TR/09/90 September 1990

METRIC STATISTICS ON THE SIMPLEX

by

W. H. FOSTER

1. Introuduction

Given a finite set of points S = $\{x_1,....,x_k\} \subseteq \Delta^{n-1}$, the standard n-1 simplex in $\mathbb{R}^n$, we define :

1. $d_i(S)=\max\left\{\left|x_{ji}-x_{mi}\right|: \forall\ j,m \in k\right\}, i \in n$

Where $x_j = \{x_{ji},....,x_{jn}\}, j \in k$.

Note that the distributions for d_i (S)are the same $\forall\ i \in n$.

2. d(S) = max{d_i (S): i∈n}

It is an easy exercise in Crofton's Theorem (2.2 of [2] or Chapter 5 of [3]) to obtain expressions for the probability distribution of d_i(S).
We present these in section 2 below.

The main purpose of this paper is to obtain an expression for the cumulative probability distribution of d(S) for a given n ≥ 2 and k≥2 (see 3.4).
This is achieved in section 3 by a crucial use of the agreement or clustering measure c(S) , see [1] , for a finite set S $\subseteq \Delta^{n-1}$.
In 3.5 we provide calculations of the above probabilities for some values of n for k = 2.
The rather messy expressions for the probabilities contrasts with the same measure d(S) for S = k points in $[0,1]^n \subseteq \mathbb{R}^n$.
It is easy to show, using the independence of the coordinates, probability d(S)≤λ = $\lambda^{n(k-1)}(k-(k-1)\lambda)^n$

<u>Notation</u>

The only non-standard notation we use is to denote for n $\in \mathbb{N}$,
$n = \{x \in \mathbb{N}: 1 \le x \le n\} = \{1,2,..,n\}$

2.<u>Distribution of d_i (S)</u>

Suppose that we have k points S = $\{x_i,....,x_k\} \subseteq \Delta^{n-1}$. Choose a coordinate, say the ith.
Let $d_i(S)$ = max $\left\{\left|x_{ji}-x_{mi}\right|: \forall\ j,m \in k\right\}$, i ∈ n
where $x_j = (x_{ji},....,x_{jn}), j \in k$

We can apply Crofton's Theorem to obtain the following expression for the cumulative probability $p^{\lambda}_{n.k}$ of $d_i(S)$.

$$P^{\lambda}_{n,k} = k(n-1)\sum_{r=o}^{k-1} \left\{\int_o^1 (-1)^{nr} \begin{bmatrix} k-1 \\ r \end{bmatrix} (\lambda - x)^{(n-1)r} x^{nk-nr-k-1+r} dx \right\} + \alpha\lambda^{k(n-1)}$$

Where $\alpha = \sum_{r=1}^{k-1} \frac{(-1)^{nr+1}\begin{bmatrix} k-1 \\ r \end{bmatrix}}{\begin{bmatrix} nk-k-1 \\ nr-r \end{bmatrix}}$

If both n and k are odd then $\alpha = 0$. This is left as a pleasant Exercise for the reader.

<u>Examples</u>

1. $P^{\lambda}_{3,3} = 6\lambda^2 - 8\lambda^3 + 3\lambda^4$

2. $P^{\lambda}_{n,2} = 1 - 2(1-\lambda)^n \left\{ \sum_{r=1}^{n-1} \frac{\begin{bmatrix} 2n-2-r \\ n-1-r \end{bmatrix}}{\begin{bmatrix} 2n-2 \\ n-1 \end{bmatrix}} \lambda^{r-1} \right\}$

3. The p.d.f. for $d_i(S)$ in the case k = 2 can be calculated Directly and we obtain:

$$2(n-1)^2 x^{2n-9} \int_x^1 \frac{(1-z)^{n-2}}{z^{2n-2}} dz$$

or by differentiating 2. we obtain the equivalent expression for the p.d.f.

$$\frac{2(n-1)!(1-x)^{n-1}}{(2n-3)!} \left\{ \sum_{r=o}^{n-2} \begin{bmatrix} 2n-r-4 \\ n-2 \end{bmatrix} x^r \right\}$$

4. The mth moment about zero for $d_i(S)$ in the case k=2 is

$$\mu_m = \frac{2(n-1)^2 m!(n-2)!}{(m+2n-2)!(m+n-1)!}$$

3.Distribution of d(S)

First we outline some basic results needed in this section.

