Elect. Comm. in Probab. **7** (2002) 97–100

ELECTRONIC COMMUNICATIONS in PROBABILITY

SUBDIAGONAL AND ALMOST UNIFORM DISTRIBUTIONS

PAUL RESSEL
Mathematisch - Geographische Fakultät, Katholische Universität Eichstätt, 85071 EICHSTÄTT, Germany.
email: `paul.ressel@ku-eichstaett.de`

submitted July 30, 2001 *Final version accepted* September 10, 2001

AMS 2000 Subject classification: 60E99, 60E15, 46A55
Subdiagonal distribution, almost uniform distribution, exchangeable random order

Abstract
A distribution (function) F on $[0,1]$ with $F(t)$ less or equal to t for all t is called subdiagonal. *The extreme subdiagonal distributions are identified as those whose distribution functions are almost surely the identity, or equivalently for which $F \circ F = F$. There exists a close connection to exchangeable random orders on $\{1,2,3,\ldots\}$.*

In connection with the characterization of exchangeable random total orders on $\mathbb{N}$ an interesting class of probability distributions on $[0,1]$ arizes, the socalled *almost uniform* distributions, defined as those $w \in M_+^1([0,1])$ for which $w(\{t \in [0,1] | w([0,t]) = t\}) = 1$, i.e. the distribution function F of w is w–a.s. the identity. The space $\mathcal{W}$ of all almost uniform distributions parametrizes in a canonical way the extreme exchangeable random total orders on $\mathbb{N}$, as shown in [1]. If ν is any probability measure on $\mathbb{R}$ with distribution function G, then the image measure ν^G is almost uniform, see Lemma 3 in [1]. In this paper we show another interesting "extreme" property of $\mathcal{W}$: calling $\mu \in M_+^1([0,1])$ *subdiagonal* if $\mu([0,t]) \le t$ for all $t \in [0,1]$, we prove that the compact and convex set $\mathcal{K}$ of all subdiagonal distributions on $[0,1]$ has precisely the almost uniform distributions as extreme points. A simple example shows that $\mathcal{K}$ is not a simplex.

Lemma. *Let $a < b, c < d$ and*

$$C := \{\varphi : [a,b] \longrightarrow [c,d] \,|\, \varphi \text{ non–decreasing},\ \varphi(a) = c,\ \varphi(b) = d\}\ .$$

Then C is compact and convex ($w.r.$ *to the pointwise topology*) *and*

$$\varphi \in ex(C) \Longleftrightarrow \varphi([a,b]) = \{c,d\}\ .$$

Proof. If $\varphi([a,b]) = \{c,d\}$ then φ is obviously an extreme point. Suppose now that $\varphi \in ex(C)$. We begin with the simple statement that on $[0,1]$ all functions $f_\alpha(x) := x + \alpha(x - x^2)$, for $|\alpha| \le 1$, are strictly increasing from 0 to 1. If $\varphi \in C$ then $\psi := (\varphi - c)/(d - c)$ increases on $[a,b]$ from 0 to 1, hence $\psi_\alpha := f_\alpha \circ \psi$ has the same property. So $\varphi_\alpha := (d-c)\psi_\alpha + c$ increases from c to d, i.e. $\varphi_\alpha \in C$ for $|\alpha| \le 1$; note that $\varphi = \varphi_0$. Now $\psi = \frac{1}{2}(\psi_\alpha + \psi_{-\alpha})$ and

$\varphi = \frac{1}{2}(\varphi_\alpha + \varphi_{-\alpha})$ which shows that φ is not extreme if $\varphi \neq \varphi_\alpha$. We note the equivalences (for $\alpha \neq 0$)

$$\begin{aligned}\varphi = \varphi_\alpha \quad &\Longleftrightarrow \quad \psi = \psi_\alpha \Longleftrightarrow f_\alpha(\psi(t)) = \psi(t) \quad \forall\, t \in [a,b] \\ &\Longleftrightarrow \quad \psi([a,b]) = \{0,1\} \\ &\Longleftrightarrow \quad \varphi([a,b]) = \{c,d\}\ .\end{aligned}$$

Hence $\varphi \in ex(C) \Longrightarrow \varphi = \varphi_\alpha \Longrightarrow \varphi([a,b]) = \{c,d\}$, which was the assertion. $\square$

Remarks.

1.) If φ is right–continuous so are the φ_α.

2.) Since $|f_\alpha(x) - x| \leq |\alpha|/4$ we get the uniform estimate $||\varphi_\alpha - \varphi|| \leq (d-c) \cdot |\alpha|/4$.

3.) $\varphi_\alpha \geq \varphi$ for $\alpha \geq 0,\ \varphi_\alpha \leq \varphi$ for $\alpha \leq 0$.

