JOSÉ GALVÃO LEITE

COMPARAÇÃO DE EXPERIMENTOS:

DOIS MÉTODOS EQUIVALENTES DE COMPARÁ-LOS

Dissertação apresentada ao Instituto de Matemática e Estatística da Universidade de São Paulo como parte dos requisitos para a obtenção do título de Mestre em Estatística Teórica e Probabilidade.

SÃO PAULO

1972

AGRADECIMENTOS

Desejamos expressar nossos agradecimentos às pessoas:

- Prof. Pedro Jesus Fernandez, pela orientação durante a execução deste trabalho;
- Prof. Luiz Arthaud Berthet, que nos iniciou na vida universitária;
- Prof. Carlos Alberto Barbosa Dantas, pela assistência que nos tem prestado junto ao Departamento de Estatística do I.M.E. da U.S.P.;
- Prof. Flávio Wagner Rodrigues, pelas valiosas discussões sôbre alguns tópicos deste trabalho.

I N D I C E

pg.

APRESENTAÇÃO

Dado um espaço mensurável $(\Omega,\mathcal{B})$ um experimento é um conjunto de medidas de probabilidade sôbre $(\Omega,\mathcal{B})$. Bohnenblust, Shapley e Sherman introduziram um método de comparar dois experimentos em têrmos dos vetores riscos associados a funções de decisão para ambos os experimentos (3.1.2). David Blackwell foi o autor de um outro método para comparar dois experimentos em têrmos de um kernel de Markov (definição 3.3.1). O que pretendemos nêste trabalho é aplicar um teorema devido a Blackwell-Sherman-Cartier (teorema 2.3.1) para mostrar que estes métodos são equivalentes (teorema 3.3.1). Para isto, deduzimos o teorema 2.3.1 como uma consequência de um teorema de Strassen (teorema 2.1.1). No capítulo 1 consideramos o necessário para a compreensão dos capítulos 2 e 3; no capítulo 2 são demonstrados os teoremas de Strassen e de Blackwell-Sherman-Cartier e no capítulo 3, são feitas considerações gerais sôbre experimentos e os métodos citados para compará-los, bem como é demonstrada a equivalência entre - tais métodos. No final do capítulo 3, são considerados os experimentos binomiais.

CAPÍTULO 1

PRELIMINARES DE PROBABILIDADE E ANÁLISE

1.1 - Kernel de Markov

Sejam $(\Omega_1,\mathcal{A}_1)$ e $(\Omega_2,\mathcal{A}_2)$ espaços mensuráveis. Um kernel de Markov de Ω_1 em Ω_2 é uma aplicação

$$P : \Omega_1 \times \mathcal{A}_2 \longrightarrow [0,1]$$

satisfazendo as seguintes condições:

a) para todo $\omega_1 \in \Omega_1$, $P(\omega_1,.)$ é uma medida de probabilidade sôbre $(\Omega_2,\mathcal{A}_2)$,

b) para todo $A_2 \in \mathcal{A}_2$, $P(.,A_2)$ é $\mathcal{A}_1$- mensurável.

PROPOSIÇÃO 1.1.1 - Se μ é uma medida de probabilidade sôbre $(\Omega_1,\mathcal{A}_1)$ e P é um kernel de Markov de Ω_1 em Ω_2 , então, $P\mu$ definida por

$$(P\mu)(A_2) = \int_{\Omega_1} P(\omega_1,A_2)\mu(d\omega_1) \quad , \quad A_2 \in \mathcal{A}_2$$

é uma medida de probabilidade sôbre $(\Omega_2,\mathcal{A}_2)$.

Prova: $(P\mu)(A_2) \geq 0$ para todo $A_2 \in \mathcal{A}_2$; para toda sequência $\{ A_n \}_{n\geq 1}$ de conjuntos disjuntos de $\mathcal{A}_2$, $P(.,A_n) \geq 0$, $n \geq 1$. Segue que

$$(P\mu)(\bigcup_{n=1}^{\infty} A_n) = \int_{\Omega_1} P(\omega_1, \bigcup_{n=1}^{\infty} A_n)\mu(d\omega_1) =$$

$$= \int_{\Omega_1} (\sum_{n=1}^{\infty} P(\omega_1,A_n)) \ \mu(d\omega_1) =$$

$$= \sum_{n=1}^{\infty} (\int_{\Omega_1} P(\omega_1,A_n)\mu(d\omega_1)) = \sum_{n=1}^{\infty} (P\mu)(A_n)$$

e portanto, $P\mu$ é σ-aditiva (ver [1], pg.36, corolário 4.13);

$$(P\mu)(\Omega_2) = \int_{\Omega_1} P(\omega_1,\Omega_2)\mu(d\omega_1) = \int_{\Omega_1} \mu(d\omega_1) = \mu(\Omega_1) = 1. \dagger$$

PROPOSIÇÃO 1.1.2 - Se uma aplicação $z : \Omega_2 \longrightarrow R$ é limitada, A_2- mensurável e P é um kernel de Markov de Ω_1 em Ω_2, então, a aplicação $zP : \Omega_1 \longrightarrow R$ definida por

$$(zP)(\omega_1) = \int_{\Omega_2} z(\omega_2)\ P\ (\omega_1,d\omega_2)$$

é limitada e A_1- mensurável.

Prova: Existe M , $M > 0$ tal que $|z(\omega_2)| \leq M$ para todo $\omega_2 \in \Omega_2$. Para todo $\omega_1 \in \Omega_1$,

$$|(zP)(\omega_1)| = |\int_{\Omega_2} z(\omega_2)\ P\ (\omega_1,d\omega_2)| \leq$$

$$\leq \int_{\Omega_2} |z(\omega_2)|\ P\ (\omega_1,d\omega_2) \leq M \int_{\Omega_2} P(\omega_1,d\omega_2) = M.$$

Logo, zP é limitada.

a) Se $z = I_A$, $A \in A_2$, $(zP)(\omega_1) = (I_A P)(\omega_1) =$

$$= \int_{\Omega_2} I_A(\omega_2)P(\omega_1,d\omega_2) = P(\omega_1,A)$$

para todo $\omega_1 \in \Omega_1$.

Portanto, $zP = I_A P = P(.,A)$ é A_1- mensurável;

b) Se z é uma função simples não negativa, isto é,

$$z = \sum_{i=1}^{n} a_i I_{A_i}$$

onde

$A_i \in \mathcal{A}_2, i=1,2,\ldots,n$; $A_i \cap A_j = \emptyset$, $i \neq j$, $i,j=1,2,\ldots,n$;
$\bigcup_{i=1}^{n} A_i = \Omega_2$ e $a_i \in R_+$, $i=1,2,\ldots,n$,

tem-se

$$(zP)(\omega_1) = \int_{\Omega_2} \sum_{i=1}^{n} a_i I_{A_i}(\omega_2) P(\omega_1, d\omega_2) =$$

$$= \sum_{i=1}^{n} (a_i \int_{\Omega_2} I_{A_i}(\omega_2) P(\omega_1, d\omega_2)) = \sum_{i=1}^{n} a_i P(\omega_1, A_i)$$

para todo $\omega_1 \in \Omega_1$.

Portanto,

$$zP = \sum_{i=1}^{n} a_i P(.,A_i) \text{ é } \mathcal{A}_1\text{- mensurável;}$$

c) Se z é uma função $\mathcal{A}_2$- mensurável não negativa,

$$z = \lim_n z_n \quad \text{onde} \quad \{z_n\}_{n\geq 1}$$

é uma sequência não decrescente de funções simples não negativas (ver [1], pg.13). Para todo $n\geq 1$, $z_n P$ é $\mathcal{A}_1$- mensurável pelo item b) e

$$\lim_n (z_n P)(\omega_1) = \lim_n \int_{\Omega_2} z_n(\omega_2) P(\omega_1, d\omega_2) = \int_{\Omega_2} \lim_n z_n(\omega_2) P(\omega_1, d\omega_2) =$$

$$= \int_{\Omega_2} z(\omega_2) P(\omega_1, d\omega_2) = (zP)(\omega_1)$$

para todo $\omega_1 \in \Omega_1$ (teorema da convergência monotônica). Como zP é o limite de uma sequência $\{z_n P\}_{n\geq 1}$ de funções $\mathcal{A}_1$- mensuráveis , zP é $\mathcal{A}_1$- mensurável;

d) Se z é uma função $\mathcal{A}_2$- mensurável qualquer

$$(zP)(\omega_1) = \int_{\Omega_2} z(\omega_2)P(\omega_1,d\omega_2) = \int_{\Omega_2} z^+(\omega_2)P(\omega_1,d\omega_2) -$$

$$- \int_{\Omega_2} z^-(\omega_2)P(\omega_1,d\omega_2) = (z^+P)(\omega_1) - (z^-P)(\omega_1) =$$

$$= (z^+P-z^-P)(\omega_1) \qquad \text{para todo} \quad \omega_1 \in \Omega_1 .$$

Então,

$$zP = z^+P - z^-P$$

$$(z^+ = zI_{[z\geq 0]} \; , \; z^- = -zI_{[z<0]}).$$

Como z^+P e z^-P são A_1- mensuráveis (item c), segue que zP é A_1- mensurável. †

Se (Ω,A) é um espaço mensurável a aplicação

$$I : \Omega\times A \longrightarrow [0,1]$$

definida por

$$I(\omega,A) = I_A(\omega) \quad , \quad \omega \in \Omega \; , \; A \in A$$

é um exemplo de um kernel de Markov de Ω em Ω. Para toda aplicação $z:\Omega \longrightarrow R$, A - mensurável tem-se $zI = z$.

