A Note on Common Prior*

Miklós Pintér
Corvinus University of Budapest
Department of Mathematics
13-15. Fővám tér, Budapest
Hungary H-1093
miklos.pinter@uni-corvinus.hu

October 20, 2005

Abstract

Harsányi introduced the concept of type space in an intuitive way. Later Heifetz and Samet formalized it. Harsányi used conditional probabilities to model the beliefs of the players, Heifetz and Samet avoided using conditional probabilities formally. We show that in both cases the concept of transition probability can reproduce the models, moreover, the transition probability approach fits to both Harsányi's intuition and the formalization of Heifetz and Samet. As a consequence, our results suggest that the concept of common prior is not appropriate to determine the players' beliefs. Two examples are also given.

1 Introduction

In his celebrated paper Harsányi [2] introduced the concept of types space to model games with incomplete information played by Bayesian players. In his original work and in textbooks as well conditional probabilities are used for deducing the beliefs of the players from the common prior.

It is known that the conditional probability of an arbitrary fixed event is a set of random variables (the Radon-Nikodym derivative) which is defined only up to sets of measure zero, so it is an element of L^1. Therefore, there can be many variants of the conditional probability, which are coincide "only" *almost surely.* The main disadvantage of conditional probability is that it is not defined that, which variants of the conditional probability are the players' beliefs. In other words, the common prior does not contain the information that the players' beliefs do.

Furthermore, the transition probability, which can be defined as a variant of the conditional probability, can determine exactly the players' beliefs at any type, so it is well defined, it goes with Harsányi's intuition.

Heifetz and Samet [3] formalized the so called *Harsányi type space.* Their formalization avoid to use the common prior assumption and conditional probabilities, so it does not share the problems mentioned above. We show that

*This work was supported by grant OTKA T046194.

their model can be built on the concept of transition probability, therefore their model also fits to our suggestion.

Our two examples illustrate the inappropriateness of conditional probability. Both examples use infinite type sets, none of our counterexamples can be simplified to finite model, in both examples the measurable structure does matter. We interpret this feature of our examples, as these examples are based on the cognitive constraints of the players.

Nevertheless, our counterexamples shed some light on a general problem of conditional probability. It is commonly believed that conditional probability answers question like: what is the probability of event A if event B happens. However, the conditional probability can answer the questions for which the probability space contains the information needed. However, in general the model (probability space) does not contains the information for the answer. If we want to get answers for questions like mentioned above, then first we have to build the answer into the model, so we need to know the answer before. See the following short example as an illustration.

We have a 6cl glass of cocktail which cocktail is two colored, the upper half (3cl) is green and the lower half is red. We use a plastic pipe to get 1cl sample from the drink. We put it in the drink randomly, and we get a one colored sample (we assume that if the lower end of the plastic pipe is in the green part of the drink, then the whole sample is one colored, green, and if it is in the red part of the drink, then the sample is also one colored, red. What is the probability that we get a 1cl red sample, if we previously got an other 1cl sample which was green? In other words, what is the probability of getting a red sample conditional on we have already got a green sample before?

One could calculate in such a way that if we got a green sample before, then there are 3cl red and 2cl green components in the drink, so the answer is 3/5. However, this argument is totally wrong, since we do not know what happens after we get the first sample, e.g. it is possible that if the balance between the two components ends, then the drink becomes a yellow mass. The assumption that, there will be 3cl red and 2cl green components in the drink is not from the description of the problem. Actually, the description of the problem does not contain the information what happens after we get the first sample, so the correct answer is: we cannot answer this question.

Summing up, the conditional probability is much less than it is commonly believed. One can think that our example above is only a trick with the words, but our examples later demonstrate that it is not the case. In spite of our examples' infinitary feature, as we demonstrated, the informational problem discussed above appears in finite models as well. Our examples only emphasize the problem, which we cannot formalize in a finite model.

In the next section we give our two examples. In the third section we show that Heifetz and Samet's definition of type space can be deduced from transition probabilities (Theorem 21.). In the last section we summarize our results, and introduce a slight modification of Heifetz and Samet's definition of Harsányi type space, which formalization emphasizes our suggestion: the concept of transition probability.

2 The examples

The following two examples are counterexamples, and they show two kinds of failure of conditional probability. We present here two, instead of one counterexample, because the second counterexample calls for the use of the axiom of choice, and we want to give counterexamples with and without axiom of choice as well.

