Disintegration into conditional measures: Rokhlin's theorem

Let Z be a compact metric space, μ be a Borel probability measure on Z, and $\mathcal{P}$ be a partition of Z into measurable subsets. Let $\pi : Z \to \mathcal{P}$ be the map associating to each $z \in Z$ the atom $P \in \mathcal{P}$ that contains it. By definition, $\mathcal{Q}$ is a measurable subset of $\mathcal{P}$ if and only if $\pi^{-1}(\mathcal{Q})$ is a measurable subset of Z. Let $\hat{\mu}$ be the push-forward of μ under π, in other words, $\hat{\mu}$ is the probability measure on $\mathcal{P}$ defined by $\hat{\mu}(\mathcal{Q}) = \mu(\pi^{-1}(\mathcal{Q}))$ for every measurable set $\mathcal{Q} \subset \mathcal{P}$.

Definition 1. A *system of conditional measures of μ with respect to $\mathcal{P}$* is a family $(\mu_P)_{P\in\mathcal{P}}$ of probability measures on Z such that

1. $\mu_P(P) = 1$ for $\hat{\mu}$-almost every $p \in \mathcal{P}$;
2. given any continuous $\varphi : Z \to \mathbb{R}$, the function $\mathcal{P} \ni P \mapsto \int \varphi \, d\mu_P$ is measurable and $\int \varphi \, d\mu = \int \left(\int \varphi \, d\mu_P\right) d\hat{\mu}(P)$.

Proposition 1. *If $(\mu_P)_{P\in\mathcal{P}}$ is a system of conditional measures of μ relative to the partition $\mathcal{P}$, then $\mathcal{P} \ni P \mapsto \int \psi \, d\mu_P$ is measurable and $\int \psi \, d\mu = \int \left(\int \psi \, d\mu_P\right) d\hat{\mu}(P)$, for any μ-integrable function $\psi : Z \to \mathbb{R}$.*

Proof. The class of functions that satisfy the conclusion of the lemma contains all the continuous functions, and is closed under dominated pointwise convergence. Therefore, it contains all bounded measurable functions. $\square$

In particular, $P \mapsto \mu_P(E)$ is mesurable, and $\mu(E) = \int \mu_P(E) \, d\hat{\mu}(P)$, for any measurable set $E \subset Z$.

Conditional measures, when they exist, are unique almost everywhere:

Proposition 2. *If $(\mu_P)_{P\in\mathcal{P}}$ and $(\nu_P)_{P\in\mathcal{P}}$ are two systems of conditional measures of μ with respect to $\mathcal{P}$, then $\mu_P = \nu_P$ for $\hat{\mu}$-almost every $P \in \mathcal{P}$.*

Proof. Suppose otherwise, that is, there exists a measurable set $\mathcal{Q}_0 \subset \mathcal{P}$ with $\hat{\mu}(\mathcal{Q}_0) > 0$ such that $\mu_P \neq \nu_P$ for every $P \in \mathcal{Q}_0$. Using the fact that $C^0(Z, \mathbb{R})$ admits countable subsets dense with respect to the uniform norm, one gets that there exists $\varphi \in C^0(Z, \mathbb{R})$ and a subset $\mathcal{Q}$ of $\mathcal{Q}_0$ such that $\hat{\mu}(\mathcal{Q}) > 0$ and (interchanging the roles of μ_P and ν_P, if necessary) $\int \varphi \, d\mu_P > \int \varphi \, d\nu_P$ for every $P \in \mathcal{Q}$. Then

$$\int_{\mathcal{Q}} \left(\int \varphi \, d\mu_P \right) d\hat{\mu}(P) > \int_{\mathcal{Q}} \left(\int \varphi \, d\nu_P \right) d\hat{\mu}(P). \tag{1}$$

On the other hand, by Proposition 1,

$$\int (\varphi \mathcal{X}_{\pi^{-1}(\mathcal{Q})}) \, d\mu = \int \left(\int (\varphi \mathcal{X}_{\pi^{-1}(\mathcal{Q})}) \, d\mu_P \right) d\hat{\mu}(P).$$