1.2 - Funções sublineares e teorema de Hahn-Banach.

DEFINIÇÃO 1.2.1 - Seja E um espaço vetorial sôbre o corpo dos números reais. Uma aplicação $p:E \longrightarrow R$ se diz sublinear se as seguintes condições se verificarem:

a) $p(x+y)\leq p(x)+p(y)$ para todo $x,y \in E$;

b) $p(\alpha x) = \alpha p(x)$ para todo $x \in E$ e para todo $\alpha \in R_+$.

A condição a) diz que p é subaditiva e a condição b) diz que p é homogênea não negativa.

PROPOSIÇÃO 1.2.1 - Se p é uma função sublinear definida sôbre E, então,

1) $p(0) = 0$,

2) $p(x) \geq - p(-x)$ para todo $x \in E$.

Prova: Seja $x \in E$; $p(0) = p(0.x) = 0.p(x) = 0$; para todo $x \in E$, $0 = p(0) = p(x+(-x)) \leq p(x)+p(-x)$ e portanto, vale 2). †

PROPOSIÇÃO 1.2.2 - Seja E um espaço normado e seja p uma função sublinear definida sôbre E. São equivalentes as seguintes propriedades:

a) p é contínua

b) $\sup\limits_{||x|| \leq 1} |p(x)| < \infty$.

Prova:

a) $\Longrightarrow$ b)

p sendo contínua na origem, dado $\varepsilon > 0$ existe $\delta > 0$ tal que $||x|| < \delta$ implica $|p(x)| \leq \varepsilon$ e portanto, $||x|| \leq 1$ implica $|p(x)| \leq {}^{\varepsilon}/\delta$.

b) $\Longrightarrow$ a)

Como $\sup\limits_{||x|| \leq 1} |p(x)| \leq M < \infty$ seque que $|p(x) - p(y)| \leq M \, ||x-y||$ para todo $x,y \in E$. Com efeito, se $x = y$ então $|p(x)-p(y)| = = 0 = M||x-y||$. Suponhamos $x \neq y$. Se $p(x) \geq p(y)$ tem-se

$$p(x) = p(x-y+y) \leq p(x-y) + p(y) .$$

Logo,

$$0 \leq p(x) - p(y) \leq p(x-y)$$

e portanto,

$$|p(x) - p(y)| \leq |p(x-y)| = \left|p\left(\frac{x - y}{||x-y||}\right)\right| \; ||x-y|| \leq$$

$$\leq \sup_{||z|| \leq 1} |p(z)| \; ||x-y|| \leq$$

$$\leq M \, ||x-y|| \; ;$$

se $p(x) < p(y)$ tem-se

$$p(y) = p(y-x+x) \leq p(y-x) + p(x).$$

Logo,

$$0 < p(y) - p(x) \leq p(y-x)$$

e portanto,

$$|p(x) - p(y)| = |p(y) - p(x)| \leq |p(y-x)| =$$

$$= |p(\frac{y-x}{||y-x||})| \; ||y-x|| \leq$$

$$\leq \sup_{||z|| \leq 1} |p(z)| \; ||y-x|| \leq M \, ||x-y|| \, .$$

Seja $x_o \in E$. Dado $\varepsilon > 0$, arbitrário, $||x-x_o|| \leq \varepsilon/M$ implica

$$|p(x) - p(x_o)| \leq M \, ||x-x_o|| \leq M \, \varepsilon/M = \varepsilon \, . \dagger$$

TEOREMA 1.2.1 (de Hahn-Banach) - Seja E um espaço vetorial real e p uma função sublinear definida sobre E ; seja E_o um subespaço vetorial de E e seja ℓ_o uma forma linear definida sobre E_o e tal que $\ell_o(y) \leq p(y)$ para todo $y \in E_o$. - Então, existe uma forma linear ℓ definida sobre E , prolongando ℓ_o e tal que $\ell(x) \leq p(x)$ para todo $x \in E$.(*)

Ver demonstração deste teorema em [9] , pg. 271.

Seja E um espaço de Banach e E^* o dual algébrico de E . Para cada função sublinear contínua p definida sôbre E, consideremos o conjunto

$$A_p = \{ x^* : x^* \in E^* \text{ e } x^* \leq p \}$$

(*) Consideraremos sòmente espaços vetoriais reais.

($x^* \in E^*$; $x^* \leq p$ se e sòmente se $x^*(x) \leq p(x)$ para todo $x \in E$). Observamos que se $x^* \in A_p$, então, x^* é contínua sôbre E .

PROPOSIÇÃO 1.2.3 - A_p é um subconjunto não vazio, convexo de E^* e $\sup A_p = p$.

<u>Prova</u>: Se $E = \{0\}$ a proposição acima vale. Suponhamos $E \neq \{0\}$ e seja $x_o \in E$. Seja $u : [x_o] \longrightarrow R$ ($[x_o]$ é o subespaço vetorial de E gerado por x_o) definida por $u(y) = \alpha p(x_o)$, onde $y = \alpha x_o$, $\alpha \in R$. Afirmamos que $u(y) \leq p(y)$ para todo $y \in [x_o]$. De fato, para todo $y \in [x_o]$, $y = \alpha x_o$ com $\alpha \in R$. Se $\alpha \geq 0$, $u(y) = \alpha p(x_o) = p(\alpha x_o) = p(y)$. Se $\alpha < 0$, $u(y) = \alpha p(x_o) = (-\alpha)(-p(x_o))$. Pela proposição 1.2.1 , $-p(x_o) \leq$ $\leq p(-x_o)$. Segue que $u(y) = (-\alpha)(-p(x_o)) \leq (-\alpha)p(-x_o) =$ $= p((-\alpha)(-x_o)) = p(\alpha x_o) = p(y)$. Notando que u é uma forma linear definida sôbre $[x_o]$, pelo teorema de Hahn-Banach, existe uma forma linear x^* definida sobre E , prolongando u e tal que $x^*(x) \leq p(x)$ para todo $x \in E$. Logo $x^* \in A_p$ e $A_p \neq \emptyset$. Quaisquer que sejam $x^*, y^* \in A_p$, $x^* \leq p$ e $y^* \leq p$. Logo, para todo $\lambda \in (0,1]$,

$$\lambda x^* \leq \lambda p \ , \ (1-\lambda)y^* \leq (1-\lambda)p$$

e portanto,

$$\lambda x^* + (1-\lambda)y^* \leq \lambda p + (1-\lambda)p = p \ .$$

Se $\lambda = 0$, $\lambda x^* + (1-\lambda)y^* = y^* \leq p$. Então,

para todo $\lambda \in [0,1]$, $\lambda x^* + (1-\lambda)y^* \in A_p$, ou seja, A_p é convexo.

Finalmente, para mostrar que $\sup A_p = p$, seja a função $f : E \longrightarrow R$ definida por

$$f(x) = \sup_{x^* \in A_p} x^*(x) \ .$$

É imediato que $f \leq p$. Afirmamos que $f \geq p$. Com efeito, se $f \not\geq p$ existe $x_o \in E$, $x_o \neq 0$, tal que $f(x_o) < p(x_o)$. Seja $u : [x_o] \longrightarrow R$ definida por $u(y) = \alpha p(x_o)$, onde $y = \alpha x_o$, $\alpha \in R$. Pela demonstração feita acima existe

$$y^* \in A_p \quad \text{tal que} \quad y^*\big|[x_o] = u$$

Então,

$$y^*(x_o) \leq f(x_o) < p(x_o) = u(x_o) \qquad \text{o que é absurdo, pois}$$

$$y^*(x_o) = u(x_o) \text{ . Logo, } f = p = \sup A_p \text{ .}$$

†

1.3 - Família de funções sublineares fracamente mensuráveis e limitadas.

Seja E um espaço de Banach e $(\Omega,\mathcal{B},\mu)$ um espaço de probabilidade. Denotaremos por S_E o conjunto das funções - sublineares definidas sôbre E.

DEFINIÇÃO 1.3.1 - A aplicação

$$\omega \in \Omega \longrightarrow h_\omega \in S_E$$

se diz

a) fracamente mensurável se, para cada $x \in E$, a aplicação

$$\omega \in \Omega \longrightarrow h_\omega(x) \in R$$

é $\mathcal{B}$- mensurável;

b) limitada se existe uma constante K, $K > 0$, tal que

$$|h_\omega(x)| \leq K \,||x|| \text{ , para todo } \omega \in \Omega \text{ e para todo}$$

$x \in E$.

A relação $|h_\omega(x)| \leq K\,||x||$ implica, pela proposição 1.2.2, a continuidade das funções h_ω sôbre E .

PROPOSIÇÃO 1.3.1 - Se a aplicação $\omega \in \Omega \longrightarrow h_\omega \in S_E$ for fracamente mensurável e limitada, então, a função

$$h : E \longrightarrow R$$

definida por

$$h(x) = \int_\Omega h_\omega(x)\mu(d\omega)$$

é uma função sublinear contínua definida sôbre E.

