Capítulo 1

Juegos estáticos y juegos estocásticos

La teoría de juegos estudia modelos matemáticos de situaciones de cooperación o de conflicto en el que participan dos o más entidades (personas, empresas, países, etc.) las cuales eligen acciones o toman decisiones que producen alguna "ganancia"a cada uno de los participantes. En este contexto, la palabra "juego"se refiere a la situación de conflicto (económico, social, natural, etc.) y a los participantes se les llama jugadores.

Este trabajo está orientado a los modelos no-cooperativos, es decir, modelos para los cuales los jugadores actúan independientemente y cada uno desea alcanzar su propio objetivo. Dicho objetivo consiste en maximizar sus ganancias tomando decisiones de acuerdo a determinadas reglas llamadas estrategias.

En la teoría de juegos la ganancia que recibe cada jugador depende tanto de las decisiones que toman los otros jugadores como de la propia, en consecuencia un jugador no necesariamente puede alcanzar su ganancia máxima posible. Por tal motivo, se introduce el concepto de equilibrio de Nash, el cual se interpreta como una situación en la que para ninguno de los jugadores es favorable cambiar de estrategia ya que cualquier cambio implica una disminución en sus ganancias.

1.1. Juegos estáticos no-cooperativos

Un juego estático no-cooperativo de N jugadores es un sistema

$$\Gamma = (I, \{A_i\}_{i\in I}, \{r_i\}_{i\in I}) \tag{1.1}$$

que consiste en el conjunto de jugadores $I = \{1, ..., N\}$, el conjunto de acciones o estrategias puras A_i del jugador $i \in I$ y la función de ganancia o pago $r_i : A \to \mathbb{R}$ para el jugador i, donde

$A := A_1 \times \cdots \times A_N$. Un juego es finito, si el conjunto de acciones de cada jugador es un conjunto finito.

En un juego estático no-cooperativo cada jugador escoge una acción de su conjunto de acciones y entonces el jugador $i \in I$, recibe una ganancia $r_i(a_1, ..., a_N)$, donde $a_j \in A_j$ es la acción elegida por el jugador j.

Note que la ganancia $r_i(a_1, ..., a_N)$ puede tomar valores negativos lo cual se interpreta como una pérdida o costo para el jugador. Por lo tanto, el objetivo de cada jugador es elegir la mejor acción o estrategia pura para maximizar su ganancia. Un punto importante que el jugador debe considerar es que su ganancia depende de todas las acciones elegidas por cada uno de los jugadores, por tal motivo, se define a continuación el concepto de equilibrio el cual fue introducido por J. F. Nash en [15].

Definición 1.1.1. *Un punto* $(a_1^*, ..., a_N^*) \in A$ *es un equilibrio para el juego* Γ *si*

$$r_i(a_1^*, ..., a_N^*) = \max_{a_i \in A_i} r_i(a_1^*, ..., a_i, ..., a_N^*) \quad \forall i \in I.$$

Nótese que si $(a_1^*, ..., a_N^*)$ es un punto de equilibrio y el jugador j elige la acción a_j en lugar de elegir la acción a_j^*, entonces

$$r_j(a_1^*, ..., a_N^*) \geq r_j(a_1^*, ..., a_j, ..., a_N^*).$$

Es decir, ninguno de los jugadores está motivado para cambiar su decisión en la estrategia de equilibrio ya que cualquier cambio no aumenta su ganancia.

Ejemplo 1.1.2. *Consideremos un juego de dos jugadores con* $A_1 = \{a_1, a_2\}, A_2 = \{b_1, b_2\}$ *y las correspondientes funciones de pago dadas por*

$$\begin{array}{ll} r_1(a_1, b_1) = 4 & r_1(a_1, b_2) = -1 \\ r_1(a_2, b_1) = 0 & r_1(a_2, b_2) = 1, \\ r_2(a_1, b_1) = 1 & r_2(a_1, b_2) = -1 \\ r_2(a_2, b_1) = 0 & r_2(a_2, b_2) = 4. \end{array}$$

El juego tiene los equilibrios (a_1, b_1) *y* (a_2, b_2) *en* $A_1 \times A_2$.