By assumption $\mu_P(P) = 1$ for $\hat{\mu}$-almost every $P \in \mathcal{P}$. For any such P, we have

$$\int (\varphi \mathcal{X}_{\pi^{-1}(\mathcal{Q})}) \, d\mu_P = \mathcal{X}_{\mathcal{Q}}(P) \int \varphi \, d\mu_P.$$

Therefore,

$$\int (\varphi \mathcal{X}_{\pi^{-1}(\mathcal{Q})}) \, d\mu = \int \left(\mathcal{X}_{\mathcal{Q}}(P) \int \varphi \, d\mu_P \right) d\hat{\mu}(P) = \int_{\mathcal{Q}} \left(\int \varphi \, d\mu_P \right) d\hat{\mu}(P).$$

Analogously, we find

$$\int (\varphi \mathcal{X}_{\pi^{-1}(\mathcal{Q})}) \, d\mu = \int_{\mathcal{Q}} \left(\int \varphi \, d\nu_P \right) d\hat{\mu}(P).$$

These two last equalities contradict (1). Therefore, $\mu_P = \nu_P$ for $\hat{\mu}$-almost every P, as claimed. $\square$

Definition 2. $\mathcal{P}$ is a *measurable partition* if there exist measurable subsets $E_1, E_2, \ldots, E_n, \ldots$ of Z such that

$$\mathcal{P} = \{E_1, Z \setminus E_1\} \vee \{E_2, Z \setminus E_2\} \vee \cdots \vee \{E_n, Z \setminus E_n\} \vee \cdots \qquad \text{mod } 0.$$

In other words, there exists some full μ-measure subset $F_0 \subset Z$ such that, given any atom P of $\mathcal{P}$ we may write

$$P \cap F_0 = E_1^* \cap E_2^* \cap \cdots \cap E_n^* \cap \cdots \cap F_0 \tag{2}$$

where E_j^* is either E_j or its complement $Z \setminus E_j$, for every $j \geq 1$.

Example 1. Every finite or countable partition is a measurable partition. In fact, $\mathcal{P}$ is measurable if and only if there exists a non-decreasing sequence of finite or countable partitions $\mathcal{P}_1 \preceq \mathcal{P}_2 \preceq \cdots \preceq \mathcal{P}_n \preceq \cdots$ such that $\mathcal{P} = \vee_{n=1}^{\infty} \mathcal{P}_n \mod 0$.

Example 2. Let $Z = X \times Y$, where X and Y are compact metric spaces, and $\mathcal{P}$ be the partition of Z into horizontal lines $X \times \{y\}$, $y \in Y$. Then $\mathcal{P}$ is a measurable partition of Z.

Theorem 3. *(Rokhlin [R]) If $\mathcal{P}$ is a measurable partition, then there exists some system of conditional measures of μ relative to $\mathcal{P}$.*

Proof. For the purpose of the conclusion of the theorem, we may replace the space Z by any full measure subset. So, it is no restriction to suppose that the set F_0 in (2) actually coincides with Z, and we do so in all that follows. Let ψ be any bounded measurable real function on Z. For each $n \geq 1$ let

$$\mathcal{P}_n = \{E_1, Z \setminus E_1\} \vee \{E_2, Z \setminus E_2\} \vee \cdots \vee \{E_n, Z \setminus E_n\}$$

that is, $\mathcal{P}_n$ is the partition of Z whose atoms are the sets $E_1^* \cap \cdots \cap E_n^*$, with $E_j^* = E_j$ or $E_j^* = Z \setminus E_j$, for each $1 \leq j \leq n$. Define $\tilde{\psi}_n : Z \to \mathbb{R}$ as follows. If the atom $P_n(z)$ of $\mathcal{P}_n$ that contains z has positive μ-measure, then

$$\tilde{\psi}_n(z) = \frac{1}{\mu(P_n(z))} \int_{P_n(z)} \psi \, d\mu. \tag{3}$$

Otherwise, $\tilde{\psi}_n(z) = 0$. Clearly, the second case in the definition of $\tilde{\psi}_n$ applies only to a zero μ-measure set of points.

