stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 87/82 SEPTEMBER
(DEPARTMENT OF MATHEMATICAL STATISTICS)

W.R. VAN ZWET

AN INEQUALITY FOR RANDOM REPLACEMENT SAMPLING PLANS

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, 413 Kruislaan, 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.).

1980 Mathematics subject classification: Primary: 62D05
Secondary: 62B15

An inequality for random replacement sampling plans *)

by

W.R. van Zwet

ABSTRACT

In this paper a conjecture of Karlin concerning random replacement sampling plans is discussed.

KEY WORDS & PHRASES: *random replacement sampling plans, sampling without replacement, sampling with replacement, inequalities*

---

*) This report will be submitted for publication elsewhere.

1. INTRODUCTION

From a population $\Omega = \{1,2,\ldots,N\}$ a sample $I = (I_1, I_2, \ldots, I_n)$ of size $n \le N$ is drawn by means of a random replacement scheme as follows. Let $\pi = (\pi_1, \pi_2, \ldots, \pi_{n-1})$ be a vector of real numbers in $[0,1]$. The random replacement scheme $\mathcal{R}(\pi)$ operates by selecting the first sample element I_1 from the elements of Ω with equal probabilities. I_1 is then removed from the population with probability $(1-\pi_1)$ and replaced with probability π_1. Next I_2 is chosen with equal probabilities from the elements remaining in the population and removed with probability $(1-\pi_2)$ and replaced with probability π_2. This procedure is continued, drawing I_k with equal probabilities from the elements remaining after step $(k-1)$ and replacing it with probability π_k, until a sample $I = (I_1, \ldots, I_n)$ is obtained. Of course $\mathcal{R}(1,\ldots,1)$ and $\mathcal{R}(0,\ldots,0)$ denote sampling with and without replacement respectively. For any scheme $\mathcal{R}(\pi)$, expected values under $\mathcal{R}(\pi)$ will be denoted by E_π.

For any set Λ, let $\mathcal{C}_{\Lambda^n}$ be the class of real-valued functions ϕ on Λ^n that satisfy

(1.1) $\phi(y_1, y_2, \ldots, y_n)$ is a symmetric function of $y_1, y_2, \ldots, y_n \in \Lambda$;

(1.2) $2\phi(u,v,y_3,\ldots,y_n) \le \phi(u,u,y_3,\ldots,y_n) + \phi(v,v,y_3,\ldots,y_n)$ for all $u,v,y_3,\ldots,y_n \in \Lambda$.

The following conjecture was discussed in Karlin (1974).

KARLIN's CONJECTURE.

If $\phi \in C_{\Omega^n}$ *and* $\pi = (\pi_1,\ldots,\pi_{n-1})$ *and* $\pi' = (\pi'_1,\ldots,\pi'_{n-1})$ *are such that* $0 \le \pi_k \le \pi'_k \le 1$ *for* $k = 1,\ldots,n-1$, *then*

$$(1.3)\qquad E_\pi \phi(I) \le E_{\pi'} \phi(I) .$$

A seemingly more general but equivalent formulation of the conjecture is obtained by introducing arbitrary real variate values into the set-up. Suppose that the elements $1,2,\ldots,N$ in the population carry - not necessarily distinct - variate values $a_1,a_2,\ldots,a_N$ and define $X = (X_1,X_2,\ldots,X_n) = (a_{I_1},a_{I_2},\ldots,a_{I_n})$. Then X is a sample from the more general population $\tilde{\Omega} = (a_1,\ldots,a_N)$ generated by the same sampling scheme $R(\pi)$ that produces I from Ω . For any function f on $\tilde{\Omega}^n$ we may define a function ϕ on Ω^n by taking $\phi(i_1,i_2,\ldots,i_n) = f(a_{i_1},a_{i_2},\ldots,a_{i_n})$ and then $\phi(I) = f(X)$. Moreover, if $f \in C_{\tilde{\Omega}^n}$ then clearly $\phi \in C_{\Omega^n}$. Conversely, for every $\phi \in C_{\Omega^n}$ the relation $f(a_{i_1},\ldots,a_{i_n}) = \phi(i_1,\ldots,i_n)$ defines a function $f \in C_{\tilde{\Omega}^n}$, provided the variate values $a_1,\ldots,a_N$ are distinct. The following formulation is therefore equivalent to Karlin's conjecture.

