Notes on Order Statistics and Barankin's Structures

Shintaro SONO
Associate Professor
Faculty of Economics
Hokkaido University

(These notes are based on Chapters 9 and 10 in Sono [5]. The writer wishes to thank Prof. Y. Suzuki and Prof. E.W. Barankin for their critical suggestions.)

Note A. An Application of Order Statistics to a combinatorial Problem

1. Simple Case

To clarify the idea a simple case is considered. Let X_j, $j=1,\ldots,N$, be independently identically distributed real valued random variables on some fixed probability space, $(\Omega, \mathbb{F}, P)$, satisfying

$$(1.1)\quad X_j \sim \text{uniform distribution on } I := \,]0, 1[\,,\ j=1,\ldots,N.$$

Then the usual order statistics, $0 < X_{(1)} < \ldots < X_{(N)} < 1$, a.s. (P), are defined by X_j, $j=1,\ldots,N$. It is well-known and easily obtained that

$$(1.2)\quad P(X_{(n)} \leqq x) = \frac{N!}{(n-1)!(N-n)!}\int_0^x t^{n-1}(1-t)^{N-n}\,dt,\ x \in I.$$

On the other hand,

$$(1.3)\quad P(X_{(n)} \leqq x) = P(\bigcup_{d \in \mathbb{D}} (\bigcap_{j \in d} \{X_j \leqq x\})),$$

where $\mathbb{D}$ is the set of all $d \subset \{1,\ldots,N\}$ such that $\# d=n$, i.e.,

(1.4) $\mathbb{D} := \{d; d \subset \{1,\ldots,N\} \ \& \# d=n\}$.

The right hand side of (1.3) is equal to

$$(1.5)\ \Sigma_{r=1}^{\infty} (-1)^{r-1} \Sigma_{\{d_1,\ldots,d_r\}\subset \mathbb{D}} P(\cap_{j \in d_1 \cup \ldots \cup d_r} \{X_j \leqq x\})$$

where $\Sigma_{\{d_1,\ldots,d_r\}\subset \mathbb{D}}$ runs over all $\{d_1,\ldots,d_r\} \subset \mathbb{D}$ satisfying $\# \{d_1,\ldots,d_r\}=r$. Therefore, from (1.1) and independence of X's,

(1.6) (1.5)$= Q(x)$, $x \in I$,

where

$$(1.7)\ Q(Z) := \Sigma_{r=1}^{\infty} (-1)^{r-1} \Sigma_{\{d_1,\ldots,d_r\}\subset \mathbb{D}} Z^{\#(d_1 \cup \ldots \cup d_r)} .$$

(From (1.4), $Q(Z)$ is a polynomial of Z.)
From (1.2) and (1.6), using binomial theorem and termwise integration to (1.2),

$$Q(x) = \text{the right hand side of (1.2)}$$

$$= \frac{N!}{(n-1)!} \Sigma_{k=0}^{N-n} \frac{(-1)^k}{k!(N-n-k)!} \cdot \frac{x^{n+k}}{n+k} , \ w \in I .$$

Therefore, as polynomials of Z,

$$(1.8)\ Q(Z) = \frac{N!}{(n-1)!} \Sigma_{k=0}^{N-n} \frac{(-1)^k}{k!(N-n-k)!} \cdot \frac{Z^{n+k}}{n+k} .$$

It should be noted that the similar terms of (1.7) are arranged in (1.8). And this rearrangement from (1.7) to (1.8) solves the following problem:

Problem (1): " For any nonnegative integer, m, let W_m be the set of all $\{d_1,\dots,d_r\}\subset \mathbb{D}$, $r=1,2,3,\dots$, such that $\#(d_1\cup \dots \cup d_r)=m$, and define $S(\{d_1,\dots,d_r\}):=$' If r is odd, then +1, and if r is even, then -1.' Then compute $\Sigma_{w\in W_m}S(w)$."
The answer is $\Sigma_{w\in W_m}S(w)=(-1)^k N!/((n-1)!k!(N-n-k)!(n+k))$, if $m=n+k$, $k=0,\dots,N-n$.