Lemma 4. *Given any bounded measurable function $\psi : Z \to \mathbb{R}$, there exists a full μ-measure subset $F = F(\psi)$ of Z such that $\tilde{\psi}_n(z)$, $n \geq 1$, converges to some real number $\tilde{\psi}(z)$, for every $z \in F$.*

Proof. We may always write $\psi = \psi^+ - \psi^-$, where $\psi^\pm$ are measurable, bounded, and non-negative: for instance, $\psi^\pm = (|\psi| \pm \psi)/2$. Then $\tilde{\psi}_n = \tilde{\psi}_n^+ - \tilde{\psi}_n^-$ for $n \geq 1$, and so the conclusion holds for ψ if it holds for ψ^+ and ψ^-. This shows that it is no restriction to assume that ψ is non-negative. We do so from now on.

For any $\alpha < \beta$, let $S(\alpha, \beta)$ be the set of points $z \in Z$ such that

$$\liminf \tilde{\psi}_n(z) < \alpha < \beta < \limsup \tilde{\psi}_n(z).$$

Clearly, given $z \in Z$, the sequence $\tilde{\psi}_n(z)$ diverges if and only if z is in $S(\alpha, \beta)$ for some pair of rational numbers α and β. So, the lemma will follow if we show that $S = S(\alpha, \beta)$ has zero μ-measure for all α and β.

For each $z \in S$ fix some sequence of integers $1 \leq a(1) < b(1) < \cdots < a(i) < b(i) < \cdots$ such that

$$\tilde{\psi}_{a(i)}(z) < \alpha \quad \text{and} \quad \tilde{\psi}_{b(i)}(z) > \beta \quad \text{for every } i \geq 1.$$

Define A_i to be the union of the partition sets $P_{a(i)}(z)$, and B_i to be the union of the partition sets $P_{b(i)}(z)$ obtained in this way, for all the points $z \in S$. By construction,

$$S \subset A_{i+1} \subset B_i \subset A_i \quad \text{for every } i \geq 1.$$

In particular, S is contained in the set

$$\widetilde{S} = \bigcap_{i=1}^{\infty} B_i \cap F_0 = \bigcap_{i=1}^{\infty} A_i \cap F_0 \,.$$

Given any two of the sets $P_{a(i)}(z)$ that form A_i, either they are disjoint or one is contained in the other. This is because $\mathcal{P}_n$, $n \geq 1$, is a non-decreasing sequence of partitions. Consequently, A_i may be written as a two-by-two disjoint union of such sets $P_{a(i)}(z)$. Hence,

$$\int_{A_i} \psi \, d\mu = \sum_{P_{a(i)}(z)} \int_{P_{a(i)}} \psi \, d\mu < \sum_{P_{a(i)}(z)} a\mu(P_{a(i)}) = \alpha\mu(A_i),$$

for any $i \geq 1$ (the sums are over that disjoint union). Analogously,

$$\int_{B_i} \psi \, d\mu = \sum_{P_{b(i)}(z)} \int_{P_{b(i)}} \psi \, d\mu > \sum_{P_{b(i)}(z)} b\mu(P_{b(i)}) = \beta\mu(B_i).$$

Since $A_i \supset B_i$ and we are assuming $\psi \geq 0$, it follows that

$$\alpha\mu(A_i) > \int_{A_i} \psi \, d\mu \geq \int_{B_i} \psi \, d\mu > \beta\mu(B_i),$$

for every $i \geq 1$. Taking the limit as $i \to \infty$, we find

$$\alpha\mu(\widetilde{S}) \geq \beta\mu(\widetilde{S}).$$

This implies that $\mu(\widetilde{S}) = 0$, and so $S \subset \widetilde{S}$ also has zero μ-measure. $\square$

Given any bounded measurable function $\psi : Z \to \mathbb{R}$, we shall represent as $e_n(\psi)$, $e(\psi)$, respectively, the functions $\tilde{\psi}_n$, $\tilde{\psi}$ defined by (3) and Lemma 4.

