$

So

$$r(x) > r(x-1) \quad \text{if } x - 1 < \frac{y(n_1-1)}{n_2-n_1}$$

$$r(x) = r(x-1) \quad \text{if } x - 1 = \frac{y(n_1-1)}{n_2-n_1}$$

$$r(x) < r(x-1) \quad \text{if } x - 1 > \frac{y(n_1-1)}{n_2-n_1}$$

Then lemma 2.3 completes the proof. □

4. ACKNOWLEDGEMENT

This report was written after valuable and stimulating discussions with Prof.dr. J. Hemelrijk and Prof.dr. W. Whitt. It was the latter who initially gave the proof of the implication for the binomial case with $p_1 = p_2$. He recognized the importance of the log-concavity of the function r(x) (lemma 2.4). This property turned out to be crucial in most of the proofs.