Ejemplo 1.1.3. *Para el juego de dos jugadores con* $A_1 = \{a_1, a_2\}, A_2 = \{b_1, b_2\}$ *y funciones de pago*

$$\begin{array}{ll} r_1(a_1, b_1) = 3 & r_1(a_1, b_2) = 1 \\ r_1(a_2, b_1) = 4 & r_1(a_2, b_2) = 0, \\ r_2(a_1, b_1) = 5 & r_2(a_1, b_2) = 0 \\ r_2(a_2, b_1) = 0 & r_2(a_2, b_2) = 5, \end{array}$$

el juego no tiene equilibrios en $A_1 \times A_2$.

El Ejemplo 1.1.3 muestra que en un juego no siempre existen equilibrios en el conjunto de las estrategias puras. Sin embargo, bajo ciertas condiciones, sí existen equilibrios en el conjunto de las estrategias aleatorizadas o estrategias mixtas. Esta clase de estrategias se define a continuación.

Una *estrategia aleatorizada* para el jugador i es una medida de probabilidad en el conjunto de acciones A_i. Denotaremos por $\mathbb{P}(A_i)$ al conjunto de todas las estrategias aleatorizadas para el jugador i y a cada elemento en $\mathbb{P}(A) := \mathbb{P}(A_1) \times \cdots \times \mathbb{P}(A_N)$ lo llamaremos *multi-estrategia* para los N jugadores. Cuando los jugadores usan una multi-estrategia $\mu = (\mu_1, ..., \mu_N) \in \mathbb{P}(A)$ la elección del jugador i puede interpretarse como el resultado de una variable aleatoria con distribución de probabilidad μ_i.

Para cada $\phi_i \in \mathbb{P}(A_i)$ y $\mu = (\mu_1, ..., \mu_N) \in \mathbb{P}(A)$ denótese

$$(\mu^{-i}, \phi_i) := (\mu_1, ..., \mu_{i-1}, \phi_i, \mu_{i+1}, ..., \mu_N).$$

Además, se denotará por

$$\bar{r}_i(\mu) \quad := \quad \int_{A_N} \cdots \int_{A_1} r_i(a_1, ..., a_N)\mu_1(da_1)\cdots\mu_N(da_N) \quad \mu \in \mathbb{P}(A) \tag{1.2}$$

a la ganancia esperada del jugador $i \in I$ cuando los jugadores usan la multi-estrategia $\mu \in \mathbb{P}(A)$.

Para el caso particular donde $\mu = (\mu_1, ..., \delta_i, ..., \mu_N)$ con δ_i la medida de Dirac concentrada en el punto a_i, es decir,

$$\delta_i(\tilde{a}) = \begin{cases} 1 & \text{si } \tilde{a} = a_i \\ 0 & \text{si } \tilde{a} \neq a_i \end{cases}$$

denotaremos por $\bar{r}_i(\mu^{-i}, a_i)$ a la función $\bar{r}_i(\mu^{-i}, \delta_i)$.

Definición 1.1.4. *(a) Diremos que $\phi_i \in \mathbb{P}(A_i)$ es una respuesta óptima del jugador $i \in I$ para la multi-estrategia μ si*

$$\bar{r}_i(\mu^{-i}, \phi_i) \quad = \quad \max_{\tau_i \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau_i). \tag{1.3}$$

(b)Una multi-estrategia $\mu \in \mathbb{P}(A)$ es un equilibrio de Nash para el juego Γ si

$$\bar{r}_i(\mu) \quad = \quad \max_{\phi_i \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \phi_i) \quad \forall i \in I.$$

Denotaremos por $\bar{R}_i(\mu)$ al conjunto de todas las respuestas óptimas del jugador i para μ.

Es importante señalar que, en general el máximo en (1.3) no se alcanza y en consecuencia el equilibrio de Nash no siempre existe. Por lo tanto, para garantizar la existencia de equilibrios es necesario imponer condiciones para el conjunto A_i y las funciones de pago r_i $(i \in I)$.