<u>Prova</u>: Para cada $x \in E$ a aplicação $\omega \in \Omega \longrightarrow h_\omega(x) \in R$ é integrável. Então, para todo $x,y \in E$ e para todo $\alpha \in R_+$ tem-se

$$h(x+y) = \int_\Omega h_\omega(x+y)\mu(d\omega) \leq \int_\Omega (h_\omega(x)+h_\omega(y))\mu(d\omega) =$$

$$= \int_\Omega h_\omega(x)\mu(d\omega) + \int_\Omega h_\omega(y)\mu(d\omega) = h(x) + h(y) \quad \text{e}$$

$$h(\alpha x) = \int_\Omega h_\omega(\alpha x)\mu(d\omega) = \int_\Omega \alpha h_\omega(x)\mu(d\omega) = \alpha\int_\Omega h_\omega(x)\mu(d\omega) =$$

$$= \alpha h(x).$$

Logo, h é uma função sublinear definida sobre E. Como existe K, $K > 0$, tal que $|h_\omega(x)| \leq K\,||x||$, para todo $\omega \in \Omega$ e para todo $x \in E$, segue que $|h_\omega(x)| \leq K$ para todo $\omega \in \Omega$ e para todo $x \in E$, com $||x|| \leq 1$. Então,

$$|h(x)| = \left|\int_\Omega h_\omega(x)\mu(d\omega)\right| \leq \int_\Omega |h_\omega(x)|\mu(d\omega) \leq K$$

para todo $x \in E$ com $||x|| \leq 1$.

Portanto,

$$\sup_{||x||\leq 1} |h(x)| \leq K < \infty$$ e pela proposição 1.2.2 h é

contínua. $\dagger$

1.4 - Funções convexas, côncavas e afins

Seja E um espaço vetorial topológico localmente convexo e Ω um subconjunto convexo, compacto de E .

DEFINIÇÃO 1.4.1 - A função $f:\Omega \longrightarrow R$ se diz

a) convexa sôbre Ω se, para todo $x,y \in \Omega$ e para todo $t \in [0,1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) ;$$

b) côncava sôbre Ω se, para todo $x,y \in \Omega$ e para todo $t \in [0,1]$,

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y) ;$$

c) afim sobre Ω se, para todo $x,y \in \Omega$ e para todo $t \in [0,1]$,

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y).$$

Seja F o conjunto de todas as funções convexas, contínuas sôbre Ω . Do fato: se f é convexa (côncava), contínua sôbre Ω , então, $-f$ é côncava (convexa), contínua sôbre Ω , segue que $-F = \{ -f:f \in F \}$ é o conjunto de todas as funções côncavas, contínuas sôbre Ω . É claro que $F \cap -F$ é o conjunto de todas as funções afins, contínuas sôbre Ω .

PROPOSIÇÃO 1.4.1 - Seja F^* o conjunto de todas as funções convexas, contínuas sobre E . Então,

$$E' + R = F^* \cap -F^*$$

(E' é o dual topológico de E).

Prova: É imediato que $E' + R \subset F^* \cap -F^*$. Reciprocamente, se

$f \in F^* \cap -F^*$ seja $g : E \longrightarrow R$ definida por $g(x)=f(x)-f(0)$. Como f é contínua g é contínua. A função g é linear; com efeito, para todo $x,y \in E$ e para todo $\alpha \in R$,

a) se $0 \leq \alpha \leq 1$, $g(\alpha x) = g(\alpha x + (1-\alpha)0) = f(\alpha x + (1-\alpha)0) -$
$- f(0) = \alpha f(x)+(1-\alpha)f(0) - f(0) =$
$= \alpha(f(x) - f(0)) = \alpha g(x)$;

b) se $\alpha > 1$, $g(x) = g(\frac{\alpha x}{\alpha}) = \frac{1}{\alpha} g(\alpha x)$ o que implica
$g(\alpha x) = \alpha g(x)$;

c) $g(x+y) = g(2\frac{x+y}{2}) = 2g(\frac{x+y}{2}) = 2[f(\frac{x+y}{2}) - f(0)] =$

$= 2[\frac{1}{2} f(x) + \frac{1}{2} f(y) - f(0)] = f(x) + f(y) -$

$- 2f(0) = (f(x) - f(0)) + (f(y) - f(0))= g(x) + g(y)$;

d) se $\alpha < 0$, $g(\alpha x) - \alpha g(x) = g(\alpha x) + g(-\alpha x) = g(\alpha x - \alpha x) =$
$= g(0) = 0$ implica $g(\alpha x)=\alpha g(x)$.

Então,

$$f = g + f(0) \in E' + R \quad \text{e} \quad F^* \cap -F^* \subset E' + R \; . \; \dagger$$

1.5 - Medidas de Radon e resultante de uma medida

Indicaremos por $C(\Omega)$ o conjunto das funções contínuas definidas em Ω a valores em R ; $C(\Omega)$ munido da norma

$$f \in C(\Omega) \longrightarrow ||f||_\infty = \sup_{\omega \in \Omega} |f(\omega)| \in R_+$$

é um espaço de Banach.

DEFINIÇÃO 1.5.1 - Toda forma linear contínua sôbre $C(\Omega)$ se diz uma medida de Radon sôbre Ω .

Uma medida de Radon L sôbre Ω é positiva se, para toda fun ção $f \geq 0$, $f \in C(\Omega)$ tem-se $L(f) \geq 0$.

Seja $M^+(\Omega)$ o conjunto de todas as medidas de Radon positivas sôbre Ω . No que segue suporemos que E seja um espaço vetorial topológico localmente convexo e de Hausdorff e Ω um subconjunto convexo, compacto, metrisável de E.

DEFINIÇÃO 1.5.2 - Um vetor $a \in E$ se diz a resultante de $L \in M^+(\Omega)$ se e sòmente se, para toda $f \in E'$,

$$f(a) = L(f|_{\Omega}) .$$

Denotaremos a resultante de L por $r(L)$. Toda $L \in M^+(\Omega)$ admite uma única resultante (ver [6] , volume II , pg. 115). É imediato que, se μ é uma medida de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$ ($\mathcal{B}(\Omega)$ é a σ-álgebra de Borel de subconjuntos de Ω), a aplicação $I: C(\Omega) \longrightarrow R$ definida por

$$(1.5.1) \qquad I(f) = \int_\Omega f(\omega)\mu(d\omega)$$

é uma forma linear positiva sôbre $C(\Omega)$ tal que $I(1) = 1$ e portanto, I é uma medida de Radon sôbre Ω . A fim de simplificar as notações, usaremos um mesmo símbolo para representar uma medida de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$, e a medida de Radon positiva sôbre Ω correspondente àquela medida de probabilidade, segundo a fórmula 1.5.1 . Assim, se μ é uma medida de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$ a fórmula 1.5.1 torna-se

$$\mu(f) = \int_\Omega f(\omega)\mu(d\omega) \ , \quad f \in C(\Omega)$$

PROPOSIÇÃO 1.5.1 - Se μ é uma medida de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$ e g é uma função afim, contínua sôbre E , então, $r(\mu) \in \Omega$ e $g(r(\mu)) = \mu(g|_{\Omega})$.

Prova: $\mu \in M^1(\Omega) = \{ \lambda : \lambda \in M^+(\Omega)$ e $||\lambda|| = 1 \}$

onde

$$||\lambda|| = \sup_{||f||_\infty \leq 1} |\lambda(f)|^{(1)} \quad ; \quad \text{logo} \quad r(\mu) \in \Omega^{(2)}$$

Pela proposição 1.4.1 , $g = f+c$ onde $f \in E'$ e $c = g(0)$.

Então,

$$g(r(\mu)) = f(r(\mu)) + c = \mu(f|_\Omega) + c = \int_\Omega f(\omega)\mu(d\omega) + c =$$

$$= \int_\Omega (f(\omega) + c)\mu(d\omega) = \int_\Omega g(\omega)\mu(d\omega) = \mu(g|_\Omega) \ . \qquad \dagger$$

NOTAS E REFERÊNCIAS

(1) Para toda medida $\mu \geq 0$ sôbre Ω , tem-se $||\mu|| = \mu(1)$

Ver [5] , pg. 47, corolário 1 .

(2) Seja $M^1(\Omega) = \{ \lambda : \lambda \in M^+(\Omega)$ e $||\lambda|| = 1\}$, onde Ω é um subconjunto convexo, compacto de um espaço vetorial topológico localmente convexo e de Hausdorff. Se $\mu \in M^1(\Omega)$, então, $r(\mu) \in \Omega$.

Ver [6] , pg. 115, volume II, Proposição 26.3.

A matéria do capítulo I foi baseada nos capítulos II, IX, XI do livro do Meyer, "Probabilités et Potentiel" ; capítulo III do livro "Intégration", do Bourbaki e nos capítulos 4, volume I ; 6, volume II do livro do Choquet, "Lectures on Analysis".

CAPÍTULO 2

TEOREMA DE STRASSEN
E
TEOREMA DE BLACKWELL-SHERMAN-CARTIER

2.1 - TEOREMA 2.1.1 (Teorema de Strassen) :

Seja E um espaço de Banach separável e $(\Omega,\mathcal{B},\mu)$ um espaço de probabilidade. Suponhamos que a aplicação

$$\omega \in \Omega \longrightarrow h_\omega \in S_E$$

seja fracamente mensurável e limitada. Seja h a função sub-linear definida sôbre E por

$$h(x) = \int_\Omega h_\omega(x)\mu(d\omega) .$$

Seja $x^* \in E^*$. As seguintes propriedades são equivalentes:

i) x^* é dominada por h $(x^* \in A_h = \{y^* : y^* \in E^* \text{ e } y^* \leq h\})$;

ii) Existe uma aplicação $\omega \in \Omega \longrightarrow x^*_\omega \in E^*$ tal que

a) para todo $x \in E$ a aplicação $\omega \in \Omega \longrightarrow x^*_\omega(x) \in R$ é $\mathcal{B}$-mensurável;

b) $x^*_\omega(x) \leq h_\omega(x)$ para todo $x \in E$ e para quase todo $\omega \in \Omega$ $(x^*_\omega \in A_{h_\omega}$ para quase todo $\omega \in \Omega)$;

c) $x^*(x) = \int_\Omega x^*_\omega(x)\mu(d\omega)$ para todo $x \in E$.