EQUIVALENT CONJECTURE.

If $f \in C_{\tilde{\Omega}^n}$ *and* $\pi = (\pi_1,\ldots,\pi_{n-1})$ *and* $\pi' = (\pi'_1,\ldots,\pi'_{n-1})$ *are such that* $0 \le \pi_k \le \pi'_k \le 1$ *for* $k = 1,\ldots,n-1$, *then*

$$(1.4)\qquad E_\pi f(X) \le E_{\pi'} f(X) .$$

In fact Karlin discussed this second form of the conjecture. We prefer to consider the first formulation because it does not require the additional notation needed to distinguish between different population elements carrying the same variate value.

The conjecture was proved in Karlin (1974) for the following special cases.

(a) $\pi = (0,0,\ldots,0)$, i.e. $R(\pi)$ is sampling without replacement;

(b) $\pi' = (1,1,\ldots,1)$, i.e. $R(\pi')$ is sampling with replacement and either $n \le 12$ or $[N/(N-1)]^{n-1} \le n/(n-3)$;

(c) $\pi' = (1,1,\ldots,1)$ and either

(i) $f(x_1,\ldots,x_n) = (x_1+\ldots+x_n)^2$ or

(ii) $f(x_1,\ldots,x_n) = g(x_1+\ldots+x_n)$, where g is convex and the variate values $a_1,\ldots,a_N$ have only two distinct values.

These results generalize earlier ones of Hoeffding (1963) and Rosén (1967). For a review of the area see Marshall and Olkin (1979). The relation to the theory of comparison of experiments is discussed in Torgersen (1981).

Since the case $\pi = (0,\ldots,0)$ is settled and no results appear to be known for general $\pi \le \pi'$, it seems prudent to focus on the case $\pi' = (1,\ldots,1)$, where at least the partial results (b) and (c) are available to sustain one's optimism. However, even for this case it is easy to agree with Karlin that the matter "appears quite delicate". Rather than to attempt to prove or disprove the conjecture, it appears more feasible to try to indicate a reasonably large class of functions ϕ on Ω^n for which inequality (1.3) holds for $\pi' = (1,\ldots,1)$. The purpose of this paper is to provide such a class.

Let $\mathcal{D}_{\Omega^n}$ be the class of real-valued functions ϕ on Ω^n that satisfy

(1.5) $\phi(i_1,i_2,\ldots,i_n)$ is a symmetric function of $i_1,\ldots,i_n \in \Omega$;

(1.6) $$\frac{2}{N}\sum_{j=1}^{N}\phi(i,j,i_3,\ldots,i_n) \le \phi(i,i,i_3,\ldots,i_n) + \frac{1}{N^2}\sum_{j=1}^{N}\sum_{j'=1}^{N}\phi(j,j',i_3,\ldots,i_n)$$

for all $i,i_3,\ldots,i_n \in \Omega$.

THEOREM.

If $\phi \in \mathcal{D}_{\Omega^n}$, *then* (1.3) *holds for every* $\pi = (\pi_1,\ldots,\pi_{n-1})$ *and* $\pi' = (1,1,\ldots,1)$.

To see the connection between the classes $\mathcal{C}_\Omega n$ and $\mathcal{D}_\Omega n$ consider the inequality

$$(1.7)\qquad \sum_{\nu=1}^{N}\sum_{\nu'=1}^{N} \phi(\nu,\nu',i_3,\dots,i_n)\, c_\nu c_{\nu'} \geq 0$$

for real numbers $c_1,\dots,c_N$ with $\sum_{\nu=1}^{N} c_\nu = 0$. The class $\mathcal{C}_\Omega n$ consists of all symmetric functions ϕ satisfying (1.7) whenever, for some i and j , $c_i = -c_j = 1$ and $c_\nu = 0$ for $\nu \neq i,j$. Since (1.6) may be written in the form