1.1.1. Juegos no-cooperativos finitos

Considérese Γ como en (1.1) un juego estático finito , es decir, para cada $j \in I$,

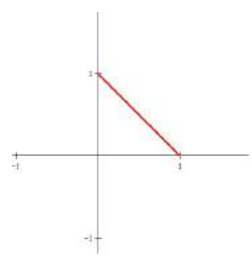
$$n_j \quad := \quad \#|A_j| < \infty.$$

Observemos que si $\mu = (\mu_1, ..., \mu_N) \in \mathbb{P}(A)$ entonces para cada $j \in I$ y cada $a_l \in A_j$, el número $\mu_j(a_l)$ es un número real no negativo menor igual que uno que representa la probabilidad de que el jugador j elija la acción a_l y tal que para cada $j \in I$,

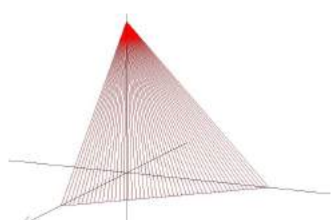
$$\sum_{l=1}^{n_j} \mu_j(a_l) = 1. \tag{1.4}$$

El espacio $\mathbb{P}(A_i)$ tiene una interpretación geométrica que varía de acuerdo al número de elementos (acciones) que tiene el conjunto A_i:

a) Cuando el conjunto de acciones es $A_i = \{a_1, a_2\}$, el conjunto $\mathbb{P}(A_i)$ puede identificarse con el subconjunto de $\mathbb{R}^2, \{(\lambda_1, \lambda_2) : \lambda_1, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1\}$



b) Cuando el conjunto de acciones es $A_i = \{a_1, a_2, a_3\}$, el conjunto $\mathbb{P}(A_i)$ puede identificarse con el subconjunto de $\mathbb{R}^3, \{(\lambda_1, \lambda_2, \lambda_3) : \lambda_1, \lambda_2, \lambda_3 \geq 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}$



En general, si $A_i = \{a_1, ..., a_{n_i}\}$, entonces el conjunto $\mathbb{P}(A_i)$ puede identificarse con $\left\{(\lambda_1, ..., \lambda_{n_i}) : \lambda_j \geq 0, \sum_{j=1}^{n_i} \lambda_j = 1\right\} \subset \mathbb{R}^{n_i}$ (simplejo de dimensión $n_i - 1$) y el espacio $\mathbb{P}(A)$ puede identificarse con un subconjunto de $\mathbb{R}^m$ donde $m = \sum_{i=1}^{N} n_i$. Para lo que sigue, dotaremos al conjunto $\mathbb{P}(A)$ con la métrica de $\mathbb{R}^m$.

Nótese que para una multi-estrategia $\mu = (\mu_1, ..., \mu_N)$, la función $\bar{r}_i$ definida en (1.2) está dada por

$$\begin{aligned}\bar{r}_i(\mu) &= \sum_{l_N=1}^{n_N} \cdots \sum_{l_1=1}^{n_1} r_i(a_{l_1}, ..., a_{l_N})\mu_1(a_{l_1})\cdots\mu_N(a_{l_N}) \\ &= \sum_{a\in A}\prod_{j\in I} r_i(a)\mu_j(a_{l_j}).\end{aligned}$$

Así tenemos que

$$\begin{aligned}\bar{r}_i(\mu^{-i}, \phi_i) &= \sum_{a\in A}[\prod_{j\in I\setminus\{i\}} r_i(a)\mu_j(a_{l_j})]\phi_i(a_{l_i}), \\ \bar{r}_i(\mu^{-i}, a_i) &= \sum_{a_j\in A_j, j\neq i}[\prod_{j\in I\setminus\{i\}} r_i(a)\mu_j(a_{l_j})].\end{aligned}$$

1.1.2. Equilibrios de Nash para juegos no-cooperativos finitos

En esta sección se demostrará que si Γ es un juego no-cooperativo finito de N jugadores entonces existe al menos una multi-estrategia en $\mathbb{P}(A)$ que es un equilibrio de Nash para el juego Γ.