Let $\{\varphi_k : k \in \mathbb{N}\}$ be some countable dense subset of $C^0(Z, \mathbb{R})$, and let

$$F_* = \bigcap_{k=1}^{\infty} F(\varphi_k),$$

where $F(\varphi_k)$ is as given by Lemma 4.

Lemma 5. *Given any continuous function $\varphi : Z \to \mathbb{R}$, the sequence $e_n(\varphi)(z)$ converges to $e(\varphi)(z)$ as $n \to \infty$, for every $z \in F_*$.*

Proof. The class of functions that satisfy the conclusion of the lemma is closed under uniform convergence and, by definition, it contains a dense subset of the space of all continuous functions. □

Let $\varphi : Z \to \mathbb{R}$ be continuous. By construction, $e_n(\varphi)$ is constant on each $P_n \in \mathcal{P}_n$, and so it is also constant on each atom P of $\mathcal{P}$, for every $n \geq 1$. Therefore, $e(\varphi)$ is constant on $P \cap F_*$ for every $P \in \mathcal{P}$. Let $e_n(\varphi)(P_n)$ represent the value of $e_n(\varphi)$ on each $P_n \in \mathcal{P}_n$. Similarly, $e(\varphi)(P)$ represents the value of $e(\varphi)$ on $P \cap F_*$ whenever the latter set is non-empty. Then, since (3) defines $e_n(\varphi)$ on a full μ-measure subset of Z,

$$\int \varphi \, d\mu = \sum_{\mu(P_n)>0} \int_{P_n} \varphi \, d\mu = \sum_{\mu(P_n)>0} \mu(P_n)\, e_n(\varphi)(P_n) = \int e_n(\varphi)\, d\mu.$$

Observe also that $|e_n(\varphi)| \leq \sup |\varphi| < \infty$ for every $n \geq 1$. Therefore, we may use the dominated convergence theorem to conclude that

$$\int \varphi \, d\mu = \int e(\varphi)\, d\mu. \tag{4}$$

Now we are in a position to construct a system of conditional measures of μ. Let P be any atom of $\mathcal{P}$ such that $P \cap F_*$ is non-empty. It is easy to see that

$$C^0(Z, \mathbb{R}) \ni \varphi \to e(\varphi)(P) \in \mathbb{R}$$

is a non-negative linear functional on $C^0(Z, \mathbb{R})$. By the Riesz-Markov theorem, there exists a unique probability measure μ_P on Z such that

$$\int \varphi \, d\mu_P = e(\varphi)(P). \tag{5}$$

For completeness, we should define μ_P also when P does not intersect F_*. In this case we let μ_P be any probability measure on Z: since the set of all these atoms P has zero $\hat{\mu}$-measure in $\mathcal{P}$ (in other words, their union has zero μ-measure in Z), the choice is not relevant. In view of these definitions, (4) may be rewritten as

$$\int \varphi \, d\mu = \int \left(\int \varphi \, d\mu_P \right) \, d\hat{\mu}(P),$$

the fact that $\mathcal{P} \ni P \mapsto \int \varphi \, d\mu_P$ is a measurable function being a direct consequence of (5). Therefore, to conclude that $(\mu_P)_{P\in\mathcal{P}}$ do form a system of conditional measures of μ with respect to $\mathcal{P}$ we only have to prove

Lemma 6. $\mu_P(P) = 1$ *for* $\hat{\mu}$*-almost every* $P \in \mathcal{P}$.

We use the following auxiliary result.