$$(1.8)\qquad \phi(i,i,i_3,\dots,i_n) - \frac{2}{N-1}\sum_{j\neq i}\phi(i,j,i_3,\dots,i_n) + \\ + \frac{1}{(N-1)^2}\sum_{j\neq i}\sum_{j'\neq i}\phi(j,j',i_3,\dots,i_n) \geq 0 ,$$

the class $\mathcal{D}_\Omega n$ consists of all symmetric functions ϕ satisfying (1.7) whenever, for some i , $c_i = 1$ and $c_\nu = -1/(N-1)$ for $\nu \neq i$. Both classes contain the set of symmetric functions ϕ satisfying (1.7) whenever $\sum c_\nu = 0$. Special cases of this set were studied in Bickel and Van Zwet (1980).

2. PROOF OF THE THEOREM

By a simple induction argument (cf. Karlin (1974), lemma 3.1) it suffices to prove the theorem for the case where $\pi = (0,1,1,\dots,1)$ and $\pi' = (1,1,\dots,1)$. Hence the only difference between π and π' is that, when using the scheme $\mathcal{R}(\pi)$, the first element sampled is not replaced.

For $j = 1,\dots,N$, let R_j denote the number of times that element j occurs in the sample and let $R = (R_1,\dots,R_N)$. Obviously $\sum_{j=1}^{N} R_j = n$, the sample size. Let $\delta_{j,j} = 1$ and $\delta_{i,j} = 0$ if $i \neq j$, so that $e_j = (\delta_{1,j},\delta_{2,j},\dots,\delta_{N,j})$ is the j-th unit vector in $\mathbb{R}^N$. Choose any function $\phi \in \mathcal{D}_\Omega n$. Because of (1.5), $\phi(I)$ is a function of R only, say $\phi(I) = \psi(R)$. In view of (1.6), or equivalently (1.8), ψ satisfies

$$(2.1)\qquad \psi(s+2e_i) - \frac{2}{N-1}\sum_{j\neq i}\psi(s+e_i+e_j) + \frac{1}{(N-1)^2}\sum_{j\neq i}\sum_{j'\neq i}\psi(s+e_j+e_{j'}) \geq 0$$

for every vector $s = (s_1,\ldots,s_N)$ having integer co-ordinates $s_j \geq 0$ with $\sum s_j = n-2$ and for $i = 1,\ldots,N$. Because every random replacement scheme is invariant under permutation of the population elements, it is no loss of generality to assume that ψ is also invariant, i.e. that $\psi(r_1,\ldots,r_N)$ is a symmetric function of its N arguments. Hence it suffices to show that

$$(2.2)\qquad E_\pi\psi(R) \leq E_{\pi'}\psi(R)$$

for any symmetric function ψ satisfying (2.1) and for $\pi = (0,1,1,\ldots,1)$ and $\pi' = (1,1,\ldots,1)$.

Define

$$(2.3)\qquad \chi(r) = E_{\pi'}(\psi(R)|R_1=r) .$$

Since $\mathcal{R}(\pi')$ is sampling with replacement,

$$(2.4)\qquad \begin{aligned}&\chi(r+2) - 2\chi(r+1) + \chi(r) = E\{\psi((r+2)e_1+S) +\\ &- \frac{2}{N-1}\sum_{j=2}^{N}\psi((r+1)e_1+e_j+S) + \frac{1}{(N-1)^2}\sum_{j=2}^{N}\sum_{j'=2}^{N}\psi(re_1+e_j+e_{j'}+S)\}\end{aligned}$$

where $S = (0,S_2,\ldots,S_N)$ and $(S_2,\ldots,S_N)$ has a multinomial distribution with cell-probabilities $(N-1)^{-1},\ldots,(N-1)^{-1}$ and sample size $(n-r-2)$. It follows from (2.1) that the right-hand side in (2.4) is nonnegative and hence that χ is a convex function defined at the points $0,1,\ldots,n$. If we extend the definition of χ to the interval $[0,n]$ by linear interpolation, i.e. $\chi(r+\alpha) = (1-\alpha)\chi(r) + \alpha\chi(r+1)$ for $\alpha \in (0,1)$, then χ is convex on $[0,n]$.