Observación 1.1.5. *En general, la estructura del espacio $\mathbb{P}(A)$ es muy importante para la existencia de equilibrios. En particular cuando los conjuntos de acciones son finitos tenemos las siguientes propiedades:*

1) *Para cada $\beta \in [0,1]$ y $\mu_1, \mu_2 \in \mathbb{P}(A)$ se tiene que*

$$\beta\mu_1 + (1-\beta)\mu_2 \in \mathbb{P}(A),$$

es decir, $\mathbb{P}(A)$ es un conjunto convexo.

2) *El conjunto $\mathbb{P}(A)$ es un conjunto acotado en $\mathbb{R}^m$ puesto que para cada $\mu = (\mu_1, ..., \mu_N) \in \mathbb{P}(A)$,*

$$\sum_{i=1}^{N}\mu_i(a_{l_i}) \leq \sum_{i=1}^{N} 1 = N \qquad \forall (a_{l_1}, ..., a_{l_N}) \in A.$$

3) *El espacio $\mathbb{P}(A)$ es cerrado y en consecuencia compacto.*

4) *La función $\bar{r}_i(\cdot)$ es continua en $\mathbb{P}(A)$.*

El siguiente teorema, sin duda alguna, marcó la historia de la teoría de juegos no-cooperativos. La demostración original de John F. Nash se publicó en [15] y la demostración que se presenta a continuación se tomó de [21].

Teorema 1.1.6. *Si* Γ *es un juego no-cooperativo finito de* N *jugadores entonces existe un equilibrio de Nash en* $\mathbb{P}(A)$.

Demostración. Sea $\Psi : \mathbb{P}(A) \twoheadrightarrow \mathbb{P}(A)$ la multifunción definida por

$$\Psi(\mu) := \prod_{i=1}^{N} \bar{R}_i(\mu).$$

Por la Observación 1.1.5, 3) y 4), el conjunto $\mathbb{P}(A_i)$ es compacto y $\bar{r}_i(\mu^{-i}, \cdot)$ es continua en $\mathbb{P}(A_i)$. En consecuencia, se tiene que $\bar{R}_i(\mu)$ es un conjunto no vacío. Por lo tanto,

$$\Psi(\mu) \neq \varnothing \quad \forall \mu \in \mathbb{P}(A).$$

Ahora se probará que $\Psi(\mu)$ es convexo. Sean $\phi_1, \phi_2 \in \bar{R}_i(\mu)$ y $\beta \in [0,1]$ y observe que

$$\begin{aligned} \bar{r}_i(\mu^{-i}, \beta\phi_1 + (1-\beta)\phi_2) &= \beta\bar{r}_i(\mu^{-i}, \phi_1) + (1-\beta)\bar{r}_i(\mu^{-i}, \phi_2) \\ &= \beta \max_{\tau_i \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau_i) + (1-\beta) \max_{\tau_i \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau_i) \\ &= \max_{\tau_i \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau_i); \end{aligned}$$

por lo tanto, $\beta\phi_1 + (1-\beta)\phi_2 \in \bar{R}_i(\mu)$.

Se demostrará ahora que la multifunción Ψ es semicontinua superiormente. Sean $(\mu_n)_{n=1}^{\infty}$ una sucesión convergente a μ en $\mathbb{P}(A)$ y $\phi_n \in \Psi(\mu_n)$ una sucesión convergente a ϕ. Para probar la semicontinuidad superior de Ψ demostraremos que $\phi \in \Psi(\mu)$.