2.1 The first example

The following example demonstrates that, conditional probabilities cannot be used for determining the players' beliefs.

Denote i, j the two players, and assume that the uncertainties stem from the players only. The sets the players' types are T_i, T_j such that $T_i = T_j = [0,1]$. The measurable structures of the type sets $\mathcal{M}_i$, $\mathcal{M}_j$ are the Borel sets of $[0,1]$, therefore, $(T_i, \mathcal{M}_i) = (T_j, \mathcal{M}_j)$ are the type sets of the players respectively. Let $(T, \mathcal{M}) = (T_i \times T_j, \mathcal{M}_i \otimes \mathcal{M}_j)$ be the set of states of the world. Denote p the common prior, where p is the Lebesgue measure on $(T, \mathcal{M})$, hence $(T, \mathcal{M}, p)$ is a probability measure space.

It is our aim to determine the conditional probabilities on any point of T_i (we define what we mean on this later). The problem is that, $p(pr_i^{-1}(\{t_i\})) = 0$ $\forall t_i \in T_i$ (where pr_i denotes the coordinate projection from T to T_i), so we cannot calculate the conditional probability by the intuitive way, we mean the formula $p(A \mid t_i) = \dfrac{p(A \cap pr_i^{-1}(\{t_i\}))}{p(pr_i^{-1}(\{t_i\}))}$ is meaningless, since the denominator is zero, where $A \in \mathcal{M}$.

The information, which is available for player i can be represented by the σ-algebra $pr_i^{-1}(\mathcal{M}_i)$, therefore we are interested in a conditional probability on this σ-algebra.

Definition 1. *The function $f'_i : T \times \mathcal{M} \to \mathbb{R}$ is a conditional probability on $pr_i^{-1}(\mathcal{M}_i)$ (henceforth $f'_i(\cdot \mid pr_i^{-1}(\mathcal{M}_i))$), if $\forall A \in \mathcal{M}$*

1. *$f'_i(\cdot, A)$ is $pr_i^{-1}(\mathcal{M}_i)$-measurable,*
2. *$p(A \cap B) = \int_B f'_i(\cdot, A) \; dp \; \forall B \in pr_i^{-1}(\mathcal{M}_i)$.*

The existence of conditional probability is a direct corollary of the existence of the Radon-Nikodym derivative. However, it is easy to see that, the conditional probability is defined up to measure zero sets, so strictly speaking, the conditional probability is a set of functions, a point in $L^1(p)$.

In our special case (product space case) we can define the conditional probability in an alternative way:

Definition 2. *Let $p_i(B) \triangleq p(pr_i^{-1}(B))$ $\forall B \in \mathcal{M}_i$. It is easy to see that p_i is a probability measure on $(T_i, \mathcal{M}_i)$. The function $f_i : T_i \times \mathcal{M} \to \mathbb{R}$ is a conditional probability on $\mathcal{M}_i$ (henceforth $f_i(\cdot \mid \mathcal{M}_i)$), if $\forall A \in \mathcal{M}$*

1. *$f_i(\cdot, A)$ is $\mathcal{M}_i$-measurable,*
2. *$p(A \cap pr_i^{-1}(B)) = \int_B f_i(\cdot, A) \; dp_i \; \forall B \in \mathcal{M}_i$.*

Since the conditional probability is some set of functions, we can speak about variants of it. However, there are variants of the conditional probability $f_i(\cdot \mid \mathcal{M}_i)$ which are not measures at a fixed point $x \in T_i$. As a result, the conditional probability is not a probability measure at all.

Definition 3. *Conditional probability $f_i(\cdot \mid \mathcal{M}_i)$ is regular, if there is such a function $f : T_i \times \mathcal{M} \to [0,1]$ that*

1. $f(\cdot, A) = f_i(A \mid \mathcal{M}_i)$ *p_i-a.s.* $\forall A \in \mathcal{M}$,
2. $f(x, \cdot)$ *is a probability measure on* $\mathcal{M}$ $\forall x \in T_i$.

In the next step, we define such two variants of our conditional probability that they meet points *1.* and *2.* in the definition above. We define them on the so called measurable rectangles (sets with form $A_i \times A_j$, $A_i \in \mathcal{M}_i$, $A_j \in \mathcal{M}_j$) only, since at any fixed $x \in T_i$ these functions are probability measures, so these measures can be extended to $\mathcal{M}$ in a unique way. We denote these variants by f_i^1 and f_i^2.