The polynomial, (1.7), was used in Sono [4]. In this Note (1.1)-(1.8), Problem(1), and its answer are generalized.

2. Generalization

For given positive integers, M, N, n_i, $i=1,\dots,M$, satisfying

$$\Sigma_{i=1}^{M} n_i \leq N,$$

the polynomial, $Q(Z_1,\dots,Z_M)$, is defined:

(2.1) $Q(Z_1,\dots,Z_M):=$

$$\Sigma_{r=1}^{\infty}(-1)^{r-1}\Sigma^{(\#)}_{\{d^1,\dots,d^r\}\subset \mathbb{D}_M}\Pi_{i=1}^{M} Z_i^{\#(d_i^1\cup\dots\cup d_i^r)},$$

where $\mathbb{D}_M$ is the set of all ordered pairs, $(d_1,\dots,d_M)$, such that $\#d_i=n_i$, $d_i\subset\{1,\dots,N\}$ and $d_i\cap d_j=\emptyset$ $(i\neq j)$, for all $i,j=1,\dots,M$. $\Sigma^{(\#)}$ in (2.1) is the summation of all $\{d^1,\dots,d^r\}\subset\mathbb{D}_M$ satisfying

$$(\#)\ \#\{d^1,\dots,d^r\}=r \text{ and } d_i^{\ell}\cap d_j^{k}=\emptyset\ (i\neq j),\ i,j=1,\dots,M,\ \ell,k=1,\dots,r.$$

(In general, put $d^{\ell}:=(d_1^{\ell},\dots,d_M^{\ell})$, $d^{\ell}\in\mathbb{D}_M$, & $\Sigma_{\emptyset}^{(\#)}:=0$.)

Then consider the following problem:

Problem(M): " For any ordered pair of nonnegative integers, $m=(m_1,\dots,m_M)$, let W_m be the set of all $\{d^1,\dots,d^r\}\subset \mathbb{D}_M$, $r=1,2,3,\dots$, such that

$$\# \cup_{\ell=1}^{r} d_i^{\ell}=m_i,\ i=1,\dots,M, \text{ under the condition, } (\#).$$

Define $S(\{d^1,\dots,d^r\}):=$' If r is odd, then $+1$, and if r is even, -1.' Then compute $\Sigma_{w\in W_m} S(w)$."

It is clear that the collection of similar terms in (2.1) solves Problem(M).

Let X_j , $j=1,\dots,N$, be independently identically distributed random variables on some probability space, $(\Omega, \mathbb{F}, P)$, such that

(2.2) X_j is $\cup_{i=1}^{M} I_i\,(I_i:=\{i\}\times]0,1[)$-valued and

$$P(X_j\in\{i\}\times]0,x])=p_i\cdot x,\ x\in I,\ i=1,\dots,M,$$

where p_i's are fixed positive numbers satisfying $\Sigma_{i=1}^{M} p_i=1$. (Of course, for any such p's, there exist X's and $(\Omega, \mathbb{F}, P)$ satisfying (2.2).)

Then the order statistics, $X_{(i,n)}(\omega)$, $i=1,\dots,M$, $n=1,\dots,N$, $\omega\in\Omega$ are defined:

(2.3) If there exists $\{j_1,\dots,j_{N(i)}\}\subset\{1,\dots,N\}$ such that $N(i)\geqq n$, $\{j_1,\dots,j_{N(i)}\}=\{j\in\{1,\dots,N\};$

$X_j(\omega)\in I_i\}$, and $X_{j_1}(\omega)<\dots<X_{j_{N(i)}}(\omega)$,

then $X_{(i,n)}(\omega):=X_{j_n}(\omega)$, otherwise $:=1$.

Informally $X_{(i,n)}$ is the n-th order statistic in I_i if exists. Like (1.2), for M, N, n_i, $i=1,\dots,M$, in (2.1), it is easily obtained that

(2.4) $P(\cup_{i=1}^{M}\{X_{(i,n_i)} \in \{i\}\times]0,\ x_i]\})=$

$$\frac{N!}{M!(N-\Sigma_{i=1}^{M} n_i)!\ \Pi_{i=1}^{M}\ (N_i-1)!}\ .$$

$$\int_0^{p_1x_1} dt_1 \ldots \int_0^{p_Mx_M} dt_M\ \Pi_{i=1}^{M} t_i^{n_i-1}\ .\ (1-\Sigma_{i=1}^{M}\ t_i)^{N-\Sigma_{i=1}^{M} n_i}\ ,$$

for $x_i \in I_i$, $i=1,\ldots,M$.