3.1 Agreement Measure Background

Let $S = \{x_1,...,x_k\} \subseteq \Delta^{n-1}$ where $x_j = (x_{j1},...,x_{jn})$, $j \in k$ be a random set of k not necessarily distinct points.

Let $a_i = \min \{ x_{ji} : \forall\ j \in k \}$.

In [1] we defined $\subset(S) = 1 - \sum_{i=1}^{n} a_i$ and it was shown there that $\subset S)$ was distributed as a beta distribution with parameters (n-1)(k-1), n.

Furthermore, given S we obtain an ordered k-partition of n, $A_i,...,A_k$, where $A_j = \{i \in n : a_i = x_{ji}\}$.(we have to assume that S is in general position-see 2.5 in [1] - also at least two of the sets $A_1,...A_k$ are non empty.)

If $r_j = |A_j|$, $j = 1,...,k$, then the probability that S has the ordered k-partition $\Pi = \{A_1,...,A_k\}$ associated to it is

$$P(\Pi) = \frac{(nk-n-k)\ !\ ((n-1)!)^{k}}{(n-r_1-1)\ !\ (n-r_2-1)\ !....(n-r_k-1)\ !\ (nk-k)!}$$

(see theorem 20 in [1])

3.2 Subsets of Δ^{n-1} and their measures.

Our approach in this section is to find probabilities by considering the measures of subjects of Δ^{n-1}.

Let $B^n = \left\{(x_1,...,x_n) : \sum_{i=1}^{n} x_i \leq 1,\ 0 \leq x_i \leq 1, \forall\ i \in n\right\}$,

then given $x \in B^n, x = (x_1,...,x_n)$ we define the unique regular subsimplex $\Delta(x)$ of Δ^{n-1} spanned by $x(i) = (y_1,...,y_n), i \in n$ where

$y_j = x_j$, $j \neq i$,

$y_i = 1 - \sum_{j=1}^{n} x_j + x_i$

see 2.4 in [1]).

or given, $\in n$, $a \in \mathbb{R}$,

Let $H(1,R = (y \in \mathbf{R}^n \; y=(y_1,...,y_n)$, $y_L \geq a\,\lambda$

We then have for $x=(x_1,...,x_n) \in B^n$

$\Delta(x) = \Delta^{n-1} \cap H(1,x_1) \cap ... \cap H(n_1 x_n)$ or

$\Delta(x) = \{y \in \Delta^{n-1} : y=(y_1,...,y_n)$, $y_i \geq x_i\}$...........(1).

Also we have that the measure or volume of $\Delta(X)$ is $V(\Delta(X)) = \frac{y_n}{(n-1)!}(1-\sum_{i=1}^{n} x_i)^{n-1}$, regarded as a subset of $\mathbb{R}^{n-1}$.

<u>Definition</u>

Let a , $b \in \mathbb{R}^n$ where $a=(a_1,...,a_n)$ $b=(b_1,...,b_n)$

We define $a \oplus b = (c_1,...,c_n)$ where $c_i = \max\{a_i , b_i\}$.

Note that it is possible for $a,b \in B^n$ but $a \oplus b \notin B^n$.

Clearly, if a ,b ,...,$w \in \mathbb{R}^n$ then $a \oplus b \oplus .. \oplus w$ is well defined.

<u>Lemma 1</u>

Let a ,b ,...,$w \in B^n$,

1, $\Delta(a) \cap \Delta(b) \cap .. \cap \Delta(w) = \Delta(a \oplus b \oplus .. \oplus w)$ if $a \oplus b \oplus .. \oplus w \in B^n$

2. $\Delta(a) \cap \Delta(b) \cap .. \cap \Delta(w) = \phi$ if $a \oplus b \oplus .. \oplus w \notin B^n$

<u>Proof</u>

Follows directly from (1) above.