<u>Prova:</u> ii) $\Longrightarrow$ i)

$$x^*(x) = \int_\Omega x^*_\omega(x)\mu(d\omega) \leq \int_\Omega h_\omega(x)\mu(d\omega) = h(x)$$

para todo $x \in E$.

i) $\Longrightarrow$ ii)

Vamos supor que o espaço de probabilidade $(\Omega,\mathcal{B},\mu)$ seja completo [(1)]. Se $(\Omega,\mathcal{B},\mu)$ não fôr completo, ver referência à pg. 35.

Seja L o espaço vetorial das funções $\mathcal{B}$-mensuráveis definidas em $(\Omega,\mathcal{B})$ a valores em $(E,\mathcal{B}(E))$ ($\mathcal{B}(E)$ é a σ-álgebra - de Borel de subconjuntos de E), que assumem sòmente um número finito de valores, sendo duas funções X e X' em L consideradas iguais se $X = X'$ em quase toda parte com relação a μ. Seja $X \in L$ e suponhamos que $Im(X) = \{x_1,x_2,\ldots,x_m\}$. É imediato que $(A_i)_{1\le i\le m}$, onde $A_i = X^{-1}(\{x_i\})$; $i=1,2,\ldots,m$ é uma partição mensurável finita de Ω . Considerando as funções $g,g_1,g_2,\ldots,g_m$ de Ω em R , definidas por $g(\omega) = h_\omega(X(\omega))$, $g_i(\omega) = h_\omega(x_i)$; $i=1,2,\ldots,m$, tem-se

$$g = \sum_{i=1}^{m} g_i I_{A_i}$$

e portanto, g é $\mathcal{B}$-mensurável.

Como existe K, $K > 0$ tal que $|h_\omega(X(\omega))| \le K \max \{||x_1||, ||x_2|| ,\ldots,||x_m||\}$ para todo $\omega \in \Omega$, segue que

$$\int_\Omega g(\omega)\mu(d\omega) = \int_\Omega h_\omega(X(\omega))\mu(d\omega) < \infty$$

Portanto, se $X \in L$ a aplicação

$$\omega \in \Omega \longrightarrow h_\omega(X(\omega)) \in R$$

é $\mathcal{B}$-mensurável e integrável. Consideremos a função $H: L \longrightarrow R$, definida por

$$(2.1.1) \qquad H(X) = \int_\Omega h_\omega(X(\omega))\mu(d\omega).$$

Se $X,X' \in L$ e $\alpha \in R_+$, então,

$$H(X+X') = \int_\Omega h_\omega((X+X')(\omega))\mu(d\omega) = \int_\Omega h_\omega(X(\omega)+X'(\omega))\mu(d\omega) \leq$$

$$\leq \int_\Omega (h_\omega(X(\omega)) + h_\omega(X'(\omega)))\mu(d\omega) = \int_\Omega h_\omega(X(\omega))\mu(d\omega) +$$

$$+ \int_\Omega h_\omega(X'(\omega))\mu(d\omega) = H(X) + H(X')$$

e

$$H(\alpha X) = \int_\Omega h_\omega((\alpha X)(\omega))\mu(d\omega) = \int_\Omega h_\omega(\alpha X(\omega))\mu(d\omega) =$$

$$= \int_\Omega \alpha h_\omega(X(\omega))\mu(d\omega) = \alpha H(X).$$

Logo, H é uma função sublinear sôbre L.

Seja $\tilde{L} = \{ \tilde{X}:\tilde{X} \in L$ e $\tilde{X}$ é constante em quase toda parte com relação a $\mu \}$. É imediato que $\tilde{L}$ é um subespaço vetorial de L e a função $\tilde{Q}:\tilde{L} \longrightarrow R$ definida por $\tilde{Q}(\tilde{X}) = x^*(x)$, onde x é o valor quase certo de $\tilde{X}$ com relação a μ, é uma forma linear sôbre $\tilde{L}$. (*) Notando que vale i) tem-se

$$\tilde{Q}(\tilde{X}) = x^*(x) \leq h(x) = \int_\Omega h_\omega(x)\mu(d\omega) = \int_\Omega h_\omega(\tilde{X}(\omega))\mu(d\omega) = H(\tilde{X})$$

para todo $\tilde{X} \in \tilde{L}$. Então, pelo teorema de Hahn-Banach, existe uma forma linear Q definida sôbre L, prolongando $\tilde{Q}$ e tal que

$$Q(X) \leq H(X) \quad \text{para todo } X \in L. \tag{2.1.2}$$

Se $x \in E$ e $A \in B$, como a função $xI_A:\Omega \longrightarrow E$ definida por

---

(*) $\tilde{X}(\omega) = x$ exceto para os ω pertencentes a um conjunto de μ-medida zero.

$$(xI_A)(\omega) = xI_A(\omega)$$

é $\mathcal{B}$-mensurável e assume sòmente os valores 0 e x, então, - $xI_A \in L$.

Por (2.1.1) e (2.1.2) tem-se

$$(2.1.3) \qquad Q(xI_A) \leq H(xI_A) = \int_\Omega h_\omega(xI_A(\omega))\mu(d\omega) =$$

$$= \int_A h_\omega(xI_A(\omega))\mu(d\omega) + \int_{A^c} h_\omega(xI_A(\omega))\mu(d\omega) =$$

$$= \int_A h_\omega(x)\mu(d\omega)$$

e

$$(2.1.4) \qquad -Q(xI_A) = Q(-xI_A) \leq H(-xI_A) = \int_\Omega h_\omega(-xI_A(\omega))\mu(d\omega) =$$

$$= \int_A h_\omega(-xI_A(\omega))\mu(d\omega) + \int_{A^c} h_\omega(-xI_A(\omega))\mu(d\omega) =$$

$$= \int_A h_\omega(-x)\mu(d\omega).$$

De (2.1.3) e (2.1.4) e pelo fato da aplicação

$$\omega \in \Omega \longrightarrow h_\omega \in S_E$$

ser limitada, existe K, $K > 0$, tal que

$$(2.1.5) \qquad -K\|x\|\mu(A) \leq Q(xI_A) \leq K\|x\|\mu(A).$$

Definindo-se, para cada $A \in \mathcal{B}$, $\mu_x(A) = Q(xI_A)$ tem-se que μ_x é uma medida sinalizada finita definida em $(\Omega,\mathcal{B})$ e $\mu_x \ll \mu$ (μ_x é absolutamente contínua com relação a μ).

De fato, $\mu_x(\emptyset) = 0$ e para toda sequência $\{A_n\}_{n\geq 1}$, $A_n \in \mathcal{B}$, $n \geq 1$, tal que $A_n \downarrow \emptyset$ tem-se $\lim_n \mu_x(A_n) = 0$ (ver 2.1.5) .

Então, como μ_x é aditiva, μ_x é σ-aditiva. Por (2.1.5)

μ_x é finita e $\mu_x << \mu$.[2] Logo, existe uma função

$$\omega \in \Omega \longrightarrow q_\omega(x) \in R ,$$

$\mathcal{B}$-mensurável, integrável, tal que

$$\mu_x(A) = \int_A q_\omega(x)\mu(d\omega)$$

para todo $A \in \mathcal{B}$ [3] e

$$\mu_x(A) = \int_A q_\omega(x)\mu(d\omega) = Q(xI_A) \leq \int_A h_\omega(x)\mu(d\omega) ,$$

para todo $A \in \mathcal{B}$, implica

(2.1.6) $\quad q_\omega(x) \leq h_\omega(x)$

para quase todo $\omega \in \Omega$ com relação a μ . Pelo fato de Q ser uma forma linear definida sôbre L , se $x,y, \in E$ e $a,b \in R$, então,

(2.1.7) $\quad q_\omega(ax+by) = aq_\omega(x)+bq_\omega(y)$,

para quase todo $\omega \in \Omega$ com relação a μ .

Seja E_o um subconjunto de E, enumerável, denso em E e estável por combinações lineares com coeficientes racionais. Por (2.1.6) e (2.1.7) , para cada ω pertencente ao completar - de um certo conjunto $\mathcal{B}$-mensurável A_o , com $\mu(A_o) = 0$, valem

(2.1.8) $\quad q_\omega(x) \leq h_\omega(x)$ para todo $x \in E_o$

e

(2.1.9) $\quad q_\omega(ax+by) = aq_\omega(x) + bq_\omega(y)$ para todo $x,y \in E_o$

e todo a,b racionais. Sem perda de generalidade podemos supor $A_o = \emptyset$, porque se $A_o \neq \emptyset$ basta considerar, para cada $\omega \in \Omega$, a função $g_\omega : E_o \longrightarrow R$, assim definida:

$$g_\omega(x) = \begin{cases} q_\omega(x) & \text{se } \omega \in A_o^c \\ y_\omega^*(x) & \text{se } \omega \in A_o , \text{ onde } y_\omega^* \in A_{h_\omega} . \end{cases}$$

Notamos que, para todo $x \in E_o$, a aplicação

$$\omega \in \Omega \longrightarrow g_\omega(x) \in R$$

é $\mathcal{B}$-mensurável e g_ω verifica (2.1.8) e (2.1.9) , para todo $\omega \in \Omega$.