Obsérvese que para cada $n \in \mathbb{N}, i \in I$ y para cada $\tau_i \in \mathbb{P}(A_i)$,

$$\bar{r}_i((\mu_n)^{-i}, \phi_n^i) \geq \bar{r}_i((\mu_n)^{-i}, \tau_i).$$

Por la continuidad de $\bar{r}_i$ sobre $\mathbb{P}(A)$, tenemos que

$$\bar{r}_i(\mu^{-i}, \phi^i) \geq \bar{r}_i(\mu^{-i}, \tau_i) \quad \forall \tau_i \in \mathbb{P}(A_i);$$

entonces, $\phi^i \in \bar{R}_i(\mu)$ para cada $i \in I$; por lo tanto, $\phi \in \Psi(\mu)$. Entonces, la multifunción $\Psi(\cdot)$ satisface las condiciones del teorema de punto fijo de Kakutani (Teorema 4.0.1) y por lo tanto, existe una multi-estrategia $\mu \in \mathbb{P}(A)$ tal que $\mu \in \Psi(\mu)$, es decir,

$$\bar{r}_i(\mu) = \max_{\tau_i \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau_i) \quad \forall i \in I;$$

equivalentemente, la multi-estrategia μ es un equilibrio de Nash para el juego Γ. ∎

J. Nash generalizó el concepto de equilibrio estudiado ampliamente por John Von Neumann y Oskar Morgenstern en su libro *Theory of Games and Economic Behavior*. En un juego finito de suma cero de dos jugadores $G = \{A_1, A_2, r\}$ tenemos que

$$r(a) \quad = \quad r_1(a) = -r_2(a). \tag{1.5}$$

Para un juego de suma cero, si $(\mu_1^*, \mu_2^*) \in \mathbb{P}(A_1)\times\mathbb{P}(A_2)$ es un equilibrio de Nash, entonces

$$\begin{aligned} r(\mu_1^*, \mu_2^*) &\geq r(\mu_1, \mu_2^*) \quad \forall \mu_1 \in \mathbb{P}(A_1), \\ r(\mu_1^*, \mu_2^*) &\leq r(\mu_1^*, \mu_2) \quad \forall \mu_2 \in \mathbb{P}(A_2), \end{aligned}$$

de donde se sigue que

$$\max_{\mu_1 \in \mathbb{P}(A_1)} r(\mu_1, \mu_2^*) \leq r(\mu_1^*, \mu_2^*) \leq \min_{\mu_2 \in \mathbb{P}(A_2)} r(\mu_1^*, \mu_2),$$

$$\min_{\mu_2 \in \mathbb{P}(A_2)} \max_{\mu_1 \in \mathbb{P}(A_1)} r(\mu_1, \mu_2) \leq r(\mu_1^*, \mu_2^*) \leq \max_{\mu_1 \in \mathbb{P}(A_1)} \min_{\mu_2 \in \mathbb{P}(A_2)} r(\mu_1, \mu_2).$$

Por otro lado se tiene que

$$\max_{\mu_1 \in \mathbb{P}(A_1)} \min_{\mu_2 \in \mathbb{P}(A_2)} r(\mu_1, \mu_2) \quad \leq \quad \min_{\mu_2 \in \mathbb{P}(A_2)} \max_{\mu_1 \in \mathbb{P}(A_1)} r(\mu_1, \mu_2),$$

lo cual, combinando con las desigualdades anteriores; (μ_1^*, μ_2^*) es un equilibrio de Nash si, y sólo si

$$\begin{aligned} r(\mu_1^*, \mu_2^*) &= \min_{\mu_2 \in \mathbb{P}(A_2)} \max_{\mu_1 \in \mathbb{P}(A_1)} r(\mu_1, \mu_2) \\ &= \max_{\mu_1 \in \mathbb{P}(A_1)} \min_{\mu_2 \in \mathbb{P}(A_2)} r(\mu_1, \mu_2). \end{aligned}$$

1.1.3. Juegos no-cooperativos con espacio de acciones compactos

En 1951 L. Glicksberg en [10] extiende el resultado de existencia de equilibrios de Nash para juegos estáticos no-cooperativos de N jugadores no necesariamente finitos. Él consideró un juego donde el espacio de las estrategias puras de cada jugador es un espacio métrico compacto y la función de ganancia r_i es una función continua en A. Este resultado es presentado como una aplicación del teorema de punto fijo (ver Teorema 4.0.2).

En este caso, el espacio $\mathbb{P}(A_i)$ es un subconjunto compacto y convexo del espacio $\mathscr{M}(A_i)$ formado por las medidas con signo finitas en A_i cuando se considera la topología débil.