$\forall A = A_i \times A_j$ let

$$f_i^1(x, A) \mathrel{\hat{=}} \begin{cases} p(pr_j^{-1}(A_j)), & \text{if } x \in A_i \\ 0, & \text{otherwise} \end{cases}.$$

Furthermore, let $t_i^* \in T_i$, $t_j^* \in T_j$ be arbitrary fixed, $\forall A = A_i \times A_j$

$$f_i^2(x, A) \mathrel{\hat{=}} \begin{cases} f_i^1(x, A), & \text{if } x \neq t_i^* \\ 1, & \text{if } x = t_i^* \text{ and } t_j^* \in A_j \\ 0, & \text{otherwise} \end{cases}.$$

It is easy to see that f_i^1 coincides with f_i^2 except on the point t_i^*, where f_i^2 is a Dirac measure concentrated on $t_i^* \times t_j^*$. It is also obvious that $f_i^1(\cdot, A) = f_i^2(\cdot, A)$ *p_i-a.s.* $\forall A \in \mathcal{M}$.

Now, we repeat our above argument, but we use density functions instead of probability measures (probability distributions).

Let x, y denote the points in T_i and T_j respectively. It easy to see that, the function $f(x, y) = 1$ is a density function of probability p defined above. Then, we follow Harsányi (p. 174, equations (5.2), (5.3)):

$$f(y \mid x) = \frac{f(x, y)}{\int_{[0,1]} f(x, y)\, dx}.$$

By slight calculation we get the density function of the conditional probability

$$f(y \mid x) = 1.$$

It seems contradicting our result above. However, the distribution function is also defined only up to (p) measure zero sets, so the function

$$g(y \mid x) = \begin{cases} f(y \mid x), & \text{if } x \neq t_i^* \\ 1, & \text{if } x = t_i^* \text{ and } y = t_j^* \\ 0, & \text{otherwise} \end{cases}$$

is also a density function of the conditional probability (it is the density function of a different variant of the conditional probability). Moreover, the two functions $f(y \mid x) = g(y \mid x)$ *p-a.s.*, $f(y \mid t_i^*) \neq g(y \mid t_i^*)$ *p-a.s.*, so they are different at t_i^*.

To sum up, we have got two different beliefs for player i at type t_i^*, and both beliefs are consistent with the common prior p. In other worlds, the players' beliefs cannot be determined by the common prior, some additional information is needed.

Harsányi wrote (p. 174): "The only difference is that in G the probabilities used by each player are subjective probabilities whereas on G^* these probabilities are objective (conditional) probabilities. But by the Bayesian hypothesis this difference is immaterial."

Unfortunately, our result above suggests that *the difference is material.*

In addition to, according to Harsányi's "definition" of type, it contains all information we need to determine the beliefs of any player. However, since the players' beliefs cannot be deduced from the common prior unambiguously, the model which based on the beliefs is not equivalent to the game based on the common prior. We think that this feature of conditional probability (non-uniqueness) does not go with Harsányi's intuition.

2.2 The second example

This example, which is based on Halmos' example [1], demonstrates that it is possible that the common prior not only do not determine the players' beliefs unambiguously, but there are no beliefs of the players which consistent with the common prior. In other words, we define such a probability space, which is a product space (the product of the two type sets of the players T_i, T_j), and there is no regular conditional probability (belief) on the measurable structure generated by pr_i (information available for player i).

Denote i and j the two players, and let $T_i = T_i = [0,1]$ be the sets of types of the players i and j respectively, and let $T = T_i \times T_j$. Let $\mathcal{M}_i = B([0,1])$, $\mathcal{M}_j = \sigma(B([0,1]) \cup \{M\})$, where M is a thick set[1] w.r.t. λ the Lebesgue measure on $B([0,1])$. Let $\mathcal{M} = \mathcal{M}_i \otimes \mathcal{M}_j$, $F : [0,1] \to [0,1] \times [0,1]$ be such a function that $F(x) = (x,x)$ $\forall x \in [0,1]$.

Lemma 4. *$\forall A \in \sigma(B([0,1]) \cup \{M\})$ $\exists B_1, B_2 \in B([0,1])$ such that*

$$A = (B_1 \cap M) \cup (B_2 \cap \complement M).$$

Proof. Left for the reader. Q.E.D.