On the other hand,

(2.5) the left hand side of (2.4)

$$= P(\bigcup_{(d_1,\ldots,d_M)\in \mathbb{D}_M} (\cap_{i=1}^{M} \cap_{j\in d_i} \{X_j \in \{i\}\times]0,\ x_i]\})$$

$$= \Sigma_{r=1}^{\infty} (-1)^{r-1} \sum^{(\#)}_{\{d^1,\ldots,d^r\}\subset \mathbb{D}_M} P(\cap_{i=1}^{M} \cap_{j\in \cup_{\ell=1}^{r} d_i^{\ell}}$$

$$\{X_j \in \{i\}\times\ ]0,\ x_i]\})$$

$$= Q(p_1x_1,\ldots,p_Mx_M).$$

Therefore, using multinomial theorem and termwise integration, from (2.4) and (2.5),

(2.6) $Q(Z_1,\ldots,Z_M)=$

$$\frac{N!}{M!\ \Pi_{i=1}^{M}(n_i-1)!}\ \Sigma_{k=0}^{N-\Sigma_{i=1}^{M} n_i}\ \frac{(-1)^k}{(N-k-\Sigma_{i=1}^{M} n_i)!}\ .$$

$$\sum_{(\Sigma_{i=1}^{M} k_i=k,\ k_i \geqq 0)}^{M} \frac{\Pi_{i=1}^{M}\ Z_i^{n_i+k_i}}{\Pi_{i=1}^{M}\ k_i!(n_i+k_i)}\ ,$$

as polynomials of Z's. The answer to Problem(M) is given by (2.6): If $m=(n_i+k_i\,;i=1,\ldots,M)$, where $\Sigma_{i=1}^{M} k_i \leqq N-\Sigma_{i=1}^{M} n_i$, $k_i \geqq 0$, $i=1,\ldots,M$, then

$$\Sigma_{w \in W_m} S(w) = \frac{(-1)^{\Sigma_{i=1}^{M} k_i}\; N!}{M!\,(N-\Sigma_{i=1}^{M}k_i-\Sigma_{i=1}^{M}n_i)!\;\Pi_{i=1}^{M}\,(n_i-1)!\,k_i!\,(n_i+k_i)}\;,$$

otherwise=0.

Note B. On Barankin's Formulation for Stochastic Phenomena

1. Notations

It is shown that the Barankin's structures in [1],[2], and [3] are special cases which are mathematically derived by elementary considerations of some topological space. Usual notations in set theory are used: a set, S, is called a topological space if the class of all open sets in S is defined, the power set of S, i.e., the set of all subsets in S, is denoted P(S), i.e., $P(S):=\{X;\ X\subset S\}$, the closure and interior (or open kernel) of a subset, X, in a topological space, S, are written as X^- and X^o, respectively; $(X^-)^o$, $X\cup(Y^-)$, etc. are written as X^{-o}, $X\cup Y^-$, etc..

2. P.o.s. As a Topological Space

A p.o.s. (:=partially ordered structure), $(S,\leqq)$, is considered under some topology. This topological space is used freely in the following sections.

2.1. Definition

The sets of all the upper elements and all the lower elements for any $p\in S$ are denoted U(p) and L(p), respectively, i.e., $U(p):=\{q\in S;\ q\geqq p\}$ and $L(p):=\{q\in S; q\leqq p\}$. The open sets in $(S,\leqq)$ are defined by $U(p)(p\in S)$: " A subset, $X\subset S$, is open if and only if $U(p)\subset X$ for all $p\in X$."

This topology is well-defined because the axioms of open sets are easily established. Therefore the closed sets in $(S, \leqq)$ are also well-defined, i.e., $X \subset S$ is closed if and only if $S-X$ is open. From the definition it is clear that $X^{o} = \{p \in X;\ U(p) \subset X\}$ and $X^{-} = \cup\{L(p);\ p \in X\}$.