Por (2.1.8) e (2.1.9) e pelo fato da aplicação

$$\omega \in \Omega \longrightarrow h_\omega \in S_E$$

ser limitada, para cada $\omega \in \Omega$,

$$|q_\omega(x)| \leq K||x||$$

para todo $x \in E_o$. Logo, para cada $\omega \in \Omega$, $q_\omega|_{E_o}$ é uniformemente contínua. Para cada $\omega \in \Omega$ seja x^*_ω a única extensão contínua de $q_\omega|_{E_o}$ ao espaço E . Então,

$$x^*_\omega(x) = \lim_{\substack{y\to x \\ y\in E_o}} q_\omega(y)$$

para cada $x \in E$.(4) Para cada $\omega \in \Omega$, $x^*_\omega \in E^*$. Com efeito, para todo $x,y \in E$ e todo $\alpha \in R$, existem sequências $\{x_n\}_{n\geq 1}$, $\{y_n\}_{n\geq 1}$ e $\{\alpha_n\}_{n\geq 1}$; x_n, $y_n \in E_o$ e α_n racional, $n \geq 1$, com $x_n \longrightarrow x$, $y_n \longrightarrow y$ e $\alpha_n \longrightarrow \alpha$. Então, $x_n+y_n \longrightarrow x+y$, $\alpha_n x_n \longrightarrow \alpha x$ e por (2.1.9), para cada $\omega \in \Omega$,

$$x^*_\omega(x+y) = \lim_n q_\omega(x_n+y_n) = \lim_n (q_\omega(x_n)+q_\omega(y_n)) =$$

$$= \lim_n q_\omega(x_n) + \lim_n q_\omega(y_n) = x^*_\omega(x) + x^*_\omega(y)$$

e

$$x^*_\omega(\alpha x) = \lim_n q_\omega(\alpha_n x_n) = \lim_n (\alpha_n q_\omega(x_n)) = \lim_n \alpha_n . \lim_n q_\omega(x_n) =$$

$$= \alpha x^*_\omega(x) .$$

Portanto, existe uma aplicação $\omega \in \Omega \longrightarrow x^*_\omega \in E^*$.

Provemos que os ítens a), b) e c) da propriedade ii) se verificam. Se $x \in E$ seja $\{x_n\}_{n\geq 1}$ uma sequência de elementos de E_o , $x_n \longrightarrow x$. Como $\{q_{.}(x_n)\}_{n\geq 1}$ é uma sequência de funções $\mathcal{B}$-mensuráveis e para cada $\omega \in \Omega$,

$$x_\omega^*(x) = \lim_n q_\omega(x_n) \quad ,$$

então, a aplicação

$$\omega \in \Omega \longrightarrow x_\omega^*(x) \in R$$

é $\mathcal{B}$-mensurável, o que vem verificar a) ; para cada $\omega \in \Omega$, $q_\omega(x_n) \leq h_\omega(x_n)$, $n \geq 1$ (ver 2.1.8) . Notando que h_ω é contínua, tem-se

$$x_\omega^*(x) = \lim_n q_\omega(x_n) \leq \lim_n h_\omega(x_n) = h_\omega(x) \, ,$$

o que verifica b).

Seja x um elemento qualquer de E. Para toda sequência $\{x_n\}_{n\geq 1}$ de elementos de E tal que $x_n \longrightarrow x$ e para todo $A \in \mathcal{B}$,

$$|\mu_{x_n}(A) - \mu_x(A)| = |Q(x_n I_A) - Q(xI_A)| = |Q(x_n - x)I_A| \leq$$

$$\leq K||x_n - x||\mu(A) \leq K||x_n - x|| \quad , \; n \geq 1$$

(ver 2.1.5), implica

$$0 \leq \sup_{A\in\mathcal{B}} |\mu_{x_n}(A) - \mu_x(A)| \leq K||x_n - x|| \quad , \; n \geq 1 \, .$$

Então,

$$\lim_n \sup_{A\in\mathcal{B}} |\mu_{x_n}(A) - \mu_x(A)| = 0$$

e para todo $A \in \mathcal{B}$, dado $\varepsilon > 0$, arbitrário, existe $n_o(\varepsilon)>0$ tal que, para todo $n \geq n_o$

$$\left|\int_A q_\omega(x_n)\mu(d\omega) - \int_A q_\omega(x)\mu(d\omega)\right| = |\mu_{x_n}(A) - \mu_x(A)| \leq$$

$$\leq \sup_{A\in\mathcal{B}} |\mu_{x_n}(A) - \mu_x(A)| < \varepsilon \, .$$

Logo,

$$\int_A q_\omega(x_n)\mu(d\omega) \longrightarrow \int_A q_\omega(x)\mu(d\omega)$$

uniformemente em $A \in \mathcal{B}$ o que implica

$$(2.1.10) \qquad \lim_n \int_\Omega |q_\omega(x_n) - q_\omega(x)|\mu(d\omega) = 0 \; . \; ^{(5)}$$

Consideremos uma sequência $\{y_n\}_{n\geq 1}$ de elementos de E_o tal que $y_n \longrightarrow x$. Por (2.1.10)

$$q_.(y_n) \xrightarrow{\mu} q_.(x) \; . \; ^{(6)}$$

Como existe uma subsequência $\{q_.(y_{n_K})\}_{K\geq 1}$ de $\{q_.(y_n)\}_{n\geq 1}$ tal que $\{q_.(y_{n_K})\}_{K\geq 1}$ converge quase certamente com relação a μ para $q_.(x)$ (7), segue que,

$$x_\omega^*(x) = \lim_n q_\omega(y_n) = \lim_K q_\omega(y_{n_K}) = q_\omega(x)$$

para quase todo $\omega \in \Omega$ com relação a μ e portanto,

$$\mu_x(\Omega) = \int_\Omega q_\omega(x)\mu(d\omega) = \int_\Omega x_\omega^*(x)\mu(d\omega).$$

Finalmente, como $x_\omega^*(x) = \tilde{Q}(\tilde{X})$, onde $\tilde{X} \in \tilde{L}$ e x é o valor quase certo de $\tilde{X}$ com relação a μ, tem-se

$$x^*(x) = \tilde{Q}(\tilde{X}) = Q(\tilde{X}) = Q(xI_\Omega) = \mu_x(\Omega) = \int_\Omega x_\omega^*(x)\mu(d\omega)$$

o que verifica c) . †

Se Ω for finito o teorema de Strassen segue da

PROPOSIÇÃO 2.1.1 - Seja E um espaço vetorial, $p_1, p_2, \ldots, p_n$ funções sublineares definidas sôbre E e $\lambda_1, \lambda_2, \ldots, \lambda_n$ números reais positivos tal que $\sum_{i=1}^n \lambda_i > 0$. Seja $x^* \in E$. As seguintes propriedades são equivalentes:

i) x^* é denominada por $\lambda_1 p_1 + \lambda_2 p_2 + \ldots + \lambda_n p_n$;

ii) Existem $x_1^*, x_2^*, \ldots, x_n^* \in E^*$ com $x_i^* \leq p_i$; $i=1,2,\ldots,n$ e

$$x^* = \lambda_1 x_1^* + \lambda_2 x_2^* + \ldots + \lambda_n x_n^* \quad .$$

<u>Prova</u>: ii) $\Longrightarrow$ i) é imediato

i) $\Longrightarrow$ ii)

A função $p: E^n \longrightarrow R$ definida por

$$p((x_1, x_2, \ldots, x_n)) = \lambda_1 p_1(x_1) + \lambda_2 p_2(x_2) + \ldots + \lambda_n p_n(x_n)$$

é sublinear sôbre E^n . Seja

$$F = \{(x_1, x_2, \ldots, x_n) \in E^n : x_1 = x_2 = \ldots = x_n\}$$

e $f: F \longrightarrow R$ definida por

$$f((x, \ldots, x)) = x^*(x) \; .$$

É imediato que F é um subespaço vetorial de E^n e f é uma forma linear definida sôbre F , tal que $f(y) \leq p(y)$ para todo $y \in F$. Logo, pelo teorema de Hahn-Banach, existe uma forma linear $\overline{f}$ definida sôbre E^n , prolongando f , tal que $\overline{f}(y) \leq p(y)$ para todo $y \in E^n$. Supondo, sem perda de generalidade, que $\lambda_1 > 0$; $i = 1, 2, \ldots, n$, as funções $x_i^* : E \longrightarrow R$ definidas por

$$x_i^*(x) = \frac{1}{\lambda_i} \overline{f}((0, \ldots, 0, \underset{i}{x}, 0, \ldots, 0)) \; ; \; i = 1, 2, \ldots, n,$$

são formas lineares definidas sôbre E. Para todo $x \in E$,

$$\begin{aligned} f((x, \ldots, x)) &= x^*(x) = \overline{f}((x, \ldots, x)) = \overline{f}((x, 0, \ldots, 0) + \\ &+ (0, x, 0, \ldots, 0) + \ldots + (0, 0, \ldots, x)) = \\ &= \overline{f}((x, \ldots, 0)) + \overline{f}((0, x, \ldots, 0)) + \ldots + \\ &+ \overline{f}((0, \ldots, x)) = \lambda_1 x_1^*(x) + \lambda_2 x_2^*(x) + \ldots + \\ &+ \lambda_n x_n^*(x) = (\lambda_1 x_1^* + \lambda_2 x_2^* + \ldots + \lambda_n x_n^*)(x) \end{aligned}$$

e

$$\begin{aligned} x_i^*(x) &= \frac{1}{\lambda_i} \overline{f}((0, \ldots, 0, \underset{i}{x}, 0, \ldots, 0)) \leq \frac{1}{\lambda_i} p((0, \ldots, 0, \underset{i}{x}, 0 \\ &\ldots, 0)) = \frac{1}{\lambda_i} (\lambda_1 p_1(0) + \lambda_2 p_2(0) + \ldots + \lambda_i p_i(x) + \ldots + \end{aligned}$$

$$+ \lambda_n p_n(0)) = \frac{1}{\lambda_i} (\lambda_i p_i(x)) = p_i(x) \ ;$$

$i=1,2,\ldots,n$ o que prova ii) . †

2.2 - Dilatação sôbre Ω e uma relação de ordem em $M^+(\Omega)$

Seja E um espaço vetorial topológico localmente - convexo e de Hausdorff e Ω um subconjunto convexo, compacto, metrisável de E .