Lemma 5. *$\sigma(B([0,1]) \cup \{M\}) = F^{-1}(\mathcal{M})$.*

Proof. It is clear that $diag([0,1] \times [0,1]) \in B([0,1] \times [0,1]) \subseteq \mathcal{M}$. Then $F^{-1}(\mathcal{M}) = F^{-1}(\mathcal{M} \cap diag([0,1] \times [0,1]))$, hence it is enough to concentrate on the measurable structure of the diagonal. $\mathcal{M} \cap diag([0,1] \times [0,1]) = diag([0,1] \times [0,1]) \cap pr_j^{-1}(\mathcal{M}_j)$, so $F^{-1}(\mathcal{M}) = \sigma(B([0,1]) \cup \{M\})$. Q.E.D.

Definition 6. *Let μ be such a probability measure on $\sigma(B([0,1]) \cup \{M\})$ that $\mu(A) = \lambda(B_1)$ $\forall A \in \sigma(B([0,1]) \cup \{M\})$, where $B_1 \in B([0,1])$ from Lemma 4. Let $p \circeq \mu \circ F^{-1}$ then*

$$(T, \mathcal{M}, p)$$

is a probability measure space.

[1]The thick set M is such a set that $\lambda^*(M) = 1$ and $\lambda_*(M) = 0$, where λ^* and λ_* are the outer and inner measure respectively, generated by λ.

Remark 7. It is easy to see that in this example $\lambda(B) = p(pr_i^{-1}(B))\ \forall B \in \mathcal{M}$, so p_i is the Lebesgue measure on $B([0,1])$ in Defintion 2.

Proposition 8. *The conditional probability $f_i(\cdot \mid \mathcal{M}_i)$ (see Definition 2.) is not regular, i.e. there is no $f_i : T_i \times \mathcal{M} \to [0,1]$ function such that*

1. $f_i(\cdot, A) = f_i(A \mid \mathcal{M}_i)$ *p_i-a.s.* $\forall A \in \mathcal{M}$,
2. $f_i(x, \cdot)$ *probability measure on* $\mathcal{M}$ $\forall x \in T_i$.

We split the proof into lemmata.

Assume indirectly that, there is such a function f_i that, it meets points *1.* and *2.* in Proposition 8.

Lemma 9. *Let $N = pr_j^{-1}(M) \cap diag([0,1] \times [0,1])$, and let $D = \{x \mid f_i(x, N) \neq 1\}$, then $\lambda(D) = 0$.*

Proof. Since f_i is a variant of the conditional probability on $\mathcal{M}_i$, $f(\cdot, N)$ is $\mathcal{M}_i$-measurable, so $D \in B([0,1]) = \mathcal{M}_i$.

Since $D \in B([0,1])$, $p(pr_i^{-1}(D) \cap N) = p(pr_i^{-1}(D) \cap diag([0,1] \times [0,1])) = p(pr_i^{-1}(D)) = p_i(D)$. Then

$$\int_{pr_i^{-1}(D) \cap N} 1\ dp = p(pr_i^{-1}(D)) = p_i(D) = \int_D 1\ dp_i.$$

From the definition of conditional probability (Definition 2.)

$$p(pr_i^{-1}(D) \cap N) = \int_D f(\cdot, N)\ dp_i,$$

so

$$\int_D 1\ dp_i = \int_D f(\cdot, N)\ dp_i.$$

Since $f(x, N) \leq 1$

$$p_i(\{x \in D \mid f(x, N) \neq 1\}) = 0.$$

However, $f(x, N) = 1\ \forall x \notin D$, so

$$p_i(\{x \in T_i \mid f(x, M) \neq 1\}) = 0,$$

therefore, since p_i is the Lebesgue measure $\lambda(D) = 0$. Q.E.D.

Lemma 10. *Let $\mathcal{R} = \{[a,b) \mid a, b \in \mathbb{Q}_{[0,1]},\ a \geq b\}$, and $E = \{x \mid \exists A \in \mathcal{R},\ f(x, pr_i^{-1}(A)) \neq 1_A\}$, then $\lambda(E) = 0$.*

Proof. Let $A \in \mathcal{R}$ be arbitrary fixed, and let $E_A = \{x \mid f(x, pr_i^{-1}(A)) \neq 1_A\}$. Since $f(\cdot, pr_i^{-1}(A))$ is $\mathcal{M}_i$-measurable, so $E_A \in B([0,1])$.