En la Proposición 4.0.6 se prueba que $\mathscr{M}(A_i)$ es un espacio vectorial topológico Hausdorff localmente convexo. Además, la convergencia en $\mathbb{P}(A_i)$ queda completamente caracterizada con la Proposición 7.21 en [4] p. 128, esto es, $\mu_n \to \mu$ en $\mathbb{P}(A_i)$ si, y sólo si,

$$\int_{A_i} f d\mu_n \to \int_{A_i} f d\mu \quad \forall f \in C(A_i).$$

La continuidad de la función ganancia r_i en A implica la continuidad de la función $\bar{r}_i$ en $\mathbb{P}(A)$. Esto es, si $(\mu_n)_{n=1}^{\infty}$ es una sucesión convergente a μ, entonces la convergencia débil implica que

$$\begin{aligned} \bar{r}_i(\mu_n) &= \int_{A_N} \cdots \int_{A_1} r_i(a_1, ..., a_N)\mu_n^1(da_1)\cdots\mu_n^N(da_N) \\ \downarrow & \\ \bar{r}_i(\mu) &= \int_{A_N} \cdots \int_{A_1} r_i(a_1, ..., a_N)\mu^1(da_1)\cdots\mu^N(da_N). \end{aligned}$$

Teorema 1.1.7. *Sea Γ un juego estático de N jugadores. Si para cada $i \in I$ el conjunto de las estrategias puras A_i es un espacio métrico compacto y la función ganancia r_i es continua en A, entonces el juego tiene un equilibrio de Nash en el conjunto de las multi-estrategias aleatorizadas.*

Demostración. Se define la multifunción $\Psi : \mathbb{P}(A) \twoheadrightarrow \mathbb{P}(A)$ por

$$\Psi(\mu) \ := \ \prod_{i=1}^{N} \bar{R}_i(\mu),$$

donde

$$\bar{R}_i(\mu) \ := \ \left\{ \phi \in \mathbb{P}(A_i) : \bar{r}_i(\mu^{-i}, \phi) = \max_{\tau \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau) \right\} \ i \in I.$$

La continuidad de la función $\bar{r}_i$ en el conjunto compacto $\mathbb{P}(A_i)$ asegura que el conjunto $\Psi(\mu)$ es no vacío. La linealidad de la función $\bar{r}_i$ implica que el conjunto $\Psi(\mu)$ es convexo.

Sea $(\mu_n)_{n=1}^{\infty}$ una sucesión convergente a μ en $\mathbb{P}(A)$ y $\phi_n \in \Psi(\mu_n)$ tal que la sucesión $(\phi_n)_{n=1}^{\infty}$ converge a un punto ϕ. Puesto que la función $\bar{r}_i$ es continua y como $\phi_n \in \Psi(\mu_n)$, para cada $\tau \in \mathbb{P}(A_i)$ tenemos que

$$\begin{aligned} \bar{r}_i((\mu_n)^{-i}, \phi_n^i) &\geq \bar{r}_i((\mu_n)^{-i}, \tau) \\ &\downarrow \\ \bar{r}_i(\mu^{-i}, \phi^i) &\geq \bar{r}_1(\mu^{-i}, \tau). \end{aligned}$$

Esto implica que $\phi^i \in \bar{R}_i(\mu)$ para todo $i \in I$, es decir, $\phi \in \Psi(\mu)$. Por lo tanto Ψ es semicontinua superiormente.

Finalmente, por el teorema de punto fijo de Glicksberg (ver Teorema 4.0.2), se sigue que existe un punto fijo μ de la multifunción Ψ, es decir, se satisface que

$$\bar{r}_i(\mu) = \max_{\tau \in \mathbb{P}(A_i)} \bar{r}_i(\mu^{-i}, \tau) \quad \forall i \in I.$$

Por lo tanto, μ es un equilibrio de Nash en $\mathbb{P}(A)$ para el juego Γ. ■

1.2. Juegos estocásticos no-cooperativos

En esta sección se definen los conceptos de juego estocástico, el criterio de pago descontado y el concepto de equilibrio de Nash para esta clase de juegos. En los capítulos 2 y 3 se establecen condiciones para la existencia de equilibrios.

