

Índice

1. Problemas de decisión	1
1.1. Elementos esenciales	1
1.2. Ambientes en un problema de decisión	7
1.2.1. Ambiente de certidumbre: Técnicas de optimización	9
1.2.2. Ambiente de riesgo	9
1.2.3. Ambiente de incertidumbre	12
1.2.4. Decisión con experimentación	12
1.3. Decisiones aleatorizadas	15
1.3.1. Decisiones aleatorizadas sin experimentación	17
1.3.2. Decisiones aleatorizadas con experimentación	18
1.4. Utilidad y probabilidad subjetiva	21
1.4.1. Utilidad	21
1.4.2. Probabilidad subjetiva	26
1.4.3. Utilidad monetaria	29
2. Decisión en ambiente de riesgo e incertidumbre	37
2.1. Introducción	37
2.2. Criterios de decisión en ambiente de riesgo	37
2.2.1. Criterio del valor esperado:	38
2.2.2. Criterio Media-Dispersión:	38
2.2.3. Criterio de riesgo fijo:	40
2.2.4. Criterio de máxima probabilidad:	40
2.2.5. Criterio del valor esperado con cláusula de seguridad:	40
2.2.6. Aplicación: El problema de selección de la cartera	41
2.2.7. Ejemplo adicional	43
2.3. Criterios de decisión en ambiente de incertidumbre	45
2.3.1. Criterio de Wald:	45
2.3.2. Criterio de Hurwicz:	46
2.3.3. Criterio de Laplace:	47
2.3.4. Criterio de Savage:	48

2.3.5.	Criterio de Bayes con probabilidad subjetiva:	49
3.	Decisiones Bayes y minimax	55
3.1.	Introducción	55
3.2.	Decisiones Bayes	55
3.2.1.	Interpretación geométrica	57
3.2.2.	Existencia de las acciones Bayes	61
3.3.	Admisibilidad y completitud	62
3.3.1.	Admisibilidad de las acciones Bayes	65
3.3.2.	Completitud de las acciones Bayes	66
3.3.3.	Completitud de las acciones no aleatorizadas	67
3.4.	Decisiones minimax	68
3.4.1.	Distribución menos favorable	70
3.4.2.	Teorema del minimax	71
3.4.3.	Métodos de determinación de la estrategia minimax	74
3.5.	Ejemplos adicionales	79
3.6.	* Apéndice: Teorema del hiperplano separador	86
4.	Decisión con experimentación	93
4.1.	Introducción	93
4.2.	Reglas de decisión Bayes	94
4.3.	Estadísticos suficientes	99
4.4.	Costes de experimentación	104
4.5.	La Inferencia Estadística como problema de decisión	108
4.5.1.	Estimación puntual	110
4.5.2.	Contraste de hipótesis	111
4.5.3.	Intervalos de confianza	112
4.6.	Invariancia en problemas de decisión	114
4.6.1.	Problemas de decisión invariantes	114
4.6.2.	Reglas de decisión invariantes minimax	117
4.6.3.	Admisibilidad de las reglas invariantes	118
4.6.4.	Reglas Bayes invariantes	119
5.	Decisión secuencial	131
5.1.	Introducción	131
5.2.	Reglas de decisión secuencial	131
5.3.	Reglas secuenciales Bayes	133

6. Decisiones múltiples	143
6.1. Introducción	143
6.2. Problemas de decisión múltiple monótonos	144
6.3. Problemas de clasificación	147
6.4. Problemas de deslizamiento	149
A. Solución de los ejercicios	153
Capítulo 1	153
Capítulo 2	163
Capítulo 3	170
Capítulo 4	184
Capítulo 5	237
Índice alfabético	241

1.3.1. Decisiones aleatorizadas sin experimentación

En general, el marco teórico para incluir decisiones aleatorizadas en un problema de decisión sin experimentación no es complicado. Simplemente debe suponerse que A está dotado de una σ -álgebra \mathcal{A} , que cumpla $\{a\} \in \mathcal{A}$ para cualquier $a \in A$ y respecto a la cual $L(\theta, \cdot)$ sean funciones medibles. Mientras A sea un conjunto discreto o un subconjunto de un espacio euclídeo, \mathbb{R}^n , ello no comporta ninguna dificultad, supuesto que las funciones de pérdida son razonablemente regulares.