From the definition of conditional probability (Definition 2.)

$$\begin{aligned}
&\textstyle\int_{T_i} f(\cdot, pr_i^{-1}(A))\ dp_i = p(pr_i^{-1}(A) \cap T) = p(pr_i^{-1}(A) \cap pr_i^{-1}(A)) \\
&= \textstyle\int_A f(\cdot, pr_i^{-1}(A))\ dp_i = p(pr_i^{-1}(A)) = p_i(A) \\
&= \textstyle\int_A 1_A\ dp_i = \int_{T_i} 1_A\ dp.
\end{aligned}$$

Therefore, since $0 \leq f(x, pr_i^{-1}(A)) \leq 1$ so $f(x, pr_i^{-1}(A)) = 1_A$ p_i-*a.s.*, $p_i(E_A) = 0$ and $\lambda(E_A) = 0$.

We know that the cardinality of $\mathcal{R}$ is countable, λ is σ-additive and $E = \cup_{A \in \mathcal{R}} E_A$, so $\lambda(E) = 0$ Q.E.D.

Lemma 11. $f(x, pr_i^{-1}(\{x\})) = 1 \; \forall x \in \complement E$.

Proof. Let $x \in \complement E$ be arbitrary fixed, and let $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{R}$ be such a sequence of sets that $A_n \supseteq A_{n+1} \; \forall n$ and $\cap_n A_n = \{x\}$ $(\cap_n pr_i^{-1}(A_n) = pr_i^{-1}(\{x\}))$. Then, since we assumed that $f(x, \cdot)$ is a probability measure, and from Lemma 10. $f(x, pr_i^{-1}(A_n)) = 1 \; \forall n$, so $f(x, pr_i^{-1}(\{x\})) = 1$. Q.E.D.

Lemma 12. $pr_i^{-1}(\complement D \cap \complement E) \cap diag([0,1] \times [0,1]) \subseteq N$.

Proof. Let $x \in \complement D \cap \complement E$ be arbitrary fixed. Since $f(x, N) = 1$, $f(x, pr_i^{-1}(\{x\})) = 1$, $p(diag([0,1] \times [0,1])) = 1$ and $f(x, \cdot)$ is a probability measure

$$\begin{aligned} & f(x, N \cap pr_i^{-1}(\{x\}) \cap diag([0,1] \times [0,1])) = 1 \\ & \Rightarrow N \cap pr_i^{-1}(\{x\}) \cap diag([0,1] \times [0,1]) \neq \emptyset \\ & \Rightarrow pr_i^{-1}(\{x\}) \cap diag([0,1] \times [0,1]) \in N. \end{aligned}$$

Q.E.D.

The proof of proposition 8. Since lemmata 9., 10., 11. and 12. $p_i(\complement D \cap \complement E) = 1$ and $pr_i^{-1}(\complement D \cap \complement E) \cap diag([0,1] \times [0,1]) \subseteq N$. However, it is a contradiction since M is a thick set w.r.t. λ on $B([0,1])$, and $\complement D \cap \complement E \in B([0,1])$. Q.E.D.

To sum up, in this example there is no regular conditional probability of player i, so her beliefs cannot fit to the common prior p.

Remark 13. Halmos' construction differs from ours a little bit. He suggests the reader to prove that the set

$$\{x \mid \exists A \in B([0,1]),\; f(x, pr_i^{-1}(A)) \neq 1_A\}$$

is a Lebesgue measure zero set. It is too strong for the construction, so in Lemma 10. and 11. we proved less. It also easy to prove Halmos' original suggestion, it goes as the proof of Lemma 11., i.e. it is enough to prove that $\forall x \in \complement E$ $f(x, pr_i^{-1}(A)) = 1_A$ for any A from the monotone class generated by $\mathcal{R}$ $(B([0,1]))$.

3 The definition of type space

In this section we introduce the concept of type space formally. We use the formalism of Heifetz and Samet, and we show that the underpinning idea of their definition is the transition probability.

At the beginning of this section we follow Heifetz and Samet [3].