E.O.P.

As an application of lemma 1 we consider the problem of finding the volume of the union of s subsets $S_i \subseteq A$, $i \in s$, in the special case where each of the S_i is a particular regular subsimplex of $A = \Delta^{n-1}$,

If $T \subseteq s$ we let V_T be the volume of $\bigcap_{i \in T} S_i$, where $V_\phi = o$, and $o(T)$ be the number of elements or order of T. Let V be the volume of $S_1 \cup ... \cup S_s$.

By the combinatorial principle of inclusion-exclusion, Chapter 5 Of [4], we have

$$V = \sum_{T \subseteq s} (-1)^{o(T)+1} V_T .$$

Now let $\in_i = (0,0,..,1,0,.,0)$,$\forall$ $i \in n$, be the standard unit vector with ith coordinate =1 and the rest 0.

Fix λ , $0 \leq \lambda \leq 1$, and let $S_i = \Delta(\lambda e_i)$, $i = 1,..,n$.

Not that the volume of $\Delta(\lambda e_i) = \frac{\sqrt{n}}{(n-1)!}(1-\lambda)^{n-1}$

It is easy to see that

$S_i = \{(x_1, ..., x_n) \in \Delta^{n-1} : x_i \geq \lambda\}$

Then, in this case, we have

$$\bigcup_{i=1}^{n} S_i = \{(x_1,...,x_n) \in \Delta^{n-1} : \exists \text{ j such that } x_j \geq \lambda\}$$

$$= \{(x_1, ..., x_n) \in \Delta^{n-1} : \max_{j\in n} \{x_j\} \geq \lambda\} \text{ .... } (2)$$

Now V_T = volume of $\bigcap_{i\in T} S_i$ = volume of $\bigcap_{i\in T} \Delta(\lambda e_i)$

Note that $\bigoplus_{i\in T} \lambda e_i = \lambda \sum_{i\in T} e_i \in B^n$ iff $\lambda o(T) \leq 1$ i.e $o(T) \leq \frac{1}{\lambda}$

Hence V_T = volume of $\bigcap_{i\in T} \Delta(\lambda e_i)$ = volume of $\Delta(\lambda \sum_{i\in T} e_i)$

$$= \frac{\sqrt{n}}{(n-1)} (1-o(T)^{n-1} \text{ if } 1-o(T)\lambda \geq o$$

$$= 0 \qquad\qquad i+1-o(T)\lambda \leq o$$

Thus $V = \sum_{T\subseteq n} (-1)^{o(T)+1} V_T = \frac{n}{(n-1)!} \sum_{r=1}^{m} (-1)^{r+1} \left[{n \atop r}\right] (1-r\lambda\lambda^{n-1}$

Where:

$m \in \mathbb{N}$ is such that $m = \min\{n, b\}$

and $b \in \mathbb{N}$ is such that $b \leq \frac{1}{\lambda} < b+1$.

We write $b = \left[\frac{1}{\lambda}\right]$

The following Lemma is well known, (see 2.14 of[2]) however the proof we present here is, we believe, new and is elementary depending only upon Lemma 1 and the inclusion-exclusion principle of combinatorics

Lemma 2

Let $x=(x_1,\dots,x_n)$ be a random point in Δ^{n-1}.

Then the probability that max $\{x_1,\dots,x_n\} \le \lambda$ is

$$\sum_{r=o}^{m}(-1)^r\left[{n \atop r}\right](1-r\lambda\lambda^{n-1},\ m=\min\{n,\left[\frac{1}{\lambda}\right]\}.$$

Proof

Now we have from (2) that given a random point $(x_1,\dots,x_n)\in\Delta^{n-1}$, the probability that at least one of it coordinates is $\ge \lambda$ is

$$q(\lambda(= \sum_{r=1}^{m}(-1)^{r+1}\left[{n \atop r}\right](1-r\lambda\lambda^{n-1},\ m=\min\{n,\left\lfloor\frac{1}{\lambda}\right\rfloor\}.$$

Hence the probability that none of its coordinates is $\ge\lambda$ is

$$p(\lambda(=1-q(\lambda(=\sum_{r=o}^{m}(-1)^r\left[{n \atop r}\right](1-r\lambda\lambda^{n-1},\ m=\min\{n,\left[\frac{1}{\lambda}\right]\}.$$

E.O.P

3.3 Simplexes and Barycentric Coordinates

Let $A=\{V_1,\dots,V_q\}$ be an affine independent set of vectors in $\mathbb{R}^p$.