DEFINIÇÃO 2.2.1 - Uma dilatação sôbre Ω é um kernel de Markov P de Ω em Ω (o espaço mensurável considerado é $(\Omega,\mathcal{B}(\Omega))$) tal que, para toda função afim, contínua z sôbre Ω ,

$$zP = z .$$

PROPOSIÇÃO 2.2.1 - Se P é uma dilatação sôbre Ω , então, $r(P(\omega,.)) = \omega$ para todo $\omega \in \Omega$.

<u>Prova</u>: Para toda $f \in E'$, $f|_\Omega$ é afim, contínua sôbre Ω . Então, se $\omega \in \Omega$,

$$f(r(P(\omega,.))) = P(\omega,f|_\Omega) = \int_\Omega f(\omega')P(\omega,d\omega') = (f|_\Omega P)(\omega) =$$

$$= f(\omega)$$

implica

$$f(r(P(\omega,.))-\omega) = 0$$

e portanto,

$$r(P(\omega,.)) = \omega . \ †$$

DEFINIÇÃO 2.2.2 - Se $\mu,\nu \in M^+(\Omega)$ diremos que $\mu \prec \nu$ (ν é mais difusa do que μ) se e sòmente se

$$\mu(f) \leq \nu(f) \quad ,$$

para toda função f convexa, contínua sôbre Ω .

Dizer que $\mu \prec \nu$ significa que a massa de ν é mais concentrada nas vizinhanças dos pontos extremos de Ω do que μ. A relação $\prec$ é uma relação de ordem em $\mathcal{M}^+(\Omega)$ (para uma inteira discussão sôbre o assunto ver [6], volume II, capítulo 6). - Lembrando que, a toda medida de probabilidade m sôbre $(\Omega, \mathcal{B}(\Omega))$ corresponde uma medida de Radon positiva sôbre Ω, segundo a fórmula

$$m(f) = \int_\Omega f(\omega) m(d\omega) \quad , \quad f \in C(\Omega) \quad ,$$

temos: se μ e ν são medidas de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$, então, $\mu \prec \nu$ se e sòmente se

$$\mu(f) = \int_\Omega f(\omega)\mu(d\omega) \leq \int_\Omega f(\omega)\nu(d\omega) = \nu(f) \quad ,$$

para toda função f convexa, contínua sôbre Ω.

2.3 - TEOREMA 2.3.1 (Teorema de Blackwell - Sherman - Cartier): μ, ν são medidas de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$; $\mu \prec \nu$ se e sòmente se existe uma dilatação P sôbre Ω tal que $\nu = P\mu$.(8)

<u>Prova</u>: $\Longrightarrow$

Demonstremos, inicialmente, o seguinte

<u>Lema</u> - Definindo-se

$$h_\omega(f) = \inf \{g(\omega) : g \in -F \text{ e } g \geq f\} \; , \; f \in C(\Omega), \; \omega \in \Omega \; ,$$

valem as seguintes propriedades:

a) Para cada $\omega \in \Omega$ a função h_ω é uma função sublinear sôbre $C(\Omega)$;

b) Para todo $\omega \in \Omega$ e para toda $f \in C(\Omega)$, $|h_\omega(f)| \leq ||f||_\infty$;

c) Para cada $f \in C(\Omega)$ a função

$$\omega \in \Omega \longrightarrow h_\omega(f) \in R$$

é $\mathcal{B}(\Omega)$- mensurável.

<u>Prova</u>: a) Seja $\omega \in \Omega$. Para toda $f,g \in C(\Omega)$ e para todo $\alpha \in R_+$ tem-se: i) para toda $v_1, v_2 \in -F$ com $v_1 \geq f$ e $v_2 \geq g$ - tem-se $v_1+v_2 \in -F$ e $v_1+v_2 \geq f+g$ o que implica $v_1(\omega)+v_2(\omega) \geq$ $\geq h_\omega(f+g)$. Então, $h_\omega(f+g) \leq \inf\{v_1(\omega)+v_2(\omega) : v_1, v_2 \in -F$ e $v_1 \geq f,\ v_2 \geq g\} = h_\omega(f)+h_\omega(g)$; ii) Se $\alpha=0$, $h_\omega(\alpha f) = h_\omega(0) = 0 =$ $= \alpha h_\omega(f)$. Se $\alpha>0$, para toda $v \in -F$ com $v \geq f$ tem-se $\alpha v \in -F$ e $\alpha v \geq \alpha f$. Logo, $h_\omega(\alpha f) \leq \alpha v(\omega)$ ou $\frac{1}{\alpha} h_\omega(\alpha f) \leq v(\omega)$, o que implica $h_\omega(\alpha f) \leq \alpha h_\omega(f)$. Por outro lado, para toda $v \in -F$, $v \geq \alpha f$ tem-se $\frac{1}{\alpha} v \in -F$ e $\frac{1}{\alpha} v \geq f$. Logo, $h_\omega(f) \leq$ $\leq \frac{1}{\alpha} v(\omega)$ ou $\alpha h_\omega(f) \leq v(\omega)$, o que implica $\alpha h_\omega(f) \leq h_\omega(\alpha f)$. Portanto, $h_\omega(\alpha f) = \alpha h_\omega(f)$.

b) Seja $\omega \in \Omega$. Para toda $f \in C(\Omega)$, $h_\omega(f) = \inf\{g(\omega) : g \in -F$ e $g \geq f\} \leq \sup_{t \in \Omega} |f(t)| = ||f||_\infty$. Notando que $-f \in C(\Omega)$ e h_ω é uma função sublinear sôbre $C(\Omega)$ (item a), tem-se

$$-||f||_\infty \leq -h_\omega(-f) \leq h_\omega(f) \leq ||f||_\infty$$

e portanto

$$|h_\omega(f)| \leq ||f||_\infty .$$

c) segue do fato: para cada $f \in C(\Omega)$ a função

$$\omega \in \Omega \longrightarrow h_\omega(f) \in R$$

é semi contínua superiormente (ver [6], pg. 123, volume II). †

Voltando a prova do teorema notamos que $C(\Omega)$ é um espaço de Banach separável, $(\Omega, \mathcal{B}(\Omega), \mu)$ é um espaço de probabilidade e a aplicação

$$\omega \in \Omega \longrightarrow h_\omega \in S_{C(\Omega)}$$

é fracamente mensurável e limitada (lema); então, as hipóteses do teorema de Strassen se verificam. Seja x^* a forma linear positiva sôbre $C(\Omega)$ definida por

$$x^*(f) = \int_\Omega f(\omega)\nu(d\omega).$$

Afirmamos que x^* é dominada por h, onde h é a função sub linear definida sôbre $C(\Omega)$ por

$$h(f) = \int_\Omega h_\omega(f)\mu(d\omega) .$$

Com efeito, para cada $f \in C(\Omega)$, da separabilidade de $\{ g : g \in -F \text{ e } g \geq f \}$ segue que existe uma sequência não crescente $\{g_n\}_{n\geq 1}$, $g_n \in -F$, $n \geq 1$, tal que

$$h_\omega(f) = \lim_n g_n(\omega)$$

para todo $\omega \in \Omega$. Como

$$h_\omega(f) = \inf \{g(\omega) : g \in -F \text{ e } g \geq f\} \geq f(\omega) \text{ e } g_n(\omega) \downarrow h_\omega(f),$$

para todo $\omega \in \Omega$, então,

$$x^*(f) = \int_\Omega f(\omega)\nu(d\omega) \leq \int_\Omega h_\omega(f)\nu(d\omega) ,$$

$$\int_\Omega g_n(\omega)\mu(d\omega) \downarrow \int_\Omega h_\omega(f)\mu(d\omega)$$

e

$$\int_\Omega g_n(\omega)\nu(d\omega) \downarrow \int_\Omega h_\omega(f)\nu(d\omega)$$

(ver [10], pg 40, proposição II-3-3). Pela hipótese $(\mu \prec \nu)$ tem-se

$$\int_\Omega g_n(\omega)\nu(d\omega) \leq \int_\Omega g_n(\omega)\mu(d\omega) .$$

Portanto,

$$x^*(f) = \int_\Omega f(\omega)\nu(d\omega) \leq \int_\Omega h_\omega(f)\nu(d\omega) = \lim_n \int_\Omega g_n(\omega)\nu(d\omega) \leq$$

$$\leq \lim_n \int_\Omega g_n(\omega)\mu(d\omega) = \int_\Omega h_\omega(f)\mu(d\omega) = h(f) .$$

Logo, a condição i) do teorema de Strassen se verifica e consequentemente, pela condição ii) do mesmo teorema, existe uma aplicação

$$\omega \in \Omega \longrightarrow x_\omega^* \in C(\Omega)^*$$

tal que: a) para toda $f \in C(\Omega)$ a aplicação

$$\omega \in \Omega \text{ —— } x_\omega^*(f) \in R$$

é $\mathcal{B}(\Omega)$-mensurável ; b) $x_\omega^*(f) \leq h_\omega(f)$ para toda $f \in C(\Omega)$ e para quase todo $\omega \in \Omega$; c) $x^*(f) = \int_\Omega x_\omega^*(f)\mu(d\omega)$ para toda $f \in C(\Omega)$.