Definition 14. *Let $(X, \mathcal{M})$ be an arbitrary measurable space. Let the measurable structure of $(\Delta(X, \mathcal{M}), \mathcal{A}_{HS})$ be the σ-field generated by the sets $O = \{\mu \in \Delta(X, \mathcal{M}) \mid \mu(A) \geq \alpha\}$, where $\Delta(\cdot)$ is for the set of probability measures on the given measurable space, $A \in \mathcal{M}$, and $\alpha \in [0,1]$ is arbitrary fixed.*

Definition 15 (Heifetz and Samet)**.** *The type space* $< (T_i, \mathcal{M}_i)_{i \in M \cup \{0\}}, m_{i \in M} >$ *(briefly* $< (T, \mathcal{M}),\ m) >$*) based on parameter space* S *is as follows (where* M *is the set of players, and* 0 *denotes the extra player):*

1. $T_0 = S$, $(T_i, \mathcal{M}_i)$ *is a measurable space* $\forall i \in M \cup \{0\}$,

2. $m_i : T_i \to (\Delta(T, \mathcal{M}), \mathcal{A}_{HS})$ *is a measurable function* $\forall i \in M$,

3. $m_i(t_i)|_{\Delta(T_i, \mathcal{M}_i)} = \delta_{t_i}$, *where* δ_{t_i} *is the Dirac measure concentrated on* t_i, $\forall t_i \in T_i$.

Thus, the type space consists of the types of the players and functions which make the connection between types and the beliefs of any given player. Next, let us see the definition of a transition probability:

Definition 16. *Let* $(X, \mathcal{M})$ *be an arbitrary fixed measurable space, and let* $f : X \times \mathcal{M} \to ([0,1], B([0,1]))$ *be a function, where* $B(\cdot)$ *is for the Borel sets. If*

- $\forall x \in X\ f(x, \cdot)$ *is a probability measure on* $(X, \mathcal{M})$,
- $\forall A \in \mathcal{M}\ f(\cdot, A)$ *is a measurable function,*

then we call function f *a transition probability.*

The transition probability has all the properties and more, of regular conditional probability. It is easy to see that the measure zero events do not cause any problems for transition probability.

Definition 17. *Let us introduce some notations:*

1. M *is the set of active players excluding* 0 *(nature),*

2. 0 *is an extra player,*

3. $(T_i, \mathcal{M}_i)$ *is a measurable space* $\forall i \in M \cup \{0\}$,

4. $(T, \mathcal{M}) = (\Pi_{i \in M \cup \{0\}} T_i, \otimes_{i \in M \cup \{0\}} \mathcal{M}_i)$,

5. $f_i : T_i \times \mathcal{M} \to [0,1]$ *is a transition probability* $\forall i \in M$.

It is clear that, T_i's are the type sets, and the transition probability capture the way how the players form their beliefs.

Definition 18. *Let* $m_i = (\Pi_{A \in \mathcal{M}} f_i(\cdot, A))|_{diag(T_i^{\mathcal{M}})}$, *so* $m_i : diag(T_i^{\mathcal{M}}) \to \Delta(T, \mathcal{M})$, *in other words* $m_i : T_i \to \Delta(T, \mathcal{M})\ \forall i \in M$.

Functions m_i determines the beliefs belong to the given types.

Example 19. Let $\mathcal{M} = \{\emptyset, T\}$, and let $T_i = \{t_1, t_2\}$. Then $\Pi_{A \in \mathcal{M}} f_i(\cdot, A) = f_i(\cdot, \emptyset) \times f_i(\cdot, T)$, so e.g. $m_i(t_1) = \{\mu(\emptyset)\} \times \{\mu(T)\}$, where μ is a probability measure on $\mathcal{M}$.

Lemma 20. $\forall i \in M\ m_i$ *is* $B([0,1])^{\mathcal{M}}$*-measurable.*

Proof. It is clear that the natural embedding $b : diag(T_i^{\mathcal{M}}) \to T_i^{\mathcal{M}}$ is a measurable function. $f_i(\cdot, A) \to [0,1]$ is a measurable function $\forall A \in \mathcal{M}$, so $\Pi_{A \in \mathcal{M}} f_i(\cdot, A)$ is also measurable function w.r.t $B([0,1])^{\mathcal{M}}$. We know that $m_i = (\Pi_{A \in \mathcal{M}} f_i(\cdot, A)) \circ b$, so m_i is a measurable function. Q.E.D.