Then they span a simplex $\Delta(A)$ of dimension q-1 given by

$$\Delta(a) = \{\lambda_1 v_1+\dots+\lambda_q v_q : \sum_{i=1}^{q}\lambda_i=1,\ 0\le\lambda_i\le 1,\ \forall\ i\in q\}$$

If $x\in\Delta(A)$ and $x=\lambda_1 v_1+\dots+\lambda_q v_q$ then $(\lambda_1,\dots,\lambda_q)$ are called the barycentric coordinates of x.

Using the barycentric coordinates gives an affine linear isomorphism from $\Delta(A)$ to the standard simplex Δ^{q-1}.

We then have

Lemma 3 (Notation as above)

Let x be a random point in $\Delta(A)$ with barycentric coordinates $(\lambda_1,\dots,\lambda_q)$. Then the probability that max $\{\lambda_1,\dots,\lambda_q\}\le\lambda$ is

$$\sum_{r=0}^{m}(-1)^r\left[{q \atop r}\right](1-r\lambda\lambda^{q-1},\ m=\min\{q,\left[\frac{1}{\lambda}\right]\}.$$

3.4 **Calculations for D(S)**

Let S = $\{x_1,\ldots,x_k\} \subseteq \Delta^{n-1}$ where $x_j = (x_{j1},\ldots,x_{jn})$, $j \in k$

and furthermore we assume that c(S) $c = 1 - \sum a_i$ where

$a_i = \min\{x_{ji} : \forall j \in k\}$.

Also we assume that the associated k-partition $\Pi = \{A_1, \ldots, A_k\}$ where $r_j = |A_j|$.

Note that at least two of the sets $\{A_1, \ldots, A_k\}$ are non empty.

If we let a(p)= $(a_1, \ldots, a_{p-1}, 1 - \sum a_i + a_p, \ldots, a_n)$, $P = 1,\ldots,n$

Then the set A = {a(1),a(2),..,a(n)} is an affine independent set spanning the simplex $\Delta(A)$ which is the smallest regular subsimplex of Δ^{n-1} containing S.

Let $\sigma_j = \langle \{a(i) : i \in A_j^c\} \rangle \ \forall j \in k$, and we then have:

1. $x_j \in \mathrm{Int}(\sigma_j)$
2. Since $|A_j^c| \geq 2$, σ_j is a regular simplex of edge length $\sqrt{2}c$, (see 2.4 of [1])
3. $y \in \sigma_j$, $y = (y_1,\ldots,y_n) \Rightarrow y_i = a_i$, $\forall i \in A_j$.

Fix $j \in k$.

Let $d_i = \max\{|x_{ji} - x_{si}| : 1 \leq i \leq n, 1 \leq S \leq k\}$

Also, since $x_j \in \sigma_j = \Delta(A_j^c)$ we have:

$x_j = \sum_{i \in A_j^c} \lambda_i a(i)$ where the λ_i are the barycentric coordinates of

x_j in $\Delta(A_j^c)$, i.e. $\sum_{i \in A_j^c} \lambda_j = 1$ and $0 \leq \lambda_i \leq 1$, $\forall i \in A_j^c$

Lemma 4 (Notation as above)

$d_j = c \max \{ \lambda_i : i \in A_j^{\varepsilon} \}$

Proof

We have $x_j = \sum_{i \in A_j^c} \lambda_i a(i)$.