2.2. Filter, Strong Filter, and Minimal Elements

The filter on $(S, \leqq)$, which is called p.o.s. - filter to exclude confusion with usual topological filters, is defined: " A subset $X \subset S$ is a p.o.s. - filter if and only if $X \neq \emptyset$, for any $(p,q) \in X \times X$ there exists $r \in S$ such that $r \leqq p$ & $r \leqq q$, and for any $(p,q) \in X \times S$ $p \leqq q \Rightarrow q \in X$, i.e., $U(p) \subset X$ for all $p \in X$." The strong filter on $(S, \leqq)$ is defined: " A subset $X \subset S$ is a strong filter if and only if X is a p.o.s. - filter on $(S, \leqq)$ and for any $(p,q) \in X \times X$ there exists $r \in X$ such that $r \leqq p$ & $r \leqq q$." Clearly any p.o.s. - filter on S is an open set in S and $U(p)$ is a strong filter on S for all $p \in S$. For any subset $X \subset S$ the sets of all the minimal elements and the minimum element in X are denoted $ML(X)$ and $M(X)$, respectively, i.e., $ML(X) := \{p \in X;\ L(p) \cap X = \{p\}\}$ and $M(X) := \{p \in X; q \geqq p \text{ for all } q \in X\}$. Of course $M(X) = \emptyset$ or $\#M(X) = 1$ and $M(X) \subset ML(X)$. The following propositions are obtained:

Proposition 2.1. If $X \subset Y \subset S$ and X is closed, then $ML(X) \subset ML(Y)$. (The proof is clear.)

Proposition 2.2. If $X \subset Y \subset S$, $X \neq \emptyset$, and X is closed and Y is a p.o.s. - filter, then Y is a strong filter and $M(X) = ML(X) = ML(Y) = M(Y)$.

Proof: Take any $(p,q) \in Y \times Y$. From $X \neq \emptyset$ there exists $r_0 \in X$. Since Y is a p.o.s. - filter on S, there exists $(r_1, r_2) \in S \times S$ such that $r_1 \leqq p$, $r_1 \leqq r_0$, $r_2 \leqq q$, and $r_2 \leqq r_0$. From the closedness of X, $\{r_1, r_2\} \subset L(r_0) = \{r_0\}^{-} \subset X^{-} = X \subset Y$. Since Y is a p.o.s. -filter, there exists $r_3 \in S$ such that $r_3 \in L(r_1) \cap L(r_2) \subset L(r_0) \subset X \subset Y$. Therefore, $r_3 \in Y$, $r_3 \leqq r_1 \leqq p$, and $r_3 \leqq r_2 \leqq q$. Hence Y is a strong filter. The latter half of the proposition is almost clear from the former half and Proposition 2.1. (From the latter half, $M(Y)=\emptyset$ is equivalent to $ML(X)=\emptyset$.)

3. Relation to Barankin's Structures

For a consideration of the relation between the topological space in Section 2 and Barankin's structures, "$*$-operator" is defined: A mapping, $*$, from $P(S)$ to $P(S)$, is called $*$- operator (star operator), if and only if

(3.1) for any $(X,Y) \in P(S)\times P(S)$, $X \subset Y \Rightarrow X^* \subset Y^*$, and

(3.2) for any $X \in P(S)$, $X^{-*} := (X^-)^* = X^-$.

In this paper the conditions (3.1) and (3.2) are called "star conditions". Clearly the closure operator on S, "-", satisfies star conditions. From star conditions $X^* \subset X^{-*} = X^-$ & $X^{*-} \subset X^- = X^{-*}$, especially, if $X \subset X^*$, then $X^- = X^{-*} = X^{*-}$. Using star conditions, it is shown that, for any $X \in P(S)$,

(3.3) for any $p \in X^* - X$ there exists $q \in X$ such that $p \leqq q$.