Theorem 21. *Heifetz and Samet's measurable structure* $\mathcal{A}_{HS}$ *(definition 14.) coincide with the measurable structure* $B([0,1])^{\mathcal{M}}$ *on* $(T, \mathcal{M})$*.*

Proof. Let $A \in \mathcal{M}$ be arbitrary, fixed. If $A = \emptyset$, or $A = X$, then the proof is finished. In the following, we assume that A does not coincide with both of the sets above.

Let $O = \{\mu \in \Delta(T, \mathcal{M}) \mid \mu(A) \geq \alpha\}$, then $\complement O = \{\mu \in \Delta(T, \mathcal{M}) \mid \mu(A) < \alpha\}$. We know that $\exists \nu \in \Delta(T, \mathcal{M})$ such that $\nu(A) = 0$, then $U(\nu, A) = \{\mu \in \Delta(T, \mathcal{M}) \mid |\nu(A) - \mu(A)| < \alpha\}$, $U(\nu, A) = \complement O$, or $O = \complement U(\nu, A)$.

Let $U(\nu, A) = \{\mu \in \Delta(T, \mathcal{M}) \mid |\nu(A) - \mu(A)| \geq \alpha\}$ for arbitrary ν, α. Let $p_1 = min\{\nu(A) + \alpha, 1\}$, and let $p_2 = max\{\nu(A) - \alpha, 0\}$. Denote $O_1 = \{\mu \in \Delta(T, \mathcal{M}) \mid \mu(\complement A) \geq 1 - p_2\}$. It is easy to see that $O_1 = \{\mu \in \Delta(T, \mathcal{M}) \mid \mu(A) < p_2\}$. Let a_n be such a strictly decreasing sequence that $a_n \in [0,1]$ $\forall n$, and $a_n \to p_1$, and let $O_2 = \cup_n \{\mu \in \Delta(T, \mathcal{M}) \mid \mu(A) > a_n\}$. Then $O_2 = \{\mu \in \Delta(T, \mathcal{M}) \mid \mu(A) > p_1\}$, so $U(\nu, A) = \complement(O_1 \cup O_2)$.

We proved that the set systems which generate the two measurable structures coincide. Q.E.D.

Remark 22. In the proof of theorem 21. we used intensively that the range of the measures is the unit interval of $\mathbb{R}$.

In the discussion above the starting points were the sets of types (T_i) and the transition probabilities (f_i). We take these sets and functions as given, we do not care about the existence of them.

4 Conclusions

To sum up, since the conditional probability does not determine the players' beliefs unambiguously, hence for determining the players' beliefs we need extra information. This extra information can be formalized by transition probabilities.

Finally, we suggest a slight modification of Heifetz and Samet's definition (definition 15.) as follows:

Definition 23. *The type space* $< (T_i, \mathcal{M}_i)_{i \in M \cup \{0\}}, f_{i \in M} >$ *(briefly* $< (T, \mathcal{M}), f) >$*) based on parameter space* S *is as follows:*

1. $T_0 = S$, $(T_i, \mathcal{M}_i)$ *is a measurable space* $\forall i \in M \cup \{0\}$,
2. $f_i : T_i \times \mathcal{M} \to [0,1]$ *is a transition probability function* $\forall i \in M$,
3. $\text{supp} f_i(t_i, \cdot) = pr_i^{-1}(\{t_i\})$ $\forall t_i \in T_i$ $\forall i \in M$.

The above definition emphasizes the basic method of the reasoning of the players more, and is not more complicated any way.

It is important to see that, we must not use the concept of common prior to make a complete information game from an incomplete information game. Even if we can define the transition probabilities as variants of the conditional probabilities of the common prior in the above definition, probability distribution on the product of the type sets does not contain all information that transition probabilities do. Therefore, in general, we cannot summarize the reasoning of the players into a model in which there are only the type sets and the common prior. In other words, from cognitive viewpoint the model with common prior and the model with transition probabilities (which can fit to the common prior) are not equivalent.

References

[1] Halmos P.R., *Measure Theory*, Springer-Verlag 1974

[2] Harsányi J., *Games with incomplete information played by Bayesian players part I., II., III.*, Management Science **14**, 159–182., 320–334., 486–502., 1967-68

[3] Heifetz A., and D. Samet, *Topology-free typology of beliefs*, Journal of Economic Theory **82**, 324–341., 1998