Lemma 7. *Given any bounded measurable function* $\psi : Z \to \mathbb{R}$ *there exists a full* $\hat{\mu}$*-measure set* $\mathcal{F}(\psi) \subset \mathcal{P}$ *such that the set* $P \cap F_*$ *is non-empty and* $\int \psi \, d\mu_P = e(\psi)(P)$*, for any* $P \in \mathcal{F}(\psi)$.

Proof. The class of functions that satisfy the conclusion of the lemma contains all the continuous functions, and is closed under dominated pointwise convergence. Therefore, it contains all bounded measurable functions. □

Now we can prove Lemma 6:

Proof. Define $\mathcal{F}_* = \cap_{k,P_k} \mathcal{F}(\mathcal{X}_{P_k})$, where the intersection is over the set of all the atoms $P_k \in \mathcal{P}_k$, and every $k \geq 1$. Since this is a countable set, $\mathcal{F}_*$ has full $\hat{\mu}$-measure. We claim that the conclusion of the lemma holds for every $P \in \mathcal{F}_*$. Indeed, let $k \geq 1$ and P_k be the element of $\mathcal{P}_k$ that contains P. By the definition of $\mathcal{F}_*$

$$\mu_P(P_k) = \int \mathcal{X}_{P_k} \, d\mu_P = e(\mathcal{X}_{P_k})(P). \tag{6}$$

For each $n \geq 1$, let P_n be the atom of $\mathcal{P}_n$ that contains P. Given any $z \in P \cap F_*$,

$$e_n(\mathcal{X}_{P_k})(z) = \frac{1}{\mu(P_n)} \int_{P_n} \mathcal{X}_{P_k} \, d\mu.$$

Now, for any $n \geq k$ we have $P_n \subset P_k$, and then the last term is equal to 1. Therefore,

$$e(\mathcal{X}_{P_k})(P) = e(\mathcal{X}_{P_k})(z) = \lim_{n\to\infty} e_n(\mathcal{X}_{P_k})(z) = 1.$$

Replacing this in (6) we get $\mu_P(\mathcal{P}_k) = 1$ for every $k \geq 1$. Finally,

$$\mu_P(P) = \lim_{k\to\infty} \mu_P(P_k) = 1$$

because the P_k, $k \geq 1$, are a decreasing sequence whose intersection is P. $\square$

The proof of Theorem 3 is complete. $\square$

Example 3. Let Z be the 2-dimensional torus, α be some irrational number, and $\mathcal{P}$ be the partition of Z into the straight lines of slope m. Then $\mathcal{P}$ is not a measurable partition. Indeed, the Haar (Lebesgue) measure on Z admits no system of conditional measures with respect to $\mathcal{P}$.

Example 4. (ergodic decomposition) Let $f : Z \to Z$ be a continuous transformation on a compact metric space Z, and B_f be the subset of points $z \in Z$ such that time averages are well-defined on the orbit of z: given any continuous function $\varphi : Z \to \mathbb{R}$, the sequence

$$\frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(z))$$

converges to some $\tilde{\varphi}(z) \in \mathbb{R}$ when $n \to \infty$. Let $\mathcal{P}$ be the partition of Z defined by (i) $Z \setminus B_f$ is an atom of $\mathcal{P}$ and (ii) two points z_1, $_2$ in B_f are in the same atom of $\mathcal{P}$ if and only if they have the same time averages: $\tilde{\varphi}(z_1) = \tilde{\varphi}(z_2)$ for every continuous function φ. Then $\mathcal{P}$ is a measurable partition, with respect to any probability measure μ in Z. If μ is f-invariant, then $\mu(Z \setminus B_f) = 0$ and any system of conditional measures $(\mu_P)_P$ of μ relative to $\mathcal{P}$ is such that μ_P is f-invariant and ergodic for $\hat{\mu}$-almost every $P \in \mathcal{P}$.

References

[R] V. A. Rokhlin, *On the fundamental ideas of measure theory*, Transl. Amer. Math. Soc., Series 1, 10 (1962), 1–52.