Para quase todo $\omega \in \Omega$,

$$x_\omega^*(1) \leq h_\omega(1) = 1 \quad ; \quad x_\omega^*(1) = -x_\omega^*(-1) \geq -h_\omega(-1) = 1 \ ;$$

para toda $f \in C(\Omega)$, $f \geq 0$,

$$x_\omega^*(f) = -x_\omega^*(-f) \geq -h_\omega(-f) \geq 0 .$$

Então, para cada ω pertencente ao complementar de um certo conjunto $\mathcal{B}(\Omega)$-mensurável Ω_o com $\mu(\Omega_o) = 0$, x_ω^* é uma forma linear positiva sôbre $C(\Omega)$ tal que $x_\omega^*(1) = 1$, e portanto, a x_ω^* corresponde uma medida de probabilidade $P(\omega,.)$ sôbre $(\Omega,\mathcal{B}(\Omega))$, tal que

$$x_\omega^*(f) = \int_\Omega f(\omega')P(\omega,d\omega') , \quad f \in C(\Omega) . \tag{9}$$

Supondo $\Omega_o = \emptyset$,

seja $\mathcal{D}$ a classe dos subconjuntos fechados de Ω e $\tau=\{A:A \in \mathcal{B}(\Omega)$ e $P(.,A)$ é $\mathcal{B}(\Omega)$-mensurável $\}$. Se $F \in \mathcal{D}$ existe uma sequência não crescente $\{f_n\}_{n\geq 1}$, $f_n \in C(\Omega)$ e $0 \leq f_n(\omega) \leq 1$, para todo $\omega \in \Omega$, $n \geq 1$, que converge ponto a ponto para I_F (ver [2] , pgs. 8 e 9). Logo, pelo teorema da convergência dominada de Lebergue, para cada $\omega \in \Omega$,

$$P(\omega,F) = \int_\Omega I_F(\omega')P(\omega,d\omega') = \int_\Omega \lim_n f_n(\omega')P(\omega,d\omega') =$$

$$= \lim_n \int_\Omega f_n(\omega')P(\omega,d\omega') = \lim_n x_\omega^*(f_n) .$$

Como a aplicação

$$\omega \in \Omega \longrightarrow x_\omega^*(f_n) \in R$$

é $\mathcal{B}(\Omega)$-mensurável, para todo $n \geq 1$ (item a)), segue que $P(.,F)$ é $\mathcal{B}(\Omega)$- mensurável e portanto, $\mathcal{D} \subset \tau$.

É imediato que i) $\Omega \in \tau$;

ii) Se $A_1, A_2 \in \tau$, então, $A_1 + A_2 \in \tau$ se $A_1 \cap A_2 = \emptyset$ e $A_1 - A_2 \in \tau$ se $A_1 \supset A_2$;

iii) Se $\{A_n\}_{n\geq 1}$ é uma sequência monotônica crescente, $A_n \in \tau$, $n \geq 1$, então, $\lim_n A_n = \bigcup_{n=1}^{\infty} A_n \in \tau$; isto é, τ é uma classe σ-aditiva de subconjuntos de Ω e consequentemente $\tau = \mathcal{B}(\Omega)$ (ver [10], pg. 19, exercício II-4-5) e portanto, P é um kernel de Markov de Ω em Ω. Como para toda função f côncava, contínua sôbre Ω ,

$$(fP)(\omega) = \int_\Omega f(\omega')P(\omega,d\omega') = x_\omega^*(f) \leq h_\omega(f) = f(\omega)$$

para todo $\omega \in \Omega$; para toda função g convexa, contínua sôbre Ω , $-g$ é côncava, contínua sôbre Ω e

$$-(gP)(\omega) = ((-g)P)(\omega) \leq (-g)(\omega) = -g(\omega)$$

para todo $\omega \in \Omega$ ou

$$(gP)(\omega) \geq g(\omega)$$

para todo $\omega \in \Omega$, tem-se: para toda função z afim, contínua sôbre Ω , $zP = z$, ou seja, P é uma dilatação sôbre Ω . Finalmente, pelo item c) ,

$$\int_\Omega f(\omega)\nu(d\omega) = x^*(f) = \int_\Omega x^*_\omega(f)\mu(d\omega) =$$

$$= \int_\Omega \{\int_\Omega f(\omega')P(\omega,d\omega')\}\mu(d\omega) = \int_\Omega f(\omega)(P\mu)(d\omega)$$

para toda $f \in C(\Omega)$, o que implica $\nu = P\mu$ (ver [2] , pg. 9, teorema 1.3) .

$\Longleftarrow$

Se f é uma função convexa, contínua sôbre Ω , para todo $\omega \in \Omega$,

$$f(r(P(\omega,.))) \leq P(\omega,f)$$

(ver [6], volume II, pg. 120). Então, pela proposição 2.2.1,

$$f(\omega) = f(r(P(\omega,.))) \leq P(\omega,f) = \int_\Omega f(\omega')P(\omega,d\omega') = (fP)(\omega)$$

para todo $\omega \in \Omega$, o que implica

$$\mu(f) = \int_\Omega f(\omega)\mu(d\omega) \leq \int_\Omega \{ \int_\Omega f(\omega')P(\omega,d\omega') \}\mu(d\omega) =$$

$$= \int_\Omega (fP)(\omega)\mu(d\omega) = \int_\Omega f(\omega)(P\mu)(d\omega) = \int_\Omega f(\omega)\nu(d\omega) =$$

$$= \nu(f) \ . \quad \dagger$$

NOTAS E REFERÊNCIAS

(1) Todo subconjunto de um conjunto de medida nula pertence a $\mathcal{B}$.

(2) Ver [10] , pg. 104, corolário 2.

(3) Ver [10] , pg. 105, proposição IV-1-4 (Radon-Nykodim).

(4) Sejam S um subconjunto denso de um espaço métrico M e $f:S \longrightarrow N$ uma aplicação uniformemente contínua, onde N é um espaço métrico completo. Então, f admite uma única extensão contínua $\bar{f}:M \longrightarrow N$, a qual é uniformemente contínua (ver [8], pg. 231, volume 2).

Seja $(\Omega, \mathcal{B}, P)$ um espaço de probabilidade

(5) Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias reais, integráveis e X uma variável aleatória real, integrável. $X_n \xrightarrow{L_1} X$ se e sòmente se $\int_A X_n dP \longrightarrow \int_A XdP$ uniformemente em $A \in \mathcal{B}$.

Ver [10], pg. 49, proposição II-5-3.

(6) Convergência em L^1 implica convergência em medida.

Ver [1], pg. 69.

(7) Seja $\{X_n\}_{n\geq 1}$ uma sequência de v.a.r. Se $X_n \xrightarrow{P} X$, X v.a.r, então, existe uma subsequência $\{X_{n_k}\}_{k\geq 1}$ de $\{X_n\}_{n\geq 1}$, tal que $X_{n_k} \longrightarrow X$ em quase toda parte com relação a P.

Ver [1], pg. 69.

(8) Vários autores estudaram o assunto envolvido neste teorema. Citaremos, em particular, Blackwell que estabeleceu este resultado para um intervalo da reta (ver [4]), Sherman que o demonstrou para as medidas tendo um suporte finito e um trabalho não publicado de Fell foi uma etapa essencial na solução do problema, devido a Cartier. Por outro lado, Modobodzki também estabeleceu este resultado (Meyer - Probabilités et Potentiel, pg. 288).

(9) i) Ω é metrisável; então, as σ-álgebras de Baire e de Borel são iguais (ver [10], pg. 60).

ii) A fórmula $E(f) = \int_\Omega f(\omega)P(d\omega)$, $f \in C(\Omega)$, estabelece uma correspondência biunívoca entre as formas lineares po

sitivas E sôbre $C(\Omega)$ tais que $E(1) = 1$ e as medidas de probabilidade sôbre $(\Omega, \mathcal{B}(\Omega))$.
(ver [10] , pg. 64, proposição II-7-5) .

Na demonstração do teorema de Strassen supomos $(\Omega, \mathcal{B}, \mu)$ completo. Contudo, se o espaço não fôr completo, a demonstração é praticamente a mesma, exceto em um ponto: se o conjunto A_o considerado à página 22 fôr não vazio, para cada $\omega \in \Omega$, consideramos a função $g_\omega : E_o \longrightarrow R$, definida por

$$\begin{aligned} g_\omega(x) &= q_\omega(x) \quad \text{se} \quad \omega \in A_o^c \\ &= 0 \quad \text{se} \quad \omega \in A_o \, . \end{aligned}$$

O capítulo 2 foi baseado nos capítulos: XI do livro do Meyer, "Probabilités et Potentiel" ; 6, volume II do livro do Choquet, "Lectures on Analysis" e em [12] .

CAPÍTULO 3

COMPARAÇÃO DE EXPERIMENTOS

3.1 - Experimentos: definições e um método de compará-los.