Now we have $x_{j8} = a_8 , \forall S \in A_j$ ………………(3)

$$x_{j8} = \left[\sum_{i \in A_j^c} \lambda_i a_8 \right] + \lambda_8 (1 - \sum_{t=1}^{n} a_t)$$

$$= a_8 + \lambda_8 c , \forall s \in A_j^c \quad \text{………………………(4)}$$

Since $a_8 = \min \{x_{t8} : 1 \le t \le k\}, S \in n$

We clearly have

$d_j = \max \{ |x_{ji} - x_{8i}| : 1 \le i \le n , 1 \le S \le k\}$

$= \max \{x_{ji} - a_i : 1 \le i \le n\}$

$= \max \{\lambda_8 c : s \in A_j^c\}$ (from(3),(4)above)

$= c \max \{\lambda_S : s \in A_j^c\}$

Hence result.
E.O.P

Lemma 5

Let $P_j^{\lambda,c}$ be the probability that $d_j \le \lambda$. Then we have:

$P_j^{\lambda,c} = \sum_{r=o}^{m} (-1)^r \left[{q \atop r} \right] (1 - r \underline{\lambda})^{q-1}$, $m = \min \{q , \left\lfloor \frac{c}{\lambda} \right\rfloor \}$.
$\qquad\qquad\qquad\qquad\quad C$

Where $q = n - r_j$.

Proof

We have $|A_j^c| = n - r_j$, hence σ_j is an $n - r_j - 1$ simplex.

The result then follows from Lemmas 3,4.
E.O.P

N.B. All of the above probabilities are conditional in that we assume fixed c and partition Π.

We let $p^{\Pi,c}(\lambda)$ be the probability that $d(S) \leq \lambda$ given $c = c(S)$ and the partition Π.

<u>Lemma 6</u>

$$p^{\Pi,c}(\lambda) = \prod_{j=1}^{k} p_j^{\lambda,c}$$

<u>Proof</u>

The only condition we need to check is independence, as it is clear that $d(S) \leq \lambda \Leftrightarrow d_j \leq \lambda, \forall j \in k$.

But each x_j is random point in σ_j and so can take any barycentric coordinates independently of the other points.

E.O.P.

The p.d.f. for c is $\frac{(nk-k)!c^{nk-n-k}(1-c)^{n-1}}{(nk-n-k)!(n-1)!}$ as c is distributed as a beta distribution with parameters (n-1)(k-1), n .

Hence ,given the partition Π, let $p^{\Pi}(\lambda)$ be the probability that $d(s) \leq \lambda$. We then have

$$P^{\Pi}(\lambda) = \int_o^1 p^{\Pi,c}(\lambda)\frac{(nk-k)!c^{nk-n-k}(1-c)^{n-1}}{(nk-n-k)!(n-1)!} \; dc$$

$$= \frac{(nk-k)!}{(nk-n-k)!(n-1)!}\int_o^1 p^{\Pi,c}(\lambda\lambda)\;(-c)^{n-1}\; dc$$

Where $p^{\Pi,c}(\lambda) = \prod_{j=1}^{k} p_j^{\lambda,c}$ and

$$p_j^{\lambda,c} = \sum_{r=0}^{m} (-1)^r \left[\begin{smallmatrix} q \\ r \end{smallmatrix}\right](c - r\lambda)^{q-1} \;,\; m = \min\{q, \left\lfloor \frac{c}{\lambda} \right\rfloor\}, q = n - r_j$$

The probability that we have the partition Π is

$$p(\Pi) = \frac{(nk-n-k)!((n-1)!)^{k}}{(n-r_1-1)!(n-r_2-1)!\ldots(n-r_k-1)!(nk-k)!}$$

The number of such partions is $\frac{n!}{r_1!r_2!\ldots r_k!}$.