Both the classes of subdiagonal as well as almost uniform distributions being defined via their distribution functions, we will now work directly with these and consider $\mathcal{K}$ as those distribution functions F on $[0,1]$ for which $F \leq id$. Theorem 2 in [1] can then be reformulated as

$$\mathcal{W} = \{F \in \mathcal{K} \,|\, F \circ F = F\}\ .$$

The announced result is the following:

Theorem. $ex(\mathcal{K}) = \mathcal{W}$.

Proof. "$\supseteq$": Let $F \in \mathcal{W},\ G, H \in \mathcal{K}$ such that $F = \frac{1}{2}(G+H)$. We now make use of the particular "shape" of almost uniform distribution functions: either t is a "diagonal point" of F, i.e. $F(t) = t$, or t is contained in a "flat" of F, i.e. in an interval $]a,b[$ on which F has the constant value a, cf. Lemma 2 in [1]. If $F(t) = t$ then certainly $G(t) = H(t) = t$ as well. If t is in the flat $]a,b[$ of F then

$$F(t) = a = F(a) = G(a) = H(a)$$

so $G(t) \geq a$ and $H(t) \geq a$ and therefore $G(t) = a = H(t)$. We see that $F = G = H$, i.e. $F \in ex(\mathcal{K})$.

"$\subseteq$": Assume $F \in \mathcal{K}$ and $F \circ F \neq F$; we want to show that $F \notin ex(\mathcal{K})$. There is some $s \in [0,1]$ such that $F(F(s)) < F(s)$, implying $0 < s < 1$ and $F(s) < s$. We may and do assume that $F(t) < F(s)$ for all $t < s$, otherwise with $s_0 := \inf\{t < s | F(t) = F(s)\}$ we would still have

$$F(F(s_0)) = F(F(s)) < F(s) = F(s_0)\ .$$

We shall first consider the case that F is constant in a right neighbourhood of s, i.e. for some $v \in]s,1]$ we have $F|[s,v[\ \equiv F(s)$, and again we may and do assume that v is maximal with this property, i.e. $F(v) > F(s)$.

If $F(s-) < F(s)$, then for sufficiently small $\varepsilon > 0$

$$G_\pm(t) := \begin{cases} F(t) \pm \varepsilon, \ t \in [s, v[\\ F(t), \ \text{else} \end{cases}$$

are both subdiagonal, and $F = \frac{1}{2}(G_+ + G_-)$, so $F \notin ex(\mathcal{K})$. If $F(s-) = F(s)$ we put $u := F(s)$ and have a non–degenerate interval $[u, s]$ on which F increases from $F(u)$ to u , and with $F([u, s]) \underset{\neq}{\supset} \{F(u), F(s)\}$ since $F(t) < F(s)$ for $t < s$. We apply the Lemma and Remark 1 to $F|[u, s]$ and get right–continuous functions $F_\alpha : [u, s] \longrightarrow [F(u), u]$ increasing from $F(u)$ to u , $|\alpha| \leq 1$, for which $F_\alpha \neq F$ if $\alpha \neq 0$. Put

$$G_\alpha(t) := \begin{cases} F_\alpha(t), & t \in [u, s] \\ F(t), & \text{else ,} \end{cases}$$

then G_α is a distribution function for $|\alpha| \leq 1$. Since $F = \frac{1}{2}(G_\alpha + G_{-\alpha})$ we are done once we know that G_α is subdiagonal for sufficiently small $|\alpha|$. For this to hold we only need to know that

$$(*) \qquad \inf_{u \leq t \leq s}(t - F(t)) > 0 \ ,$$

cf. Remark 2. Now by right continuity there is some $t_0 \in \,]u, s[$ such that $F(t_0) \leq \frac{1}{2}(u + F(u))$, i.e.

$$t - F(t) \geq u - \frac{u + F(u)}{2} = \frac{u - F(u)}{2} > 0$$

for $t \in [u, t_0]$; and for $t \in [t_0, s]$ we have $F(t) \leq F(s) = u$ and so $t - F(t) \geq t - u \geq t_0 - u$. Together this gives $(*)$.

It remains to consider the case $F(t) > F(s)$ for $t > s$. Choose $v \in \,]s, 1[$ such that $F(v) < \frac{1}{2}(s + F(s))$. Then again F increases on $[s, v]$ from $F(s)$ to $F(v)$ and $F([s, v]) \underset{\neq}{\supset} \{F(s), F(v)\}$ as well as

$$\inf_{s \leq t \leq v}(t - F(t)) \geq s - F(v) > \frac{s - F(s)}{2} > 0 \ ,$$

so that another application of the Lemma shows F to be not extreme in $\mathcal{K}$. $\Box$