Después, las **acciones aleatorizadas**, α , serán distribuciones de probabilidad en (A, \mathcal{A}) , que asignen una probabilidad $\alpha(A_1)$ a cada subconjunto $A_1 \in \mathcal{A}$. Con ellas, el decisor especifica las reglas que regirán el sorteo de la acción $a \in A$ que empleará.

La función de pérdida se extiende para acciones aleatorizadas mediante la expresión

$$L(\theta, \alpha) = \int_A L(\theta, a) \alpha(da). \quad (1.2)$$

Y, el espacio de todas las *acciones aleatorizadas* α sobre (A, \mathcal{A}) , que tengan una pérdida finita, para cada $\theta \in \Theta$, se designa habitualmente por A^* .

Por supuesto, puede considerarse que $A \subset A^*$, identificando cada acción $a \in A$ con la distribución causal que asigna probabilidad uno a $\{a\}$ (de ahí, la conveniencia de imponer que $\{a\} \in \mathcal{A}$).

Esta forma de proceder sustituye el problema de decisión original (A, Θ, L) por (A^*, Θ, L) , en el que L se ha extendido mediante (1.2). Con ello se amplía el conjunto de acciones posibles y el análisis del problema puede ser más sencillo en A^* que en A . Al menos, A^* tiene la ventaja de ser convexo, en el sentido de que, para cualesquiera $\alpha_1, \alpha_2 \in A^*$ y $\lambda \in [0, 1]$, la combinación lineal convexa $\lambda\alpha_1 + (1 - \lambda)\alpha_2$ vuelve a ser un elemento de A^* (o sea, una distribución en (A, \mathcal{A})).

Ejemplo 1.9

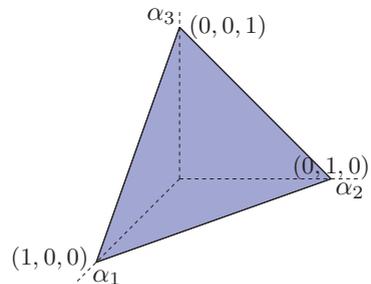
Si un problema de decisión cuenta sólo con tres acciones: $A = \{a_1, a_2, a_3\}$, entonces cada distribución de probabilidad α sobre A se caracteriza por

$$\alpha(a_1) = \alpha_1, \quad \alpha(a_2) = \alpha_2, \quad \alpha(a_3) = \alpha_3.$$

De modo que

$$A^* = \{(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}_+^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = 1\}$$

se representa como el triángulo que une los puntos $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$, en el espacio \mathbb{R}^3 . Lo mismo ocurriría en \mathbb{R}^n , si A contase con n acciones.



Supuesto que $L(\theta, a_1) = 2\theta$, $L(\theta, a_2) = \theta^2$ y $L(\theta, a_3) = \theta + 1$, el proceso de aleatorización añadiría infinitas acciones adicionales con

$$L(\theta, \alpha) = \alpha_1 2\theta + \alpha_2 \theta^2 + \alpha_3 (\theta + 1).$$

1.3.2. Decisiones aleatorizadas con experimentación

Inicialmente la idea de aleatorización en un problema de decisión con experimentación es la misma. Como se dijo, aquí el problema de decisión es (\mathcal{D}, Θ, R) , donde \mathcal{D} es la clase de todas las reglas de decisión $d : \mathcal{X} \mapsto A$ a las que (1.1) asocia un riesgo $R(\theta, d)$ finito.

Para definir una **regla de decisión aleatorizada** es necesario dotar a \mathcal{D} de una σ -álgebra, \mathfrak{D} , sobre la que definir medidas de probabilidad, δ , que indiquen la distribución con la que se sortea la regla de decisión $d \in \mathcal{D}$ que se utilizará. Dado que \mathcal{D} es un espacio funcional –de funciones de \mathcal{X} en A –, la elección de \mathfrak{D} puede ser más delicada ⁽⁵⁾. Pero no será necesario entrar en detalles técnicos sobre ello.

Lo que sí es importante es asociar con cada *regla de decisión aleatorizada* una **función de riesgo**:

$$R(\theta, \delta) = \int_{\mathcal{D}} R(\theta, e) \delta(de) \quad (1.3)$$

promedio de los riesgos $R(\theta, d)$ respecto a la distribución δ con la que la regla de decisión d es elegida.

El conjunto de todas las *reglas de decisión aleatorizadas*, δ , que tengan asociado un riesgo finito $R(\theta, \delta)$, se representa por \mathcal{D}^* . Igual que el caso sin experimentación, la aleatorización amplía el problema de decisión estadístico (\mathcal{D}, Θ, R) , sustituyéndolo por $(\mathcal{D}^*, \Theta, R)$.