Definición 1.2.1. *Un modelo de juego estocástico no-cooperativo de N jugadores, es un sistema de la forma:*

$$G := \{X, (A_i, \{A_i(x) : x \in X\}, r_i)_{i=1,\ldots,N}, Q\}$$

que consiste en:

(a) El conjunto X llamado, el espacio de estados.

(b) El conjunto $A_i, i \in I := \{1, \ldots, N\}$ es el espacio de acciones para el jugador i.

(c) El conjunto $A_i(x) \subset A_i$ $(i \in I)$ es el conjunto de acciones admisibles para el jugador i cuando el sistema está en el estado x. Definamos la multifunción de X en $A := A_1 \times \cdots \times A_N$ por

$$A(x) := \prod_{i=1}^{N} A_i(x),$$

y denótese su gráfica por,

$$\mathbb{K} := \{(x, a) : x \in X, a \in A(x)\}.$$

(d) La función $r_i : \mathbb{K} \to \mathbb{R}$ $(i \in I)$ es la función de pago o ganancia del jugador i.

(e) La ley de transición Q es un kernel estocástico en $\mathbb{P}(X|\mathbb{K})$.

Un modelo de juego estocástico representa un sistema que evoluciona en el tiempo de la siguiente manera:

1.- En el tiempo $t = 0$ el sistema se encuentra en un estado inicial x_0.

2.- Cada jugador escoge independientemente una acción de su conjunto de acciones admisibles determinando así, $a = (a_1, ..., a_N) \in A(x_0)$.

3.- El jugador i recibe un pago $r_i(x, a)$.

4.- En $t = 1$, el sistema pasa a un nuevo estado x_1, de acuerdo a la ley de transición $Q(\cdot/x_0, a)$ y el proceso se repite indefinidamente.

Para cada $t = 0, 1, 2, ...$, se define el conjunto de las $t-$historias como $\mathbb{H}_t = \mathbb{K}^t \times X$ y $\mathbb{H}_0 = X$. Obsérvese que un elemento de $\mathbb{H}_t$ es de la forma

$$h_t = (x_0, a_0, ..., x_{t-1}, a_{t-1}, x_t).$$

Una *política o estrategia* para el jugador i es una sucesión $\pi_i = \{\pi_t^i\}$ de kernels estocásticos $\pi_t^i \in \mathbb{P}(A_i|\mathbb{H}_t)$ tal que

$$\pi_t^i(A_i(x)/h_t) = 1 \quad \forall h_t \in \mathbb{H}_t.$$

Se denota por Π_i al conjunto de todas las políticas para el jugador i y a cada elemento $\pi = (\pi_1, ..., \pi_N)$ en $\Pi := \Pi_1 \times \cdots \times \Pi_N$ se le llama una *multi-estrategia* para los N jugadores. Denotaremos por $\mathbb{P}(A_i|X)$ a la familia de kernels estocásticos que satisfacen la condición $\varphi_i(A_i(x)|x) = 1$ para todo $x \in X$.

Definición 1.2.2. *Una estrategia $\pi_i = \{\pi_t^i\} \in \Pi_i$ es estacionaria si existe $\phi_i \in \mathbb{P}(A_i|X)$ tal que*

$$\pi_t^i(B/h_t) = \phi_i(B/x_t) \ \forall h_t \in \mathbb{H}_t, \ \ B \in \mathscr{B}(A_i), \ \ t \geq 0.$$

Sea Φ_i el conjunto de todas las estrategias estacionarias del jugador i e identifíquese la estrategia $\{\phi, \phi, ...\}$ con ϕ y $\Phi := \Phi_1 \times \cdots \times \Phi_N$ el conjunto de las *multi-estrategias estacionarias.*

