Distribution functions and probability measures on linearly ordered topological spaces

J. F. Gálvez-Rodríguez and M. A. Sánchez-Granero
Departamento de Matemáticas, Universidad de Almería

IX SIMPOSIO de Investigación EN CIENCIAS EXPERIMENTALES

Abstract

In [2] we describe a theory of a cumulative distribution function (in short, cdf) on a separable linearly ordered topological space (LOTS) from a probability measure defined in this space. This function can be extended to the Dedekind-MacNeille completion of the space where it does make sense to define the pseudo-inverse (see [3]). Moreover, we study the properties of both functions (the cdf and the pseudo-inverse) and get results that are similar to those which are well-known in the classical case. For example, the pseudo-inverse of a cdf allows us to generate samples of a distribution and give us the chance to calculate integrals with respect to the related probability measure. Finally, in [1] we give some conditions such that there is an equivalence between probability measures and distribution functions defined on a separable LOTS, like it happens in the classical case. What is more, we prove that the pseudo-inverse of the cumulative distribution function is univocally related to a probability measure. From this theory, some applications have arisen, such as a goodness-of-fit test.

Definition and properties of a cdf

Definition. Given a probability measure μ on a separable LOTS, $(X,\leq)$, its cdf is a function $F : X \to [0,1]$ defined by $F(x) = \mu(\leq x)$, for each $x \in X$, where $(\leq x) = \{y \in X : y \leq x\}$.

Properties.

1. F is monotonically non-decreasing.
2. F is right τ-continuous (τ is the order topology in X).
3. $\sup F(X) = 1$.
4. If there does not exist $\min X$, then $\inf F(X) = 0$.

Definition. $F_- : X \to [0,1]$ is defined by $F_-(x) = \mu(< x)$, for each $x \in X$, where $(< x) = \{y \in X : y < x\}$.

Getting the measure of a set

Let $a, b \in X$ with $a < b$, then:

- $\mu(\{a\}) = F(a) - F_-(a)$.
- $\mu(]a,b[) = F_-(b) - F(a)$.
- $\mu([a,b[) = F_-(b) - F_-(a)$.
- $\mu(]a,b]) = F(b) - F(a)$.
- $\mu([a,b]) = F(b) - F_-(a)$.

Discontinuities of a cdf

- If $\mu(\{x\}) = 0$, for each $x \in X$, then F is τ-continuous.
- The set of discontinuity points of F with respect to τ is countable.

Dedekind-MacNeille completion of a separable LOTS

Definition. Given a partially ordered set X, the Dedekind-MacNeille completion of X is defined to be $DM(X) = \{A \subseteq X : A = (A^u)^l\}$ ordered by inclusion ($A \leq B$ if, and only if $A \subseteq B$), where A^u (resp. A^l) is the set of upper (resp. lower) bounds of A.

$\phi : X \to DM(X)$ is an embedding defined by $\phi(x) = (\leq x)$, for each $x \in X$.

Proposition. $DM(X)$ is, indeed, a compactification of X and F can be extended to a cdf, $\widetilde{F}$, on $DM(X)$ by defining $\widetilde{F} : DM(X) \to [0,1]$ by $\widetilde{F}(A) = \inf F(A^u)$, for each $A \in DM(X)$.

Relationship between μ and F

Theorem. Let X be a separable LOTS such that $DM(X) \setminus \phi(X)$ is countable and $F : X \to [0,1]$ a monotonically non-decreasing and right τ-continuous function satisfying $\sup F(X) = 1$ and $\sup F(A) = \inf F(A^u)$, for each $A \in DM(X)$. Moreover, $\inf F(X) = 0$ if there does not exist the minimum of X. Then there exists a unique probability measure on X, μ, such that $F = F_\mu$.

Remark. Note that μ and F are univocally determined.

Corollary. Let X be a separable LOTS such that $DM(X) \setminus \phi(X)$ is countable and let $F_- : X \to [0,1]$ be a monotonically non-decreasing, left τ-continuous function such that $\inf F_-(X) = 0$ and $\sup F_-(A) = \inf F_-(A^u)$, for each $A \in DM(X)$. Moreover, $\sup F_-(X) = 1$ if there does not exist the maximum of X. Then there exists a unique probability measure on X, μ, such that $F_{\mu-} = F_-$.

Definition and properties of the inverse

Definition. Let F be a cdf. We define the pseudo-inverse of F by $G : [0,1] \to DM(X)$ given by $G(r) = \{x \in X : F(x) \geq r\}^l$, for each $r \in [0,1]$.

Properties.

1. G is monotically non-decreasing.
2. G is left τ-continuous.
3. $G(r) \leq \phi(x)$ if, and only if $r \leq F(x)$, for each $x \in X$ and each $r \in [0,1]$.

Relationship between μ and G

Theorem. Let X be a separable LOTS such that $DM(X) \setminus \phi(X)$ is countable and let $G : [0,1] \to DM(X)$ be a monotonically non-decreasing and left τ-continuous function such that $\sup G^{-1}(< A) = \inf G^{-1}(> A)$, for each $A \in DM(X) \setminus \phi(X)$, $G(0) = \min DM(X)$, $G^{-1}(\max DM(X)) \subseteq \{1\}$ if there does not exist the maximum of X and $G^{-1}(\min DM(X)) = \{0\}$ if there does not exist the minimum of X. Then there exists a unique probability measure on X, μ, such that G is the pseudo-inverse of F_μ.

Remark. The pseudo-inverse let us generate samples with respect to the probability measure μ by following the classical procedure. Note, also, that G and μ are univocally determined.

Decomposition of a cdf

Theorem. Each cdf F_μ defined on a separable LOTS X such that $DM(X) \setminus \phi(X)$ is countable, can be decomposed into $F_\mu = \alpha F_d + (1-\alpha)F_c$ with $0 \leq \alpha \leq 1$, where F_d is a step cdf, and F_c is a cdf satisfying that $F_{c-} = F_c$. Moreover, the decomposition is unique.

Remark. Under the above conditions we can decompose each cdf even if it is defined on an n-dimentional space.

A goodness-of-fit test in a LOTS

Suppose that we are given a random sample on a separable LOTS according to a certain cumulative distribution function, F. Our purpose is testing if F comes from a certain distribution. Let us denote by F_n the empirical cumulative distribution function of the sample. If we define the statistic $D_n = \sup_{x \in X} |F_n(x) - F(x)|$,

Theorem. Let X be a separable LOTS and suppose that μ is a probability measure on X such that $\mu(\{x\}) = 0$. Then we can write $D_n = \max_{0 \leq r \leq 1} |H_n(r) - r|$, where H_n is the empirical cdf of the image by F of the sample.

Corollary. Given a separable LOTS, X, and $n \in \mathbb{N}$, the distribution of D_n is the same for each cdf, F_μ, satisfying that $\mu(\{x\}) = 0$, for each $x \in X$.

References

[1] J. F. Gálvez-Rodríguez and M. A. Sánchez-Granero, *Equivalence between distribution functions and probability measures on a LOTS*, preprint.

[2] J. F. Gálvez-Rodríguez and M. A. Sánchez-Granero, *The distribution function of a probability measure on a linearly ordered topological space*, Mathematics, 7(9) (2019), 864.

[3] J. F. Gálvez-Rodríguez and M. A. Sánchez-Granero, *The distribution function of a probability measure on the Dedekind-MacNeille completion*, Topology and its Applications, (2019), 107010.