Ahora bien, hay un procedimiento alternativo de introducir la aleatorización en un problema de decisión con experimentación. En lugar de sortear primero la regla de decisión d que se utilizará y observar, después, el resultado \mathbf{x} del experimento, para adoptar la acción $d(\mathbf{x})$, puede hacerse al revés. Primero

⁵ Debe hacerse de tal manera que pertenezcan a \mathfrak{D} los conjuntos cilíndricos

$$\{d \in \mathcal{D} \mid d(x_1) \in A_1, \dots, d(x_n) \in A_n\} \quad (x_1, \dots, x_n \in \mathcal{X}, A_1, \dots, A_n \in \mathcal{A}).$$

Así podrá hablarse de la probabilidad con la que δ elige una regla de decisión que asocia a cada una de las observaciones x_i una acción en A_i . También debe conseguirse que los riesgos $R(\theta, \cdot)$ sean funciones medibles de $(\mathcal{D}, \mathfrak{D})$ en (\mathbb{R}, \mathbb{B}) ; lo cual vuelve a ser posible si $L(\theta, \cdot)$ son funciones suficientemente regulares.

se observa el resultado \mathbf{x} del experimento y, después, se emplea una acción aleatorizada $\gamma(\mathbf{x}) \in A^*$, que es función del resultado obtenido \mathbf{x} .

Esto significa emplear lo que se denomina una **regla de comportamiento**, definida como una aplicación $\gamma : \mathcal{X} \mapsto A^*$. Puede observarse que ello supone introducir la experimentación, en el problema ya aleatorizado (A^*, Θ, L) del apartado anterior, de manera similar a como se hizo en la sección 1.2.4.

Repasando el procedimiento allí descrito, primero será necesario dotar a A^* de una σ -álgebra, que se representará por \mathcal{A}^* ⁽⁶⁾.

Segundo, y más importante, hay que definir el riesgo asociado a cada *regla de comportamiento* γ . En consonancia con (1.1), será

$$R(\theta, \gamma) = E_{\theta}[L(\theta, \gamma(\mathbf{X}))] = \int_{\mathcal{X}} L(\theta, \gamma(\mathbf{x})) P_{\theta}(d\mathbf{x}) \quad (1.4)$$

donde $L(\theta, \gamma(\mathbf{x}))$ es la extensión de la pérdida introducida en (1.2).

El conjunto de todas las *reglas de comportamiento*, γ , con riesgo finito, se representará por Γ . Así, (Γ, Θ, R) es una nueva ampliación del problema de decisión con experimentación (\mathcal{D}, Θ, L) , diferente de la que supone $(\mathcal{D}^*, \Theta, R)$.

Existe un trabajo teórico importante, debido a Wald y Wolfowitz (1951), que establece que ambos problemas, (Γ, Θ, R) y $(\mathcal{D}^*, \Theta, R)$, son equivalentes. En el sentido de que, para cada *regla de decisión aleatorizada* δ , existe una *regla de comportamiento* γ con el mismo riesgo: $R(\theta, \delta) = R(\theta, \gamma)$. Y, recíprocamente, para cada *regla de comportamiento* γ , hay una *regla de decisión aleatorizada* δ que tiene el mismo riesgo.

Las reglas de comportamiento son más fáciles de utilizar en la práctica, pero las reglas de decisión aleatorizadas simplifican muchos razonamientos.

Ejemplo 1.10

En la situación del ejemplo 1.7, había cuatro reglas de decisión: $\mathcal{D} = \{d_1, d_2, d_3, d_4\}$, cada una de las cuales figura como una fila de la tabla

	$x = 1$	$x = 2$		$\theta = 1$	$\theta = 2$
d_1	$a = 1$	$a = 1$	d_1	2	-3
d_2	$a = 1$	$a = 2$	d_2	3/4	9/4
d_3	$a = 2$	$a = 1$	d_3	-7/4	-5/4
d_4	$a = 2$	$a = 2$	d_4	-3	4

⁶ La forma estándar de hacerlo es tomar como \mathcal{A}^* la mínima σ -álgebra que hace medibles las aplicaciones $\alpha \mapsto \alpha(A_1)$ cualquiera que sea $A_1 \in \mathcal{A}$. Cabe suponer que las pérdidas $L(\theta, \alpha)$ son funciones medibles de α , respecto a \mathcal{A}^* .