Because $p \in X^*-X \subset X^* \subset X^{-*} = X^- = \cup\{ L(q) ;\ q \in X\}$. (It is noted that if $X^* - X = \emptyset$, then (3.3) is trivial.)
If a $*$- operator is given, then the following conceptions are defined.

Definition 3.1. A subset, $X \in P(S)$, is $*$-complete if and only if

(3.4) $X^*=X$.

Therefore, from (3.2) any closed subset in S is $*$-complete.

Definition 3.2. A subset, $X \in P(S)$, is $*$-compride if and only if

$$(3.5)\quad (X^*-X)^- \cap X=\emptyset \ .$$

Therefore from (3.4) any $*$-complete subset in S is $*$-compride. (In Barankin [3] the word "compride" is used in some special sense.)(3.5) is equivalent to

(3.6) for any $(p,q) \in X^* \times X$, if $p \geq q$, then $p \in X$.

Proof: If (3.5) is assumed, then (3.6) is derived. Because if there exists $(p,q) \in X^* \times X$ such that $p \geq q$ & $p \notin X$, then, for such $p \in X^*$, $p \in X^*-X$, therefore $q \in L(p) \subset (X^*-X)^-$ & $q \in X$, this is in contradiction to (3.5). Conversely, under the condition (3.6), if there exists $q \in X$ such that $q \in (X^*-X)^-$, then, from the definition of "closedness", there exists $p \in X^*-X$ such that $p \geq q$, therefore, from (3.6), $p \in X$, this is a contradiction.

If a $*$-operator is given, then the following propositions are derived from star conditions and Definition 3.2.

Proposition 3.1. If X_1 and X_2 are closed in S, i.e., $X_1^- = X_1$ and $X_2^- = X_2$, then $X=X_1-X_2$ is $*$-compride.

Proof: $(X^*-X)^- \cap X=\emptyset$ is proved. If $(X^*-X)^- \cap X\neq\emptyset$, then there exists $q \in X$ such that $q \in (X^*-X)^-$, therefore, there exists $p \in X^*-X$ such that $p \geq q$. From star conditions, $p\in X^*-X \subset X^* \subset X_1^*=(X_1^-)^*=X_1^-=X_1$ and $p \notin X_2$(Because if $p \in X_2$, then $q \in L(p) \subset \{p\}^- \subset X_2^-= X_2$ & $q \in X= X_1- X_2$, this is a contradiction.). Hence, $p \in X= X_1- X_2$, which is in contradiction to $p \in X^*- X$.

Proposition 3.2. For any $X \in P(S)$,

$$(3.7)\quad X^- =(X^*- X)^- \cup (X^- -(X^*- X)^-),$$

where $(X^*-X)^-$ is $*$-complete and $ML((X^*-X)^-) \subset ML(S)$ (especially, if $(X^*-X)^- \neq \emptyset$ and S is a p.o.s. -filter, then $ML((X^*-X)^-)=ML(S)=M(S)$, therefore, $M(S)=\emptyset$ is equivalent to $ML((X^*-X)^-)=\emptyset$), and $X^- -(X^*-X)^-$ is $*$-compride, and if X is $*$- compride, then $X \subset X^- -(X^*-X)^-$.
(Proposition 3.2 is clear from Propositions 3.1, 2.1, and 2.2, and from Definitions 3.1 and 3.2.)

The writer defined star conditions by (3.1) and (3.2), but, of course, other some additive conditions may be satisfied by $*$-operator, for example,

(3.8) $X \subset X^*$ for any $X \in P(S)$, and

(3.9) $X^{**}= X^*$ for any $X \in P(S)$.

If star conditions with (3.8) are called "strict star conditions" and star conditions with (3.9) are called "projective star conditions", then it is almost clear that

(3.10) under strict star conditions, $(X^*)^- = X^- = (X^-)^*$ for any $X \in P(S)$, condition (3.2) is equivalent to condition (3.3), and the closure operator on S, "-", satisfies both projective star conditions and strict star conditions. Therefore, if $*$-operator satisfies (3.9), then in some intuitive sense Proposition 3.2 represents the relation in the "projections in P(S)", "-", and "$*$".