Para cada $\pi \in \Pi$ y para cada $\gamma \in \Pi_i$, se usará la siguiente notación

$$(\pi^{-i}, \gamma_i) := (\pi_1, ..., \pi_{i-1}, \gamma_i, \pi_{i+1}, ..., \pi_N).$$

1.2.1. Criterio de pago descontado

Sea $((X \times A)^\infty, \sum)$ el espacio medible donde $\sum$ es la $\sigma-$álgebra producto de $(X \times A)^\infty$ y ν una medida de probabilidad en X. Entonces por el Teorema de Ionescu-Tulcea [ver Ash (1972, p. 109), Bertsekas and Shereve (1978, pp. 140-141)], para cada multi-estrategia $\pi \in \Pi$ existe una medida de probabilidad P_ν^π y un proceso estocástico $\{(x_t, a_t) : t = 0, 1, 2, ...\}$ en

$((X \times A)^\infty, \sum)$, donde x_t representa el estado y a_t representa el vector de acciones en el tiempo t tal que

$$\begin{aligned} P_\nu^\pi(x_0 \in B) &= \nu(B), \\ P_\nu^\pi(a_t \in C/h_t) &= \pi_t(C/h_t), \\ P_\nu^\pi(x_{t+1} \in B/h_t, a_t) &= Q(B/x_t, a_t). \end{aligned}$$

para todo $B \in \mathscr{B}(X), C \in \mathscr{B}(A), h_t \in \mathbb{H}_t$ y $t \geq 0$. En el caso particular donde ν es la medida de Dirac concentrada en un estado $x \in X$, la medida de probabilidad P_ν^π la denotaremos por P_x^π y el operador esperanza con respecto a P_x^π lo denotaremos por E_x^π.

Ahora se puede establecer claramente el objetivo del juego estocástico. La ganancia se acumula durante la evolución del juego considerando el criterio de ganancia total descontada definido de la siguiente manera.

Definición 1.2.3. *Para cada $i = 1, ..., N, x \in X$ y cada multi-estrategia $\pi \in \Pi$, definimos la ganancia total esperada α-descontada para el jugador i por*

$$V_i(x, \pi) := E_x^\pi \sum_{t=0}^{\infty} \alpha^t r_i(x_t, a_t)$$

donde $\alpha \in (0,1)$ se conoce como factor de descuento.

Como en el caso de juegos estáticos, en un juego estocástico la ganancia de cada jugador depende de las estrategias que elijan todos los jugadores lo que hace necesario introducir nuevamente un concepto de equilibrio.

Definición 1.2.4. *(a) Una estrategia $\pi_i^* \in \Pi_i$ es una respuesta óptima del jugador i para la multi-estrategia $\pi \in \Pi$ si*

$$V_i(x, (\pi^{-i}, \pi_i^*)) = \max_{\tau \in \Pi_i} V_i(x, (\pi^{-i}, \tau)) \qquad \forall x \in X.$$

(b) Una multi-estrategia $\pi^* = (\pi_1^*, ..., \pi_N^*) \in \Pi$ es un equilibrio de Nash para el juego si para cada $i = 1, 2, ..., N$, la estrategia π_i^* es una respuesta óptima del jugador i para π^*, es decir,

$$V_i(x, \pi^*) = \max_{\tau \in \Pi_i} V_i(x, ((\pi^*)^{-i}, \tau)) \qquad \forall x \in X, \ \ i = 1, ..., N.$$

Uno de los principales conceptos que se utilizan en este trabajo, además de los mencionados anteriormente, es el concepto de función $W-$acotada.

Definición 1.2.5. *Sean X un espacio métrico y $W : X \to [1, \infty)$ una función medible. Para cada función $g \in \mathbb{M}(X)$ definimos la W-norma por*

$$\|g\|_W \quad = \quad \sup_{x \in X} \frac{|g(x)|}{W(x)}.$$

Si $\|g\|_W < \infty$ diremos que g es W-*acotada*. Denótese por $\mathbb{B}_W(X)$ al conjunto de todas las funciones W−acotadas y note que es un espacio de Banach con la W−norma. Además, es claro que $\mathbb{B}(X) \subset \mathbb{B}_W(X)$.