Because of the symmetry of ψ

$$E_\pi\psi(R) = E_\pi(\psi(R)|I_1=1)$$

and because the conditional distribution of R given $I_1 = 1$ under $\mathcal{R}(\pi)$ is the same as the conditional distribution of R given $R_1 = 1$ under $\mathcal{R}(\pi')$, we find

$$(2.5) \qquad E_\pi \psi(R) = E_{\pi'}(\psi(R) | R_1 = 1) = \chi(1) .$$

Obviously,

$$E_{\pi'}\psi(R) = E_{\pi'}\chi(R_1)$$

where R_1 has a binomial distribution with expected value n/N under $\mathcal{R}(\pi')$. Since χ is convex on $[0,n]$, Jensen's inequality yields

$$(2.6) \qquad E_{\pi'}\psi(R) \geq \chi(\frac{n}{N}) .$$

On the other hand, since ψ is symmetric

$$E_{\pi'}\psi(R) = E_{\pi'}(\psi(R) | I_1 = 1) = E\chi(1+B)$$

where B has a binomial distribution with expected value $(n-1)/N$. Another application of Jensen's inequality gives

$$(2.7) \qquad E_{\pi'}\psi(R) \geq \chi(1 + \frac{n-1}{N}) .$$

As $n/N \leq 1 \leq 1 + (n-1)/N$, we conclude from (2.6), (2.7) and the convexity of χ that

$$(2.8) \qquad E_{\pi'}\psi(R) \geq \chi(1) .$$

In view of (2.5) this proves (2.2) and the theorem.

3. COMMENTS

Two final remarks should be made. The first is that the proof of the theorem can also be based on a modified form of lemma 1.1 in Bickel and Van Zwet (1980). The proof given here needs less notation and is perhaps a bit more elegant.

The second remark is that it is easy to construct examples of functions belonging to one of the two classes $\mathcal{C}_\Omega n$ and $\mathcal{D}_\Omega n$ but not to both. Take $N = 3$, so $\Omega = \{1,2,3\}$ and $n = 2$ and consider the functions ϕ_1 and ϕ_2 on Ω^2 defined by

$$\phi_1(i,j) = \begin{cases} 4 & \text{if } i = j = 1 \text{ or } i = j = 2 \\ -4 & \text{if } i = j = 3 \\ -5 & \text{if } i = 1,\ j = 2 \text{ or } i = 2,\ j = 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$\phi_2(i,j) = \begin{cases} 1 & \text{if } i = 1,\ j = 2 \text{ or } i = 2,\ j = 1 \\ 4 & \text{if } i = j = 3 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that $\phi_1 \in \mathcal{C}_\Omega 2$ but $\phi_1 \notin \mathcal{D}_\Omega 2$ and that $\phi_2 \in \mathcal{D}_\Omega 2$ but $\phi_2 \notin \mathcal{C}_\Omega 2$.

The functions

$$\phi_3(i_1,\ldots,i_n) = \left(\sum_{j=1}^{n} a_{i_j}\right)^2$$

and

$$\phi_4(i_1,\ldots,i_n) = g\left(\sum_{j=1}^{n} a_{i_j}\right)$$

for convex g and only two distinct values among $a_1,\ldots,a_N$, that were discussed by Karlin in special cases (c i) and (c ii) , belong to both $\mathcal{C}_\Omega n$ and $\mathcal{D}_\Omega n$.

REFERENCES

[1] Bickel, P.J. and Van Zwet, W.R. (1980). On a theorem of Hoeffding. *Asymptotic Theory of Statistical Tests and Estimation*, I.M. Chakravarti, editor, Academic Press, New York, 307-324.

[2] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* 58, 13-30.

[3] Karlin, S. (1974). Inequalities for symmetric sampling plans I. *Ann. Statist.* 2, 1065-1094.

[4] Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

[5] Rosén, B. (1967). On an inequality of Hoeffding. *Ann. Math. Statist.* 38, 382-392.

[6] Torgersen, E. (1981). Comparison of some statistical experiments associated with sampling plans. Technical Report, Department of Mathematics, University of Oslo.

36273