To clarify the relation between the above results and Barankin's structures the conception of the infimum(or greatest lower bound) of a subset $X \subset S$ is used: for any $X \in P(S)$, an element $p \in S$ is called "the infimum of X in S" if and only if $q \geq p$ for all $q \in X$ and, for any $r \in S$, if $q \geq r$ for all $q \in X$, then $r \leq p$. Of course, the infimum of X in S may not exist or, even if exists, it may not exist in X. If the infimum of $X \in P(S)$ in S exists, then it is denoted $\wedge X$ or inf X. Some sets are defined.

Definition 3.3. For any $X \in P(S)$,

$$(3.11)\quad P_S^{(\#)}(X) := \{Y;\ Y \in P(X)\ \&\ Y \neq \emptyset\ \&\ \text{the infimum of } Y \text{ in } S \text{ exists}\},$$

and

$$(3.12)\quad P_S^{(m)}(X) := \{Y;\ Y \in P_S^{(\#)}(X)\ \&\ \#Y \leqq m\},$$

where m is any fixed cardinal number. Clearly, $P_S^{(m)}(X)$ is the set of all the subsets in X which have infimums in S and have the cardinalities equal to or smaller than m, especially $P_S^{(m)}(X) = P_S^{(\#)}(X)$ for all $m \geqq \#X$. $*(\#)$- and $*(m)$- operators are defined.

Definition 3.4. $*(\#)$- and $*(m)$-operators are the mappings from P(S) to P(S) defined by

$$(3.13)\quad X^{*(\#)} := \{\wedge Y;\ Y \in P_S^{(\#)}(X)\},$$

$$(3.14)\quad X^{*(m)} := \{\wedge Y;\ Y \in P_S^{(m)}(X)\},$$

where m is any cardinal number. Clearly, $X^{*(m)}$ is the set of all the infimums of all the subsets in X which have infimums in S and have the cardinalities $\leqq m$. The following Proposition is almost clear from Definitions 3.3 and 3.4.

Proposition 3.3. $*(\#)$- & $*(m)$-operators satisfy "strict and projective star conditions", i.e., the conditions, (3.1), (3.2), (3.8), and (3.9). Therefore, all results for $*$-operator, for example, Propositions 3.1 and 3.2, are applicable to $*(\#)$- & $*(m)$-operators.

Barankin's structures are the special cases of the structures considered above. For example, in [3], the p.o.s. structure, $(S, \leqq)$, is assumed to be "semi-lattice" and "$M(S)=\emptyset$", special cases of Proposition 3.2 are derived, and in that context the correspondences to X^-, $(X^*-X)^-$, and $X^- - (X^*-X)^-$ are called "universe", "eternal", "reach", and so on. Hence, the theory based on Barankin's formulation is the theory of $*(\#)$-opertor on a p.o.s. structure.

Remark: The topology introduced in Section 2 is a well-known topology in set theory, and some authors define the open sets by L(p) instead of U(p)(see, for example, Takeuti, G. [6], p.83.). In such case the results of Section 2 are also valid with obvious modifications by replacing inf with sup, etc..

References

[1]Barankin, E.W. (1964). Probability and east, *Annals of the Institute of Statistical Mathematics, The Twentieth Anniversary Volume* (Vol.XVI), 185-230.
[2]Barankin, E.W. (1979). A theoretical account of an empirical fact in psychology, *Annals of the Japan Association for Philosophy of Science*, Vol.5, No.4, March, 157-168.
[3]Barankin, E. W. (1980). Enlarged mathematical representations for stochastic phenomena, *Recent Developments in Statistical Inference and Data Analysis*, Matusita, K., editor, North-Holland, Amsterdam, 1-8.
[4]Sono, S. (1980). Note, Sugaku Seminar, September, 1980, Nihonhyoronshya, Tokyo, p.96, (in Japanese).
[5]Sono, S. (May, 1985). On Bayesian Inference and Bayes Decision Problems, Thesis for Doctor of Science, Department of Mathematics, Faculty of Science, Tokyo University.
[6]Takeuti, G. (1971). *An Introduction to Modern Set Theory*, Nihon-hyoronshya, Tokyo, (in Japanese).

(Dec. 29, Sun. 1985)