Let $p(\lambda)$ be the probability that $d(s) \le \lambda$

<u>THEOREM</u> (Notation as above)

$$p(\lambda) = \sum_{n}\left(P(\Pi)\frac{n!}{r_1!r_2!..r_k!}P^{\Pi}(\lambda)\right)$$

$$= n\sum_{n}\left(\binom{n-1}{r_1}\cdots\binom{n-1}{r_k}\int_0^1 P^{\Pi,c}(\lambda)(1-c)^{n-1}\,dc\right)$$

Sum over all ordered partitions $\Pi = \{r_1, .., r_k\}$ of n such that

1. $\sum_{l=1}^{k} r_l = n$
2. $r_l \le n-1, \forall\, l \in k$.

<u>Proof</u>

The set of all such partitions are exclusive and exhaustive. The Results follows afrer a small amount of algebraic manipulation.
E.O.P

3.5 Examples

We consider the case k=2 and compute $p(\lambda)$for some valurs of n.

Let e = $\left\lfloor \frac{n}{2} \right\rfloor$

We have for $\frac{1}{m+1} \le \lambda \le \frac{1}{m}$, $1 \le m < e$.

$$p(\lambda) = n\sum_{r=1}^{n-1}\left(\binom{n-1}{r}\binom{n-1}{n-r}\int_0^1 P^{\Pi.c}(\lambda)(1-c)^{n-1}\,dc\right)$$

Where

$$\int_0^1 p^{\Pi.c}(\lambda)(1-c)^{n-1}dc = \sum_{t=0}^{U}\int_{t\lambda}^{(t+1)\lambda} p^c_{r,t}(\lambda)\, p^c_{n-r}(\lambda)(1-c)^{n-1}dc$$

and

$$p^c_{r,t}(\lambda) = \sum_{a=0}^{t}(-1)^s \begin{bmatrix} r \\ s \end{bmatrix}(c-s\lambda)^{r-1}\ ,\ w = \min\{n-r-1, r-1, m\}.$$

For $o \le \lambda \le \frac{1}{e}$ we use the above with $m = e$

We used REDUCE running on an IBM PS2 to obtain the cumulative probability function $p(\lambda)$ for n=3 to n = 1O (all for k = 2 i.e. 2 points in Δ^{n-1}).

Some sample results are shown below.

n = 3

[0 , 1]	$x^2(3x^2-8x+6)$

n = 4

[O, 1/2]	$-\frac{2}{5}x^3(88x^3-29x^2+225X-80)$
[1/2, 1]	$\frac{1}{5}(16X^6-1O8X^5+27OX^4-92OX^3+18OX^2-36x+3)$

4 Application

One application of d(S),where k = 2 ,is given by decision making techniques using pairwise comparisons between n options. The resulting decision is then usually expressed in normalized form as scores $(x_1,..,x_n)$ where $\sum_{i=1}^{n} x_i = 1$, $x_i \ge 0$, $\forall i \in n$ i.e. as a point in Δ^{n-1} .

If the number of options is large then samples of the pairwise comparisons are usually taken. The resulting normalised scores from the sample can be considered as approximations to the score ideally obtained from using all the comparisons.

One method by which the sample scores can be assessed is to measure the consistency of the comparisons using weighted transitivity. If the consistency is good then it is reasonable to infer that all the comparisons follow from the sample and that

the normalised scores from the sample are reliable.
However, it is clear that a notion of distance is needed which allows statements to be made concerning the link between consistency of the sample and the likely consequent error. Also different sampling schemes need to be assessed.
d(S) gives a clear idea of the error in that if
S = {x,y} , then d(S)≤.05 means that no two components differ more than .05 or 5%.
See[5] for more details..

REFERENCES

[1] Foster , W.H. and C,E. Tripp (1990) A Clustering Statistic For Sets of Points on the Simplex. *Brunel Technical Report*, *Mathematics Dept.*, *Brunel University*.

[2] Kendall , M.G. and Moran , P.A.P. (1963) *Geometrical Probability* Griffin ,London.

[3] Solomon, H.(1978) *Geometric Probability*. Society for Industrial and Applied Mathematics , Philadelphia.

[4] Anderson , I. (1989) A *First Course in Combinatorial Mathematics* , Oxford Applied Mathematics and Computing Series , Oxford.

[5] Foster , W.H. (1990) Sampling Pairwise Comparisons . *In preparation* .