In order to see that $\mathcal{K}$ is not a simplex, consider the following four almost uniform distribution functions $F_1, ... F_4$, determined by their resp. set of diagonal points $D_1, ..., D_4$:

$$D_1 := \{0, 1\} \cup \left[\tfrac{1}{4}, \tfrac{3}{4}\right]$$

$$D_2 := \left[0, \tfrac{1}{4}\right] \cup \left\{\tfrac{1}{2}\right\} \cup \left[\tfrac{3}{4}, 1\right]$$

$$D_3 := \left[0, \tfrac{1}{4}\right] \cup \left[\tfrac{1}{2}, \tfrac{3}{4}\right] \cup \{1\}$$

$$D_4 := \{0\} \cup \left[\tfrac{1}{4}, \tfrac{1}{2}\right] \cup \left[\tfrac{3}{4}, 1\right] \ .$$

Then

$$\tfrac{1}{2}(F_1+F_2)=\tfrac{1}{2}(F_3+F_4)=\tfrac{1}{2}id+\tfrac{1}{8}\left(1_{[\frac{1}{4},1]}+1_{[\frac{1}{2},1]}+1_{[\frac{3}{4},1]}+1_{\{1\}}\right)\in\mathcal{K}\,,$$

so the integral representation in $\mathcal{K}$ is not unique.

Let us shortly describe the connection of the above theorem to exchangeable random orders. A (total) order (on $\mathbb{N}$ always) is a subset $V\subseteq\mathbb{N}\times\mathbb{N})$ with $(j,j)\in V$ for all $j\in\mathbb{N}$, with $(i,j),(j,k)\in V\Longrightarrow(i,k)\in V$, and such that either (j,k) or $(k,j)\in V$ for all $j,k\in\mathbb{N}$. The set $\mathcal{V}$ of all total orders is compact and metrisable in its natural topology, and a probability measure μ on $\mathcal{V}$ is called exchangeable if it is invariant under the canonical action of all finite permutations of $\mathbb{N}$ (see [1] for a more detailed description). A particular class of such measures arises in this way: let $X_1,X_2,...$ be an iid–sequence with a distribution $w\in\mathcal{W}$. For any $\emptyset\neq U\subseteq\mathbb{N}^2$ put

$$\mu_w(\{V\in\mathcal{V}|U\subseteq V\}):=P(X_j\leq X_k\ \forall\ (j,k)\in U)$$

This defines (uniquely) an exchangeable random total order, and the main result in [1] shows that $\{\mu_w|w\in\mathcal{W}\}$ is the extreme boundary of the compact and convex set of all exchangeable random total orders (on $\mathbb{N}$), which furthermore is a simplex.

Now, given some exchangeable random total order μ, there is a unique probability measure ν on $\mathcal{W}$ such that

$$\mu=\int\mu_w\,d\nu(w)\,,$$

and ν determines the subdiagonal distribution

$$\overline{\nu}(B):=\int w(B)\,d\nu(w),\quad B\in\mathbb{B}\cap[0,1]\,,$$

which in a way is the „first moment measure“ of ν.

One might believe that only very „simple“ probability values depend on ν via $\overline{\nu}$, but in fact, due to the defining property of almost uniform distributions, also many „higher order“ probabilities have this property. For example

$$\begin{aligned}\mu(1\preceq 2)&=\mu(\{V\in\mathcal{V}|(1,2)\in V\})\\&=\int\mu_w(1\preceq 2)\,d\nu(w)\\&=\int w\otimes w(X_1\leq X_2)\,d\nu(w)\\&=\int\int w(X_1\leq x_2)\,dw(x_2)\,d\nu(w)\\&=\int\int x_2\,dw(x_2)\,d\nu(w)\\&=\int_0^1 x\,d\bar{\nu}(x)\,,\end{aligned}$$

where $X_1,X_2:[0,1]^2\longrightarrow[0,1]$ denote the two projections.

More generally, for different $j, j_1, ..., j_n \in \mathbb{N}$

$$
\begin{aligned}
&\mu(j_1 \preceq j,\ j_2 \preceq j, ..., j_n \preceq j) \\
&= \mu(\{V \in \mathcal{V} | (j_i, j) \in V \text{ for } i = 1, ..., n\}) \\
&= \int w^{n+1}(X_{j_1} \leq X_j, ..., X_{j_n} \leq X_j)\, d\nu(w) \\
&= \int \int w^n(X_{j_1} \leq x, ..., X_{j_n} \leq x)\, dw(x)\, d\nu(w) \\
&= \int \int (w([0, x]))^n\, dw(x)\, d\nu(w) \\
&= \int \int x^n\, dw(x)\, d\nu(w) \\
&= \int_0^1 x^n\, d\bar{\nu}(x)
\end{aligned}
$$

still is a function of $\bar{\nu}$.

References.

[1] Hirth, U. and Ressel, P.: *Exchangeable Random Orders and Almost Uniform Distributions*. J. Theoretical Probability, Vol. **13** n^o 3, 2000, pp. 609 – 634.