3.1.1- DEFINIÇÕES: Seja X um espaço topológico e $\mathcal{B}(X)$ a σ-álgebra de Borel de subconjuntos de X. Um experimento α é um conjunto de N medidas de probabilidade $u_1, u_2, u_3, u_4, \ldots, u_N$ definidas sôbre $(X, \mathcal{B}(X))$. A qualquer par (α, A), onde $\alpha = \{u_1, u_2, \ldots, u_N\}$ é um experimento e A um subconjunto convexo, fechado, limitado de R^N, corresponde um problema de decisão do seguinte modo: A é considerado como o espaço das ações e um ponto $x \in X$ é selecionado segundo uma das distribuições u_i, $1 \leq i \leq N$; o estatístico observa x, escolhe uma ação $a = (a_1, a_2, \ldots, a_N) \in A$ e suponhamos que ao adotar a ação $\mathbf{a} = (a_1, a_2, \ldots, a_N)$ êle incorra em uma perda $L(i,a) = a_i$. Uma função de decisão d para (α, A) é uma aplicação $d: X \longrightarrow A$, $\mathcal{B}(X)$-mensurável. A cada ponto amostral $x \in X$ selecionado a função de decisão d associa a ação $d(x) = (a_1(x), a_2(x), \ldots, a_N(x))$. Para toda função de decisão d para (α, A), se a distribuição de probabilidade sôbre $(X, \mathcal{B}(X))$ for u_i, $1 \leq i \leq N$, temos

$$E\left[L(i, d(.))\right] = \int_X L(i, d(x))\, du_i(x) =$$

$$= \int_X L(i, (a_1(x), a_2(x), \ldots, a_N(x)) du_i(x) =$$

$$= \int_X a_i(x) du_i(x) .$$

Seja $\mathcal{D}$ o conjunto de todas as funções de decisão para (α,A) e v a aplicação

$$d \in \mathcal{D} \longrightarrow v(d) = \left(\int_X a_1(x)du_1(x), \ldots, \int_X a_N(x)du_N(x)\right) \in R^N$$

onde $a_i(x) = L(i,d(x))$; $i=1,2,\ldots,N$.

O vetor $v(d)$ se diz o vetor risco associado a d . Seja $R(\alpha,A)$ o conjunto dos vetores riscos associados a todos os elementos de $\mathcal{D}$, isto é, $R(\alpha,A) = v(\mathcal{D})$.

PROPOSIÇÃO 3.1.1 - $R(\alpha,A)$ é um subconjunto convexo, fechado , limitado de R^N e $A \subset R(\alpha,A)$.

Prova: Se $r_1,r_2 \in R(\alpha,A)$ e $\lambda \in [0,1]$, então,

$$r_1 = v(d) = \left(\int_X a_1(x)du_1(x), \ldots, \int_X a_N(x)du_N(x)\right) ,$$

$$r_2 = v(d')= \left(\int_X a_1'(x)du_1(x), \ldots, \int_X a_N'(x)du_N(x)\right),$$

onde

$$d,d' \in \mathcal{D} \quad \text{com} \quad d(x) = (a_1(x),\ldots,a_N(x)),$$
$$d'(x) = (a_1'(x),\ldots,a_N'(x))$$

e

$$\lambda r_1+(1-\lambda)r_2 = \left(\int_X (\lambda a_1(x)+(1-\lambda)a_1'(x))du_1(x), \ldots, \right.$$

$$\left. \int_X (\lambda a_N(x)+(1-\lambda)a_N'(x))du_N(x)\right).$$

Notando, pela convexidade de A , que $d''= \lambda d+(1-\lambda)d'$ é uma função de decisão para (α,A) e que $v(d'') = \lambda r_1+(1-\lambda)r_2$, segue que $R(\alpha,A)$ é convexo. Para a prova de que $R(\alpha,A)$ é fechado ver [3]. Para $i=1,2,\ldots,N$, como A é limitado, existe $M_i > 0$ tal que $|L(i,d(x))| = |L(i,(a_1(x),\ldots,a_N(x)))| =$

$= |a_i(x)| \leq M_i$, para toda função de decisão d para (α,A) ,o que implica

$$\left|\int_X a_i(x)du_i(x)\right| \leq \int_X |a_i(x)|du_i(x) \leq M_i \ .$$

Logo, $R(\alpha,A)$ é limitado. Finalmente, para todo $\mathbf{a} = (a_1,a_2,\ldots,a_N) \in A$ seja d a função de decisão para (α,A) tal que $d(X) = \{a\}$; $a = (a_1,a_2,\ldots,a_N) = v(d)$ implica que $a \in R(\alpha,A)$ e portanto, $A \subset R(\alpha,A)$.
†

Se o espaço X for finito, isto é, $X = \{ x_1,x_2,\ldots,x_n \}$ ao experimento $\alpha = \{ u_1,u_2,\ldots u_N \}$ corresponde uma matriz de Markov $P = (p_{ij})_{1\leq i\leq N \ ; \ 1\leq j\leq n}$, onde $p_{ij} = u_i(\{x_j\})$, $i=1,2,\ldots,N$; $j = 1,2,\ldots,n.$ (*)

Uma função de decisão d para (α,A) é, neste caso, especificada por uma matriz real $D = (d_{ji})_{1\leq j\leq n \ ; \ 1\leq i\leq N}$, cuja j-ésima linha, $j=1,2,\ldots,n$, é um ponto em A especificando a ação a ser adotada, quando o ponto amostral x_j for observado. Se a distribuição de probabilidade for u_i , $1\leq i\leq N$, temos

$$E[L(i,d(.))] = \sum_{j=1}^{n} u_i(\{x_j\})L(i,d(x_j)) =$$

$$= \sum_{j=1}^{n} u_i(\{x_j\})L(i,(d_{j1},d_{j2},\ldots,d_{jN})) =$$

$$= \sum_{j=1}^{n} p_{ij}d_{ji}$$

e portanto, o vetor risco associado a d é o vetor

(*) Uma matriz real $B = (b_{ij})_{i\in I \ , \ j\in J}$ é de Markov se $b_{ij}\geq 0$, para todo $i \in I$, $j \in J$ e $\sum_{j\in J} b_{ij} = 1$, para todo $i \in I$.

$$v(d) = \left(\sum_{j=1}^{n} p_{1j}d_{j1} \ , \sum_{j=1}^{n} p_{2j}d_{j2}, \ldots, \sum_{j=1}^{n} p_{Nj}d_{jN} \right).$$

Notamos que as coordenadas do vetor $v(d)$ são os elementos da diagonal da matriz PD. Estas noções são ilustradas no seguinte exemplo: seja $X = \{x_1,x_2\}$ e consideremos o experimento $\alpha = \{u_1,u_2\}$, onde

$$u_1(\{x_1\}) = \frac{1}{3} \quad , \ u_1(\{x_2\}) = \frac{2}{3}$$

e

$$u_2(\{x_1\}) = \frac{3}{5} \quad , \ u_2(\{x_2\}) = \frac{2}{5} \ .$$

Ao experimento α corresponde a matriz

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

e se $A = \{ (x,y):(x,y) \in R^2 \ \text{e} \ x^2+y^2 \leq 1 \}$, então, uma função de decisão d para (α,A) é especificada pela matriz real

$$D = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

com $(d_{j1})^2 + (d_{j2})^2 \leq 1$, $j = 1,2$. Logo,

$$PD = \begin{bmatrix} \frac{1}{3}d_{11} + \frac{2}{3}d_{21} & \frac{1}{3} d_{12} + \frac{2}{3} d_{22} \\ \frac{3}{5}d_{11} + \frac{2}{5}d_{21} & \frac{3}{5} d_{12} + \frac{2}{5} d_{22} \end{bmatrix}$$

e

$$R(\alpha,A) = \{ (x,y) : x = \frac{1}{3} d_{11} + \frac{2}{3} d_{21} \ , \ y = \frac{3}{5} d_{12} + \frac{2}{5} d_{22} \ \text{e} \ (d_{j1})^2 + (d_{j2})^2 \leq 1, \ j=1,2 \} .$$

Para todo $(a,b) \in A$, considerando $d_{11} = d_{21} = a$ e $d_{12} = d_{22} = b$ tem-se que $(a,b) \in R(\alpha,A)$, o que implica $A \subset R(\alpha,A)$. Notamos, por exemplo, que o ponto $(\frac{5}{9}, \frac{3\sqrt{8}+2\sqrt{5}}{15}) \in R(\alpha,A)$, pois $\frac{5}{9} = \frac{1}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{2}{3}$ e $\frac{3\sqrt{8}+2\sqrt{5}}{15} = \frac{3}{5} \cdot \frac{\sqrt{8}}{3} + \frac{2}{5} \cdot \frac{\sqrt{5}}{3}$, mas não é um ponto de A porque $(\frac{5}{9})^2 + (\frac{3\sqrt{8}+2\sqrt{5}}{15})^2 > 1$; ou seja, $A \subsetneq R(\alpha,A)$.

3.1.2 - Um método de comparar dois experimentos

Sejam $\alpha=\{u_1,u_2,\ldots,u_N\}$ e $\beta=\{v_1,v_2,\ldots,v_N\}$ experimentos. Dizemos que α é mais informativo do que β e indicamos $\alpha \supset \beta$, se para todo subconjunto convexo, fechado, limitado A de R^N tem-se

$$R(\alpha,A) \supset R(\beta,A) .$$

Portanto, $\alpha \supset \beta$ significa que para todo subconjunto A de R^N , nas condições acima, todo vetor risco que se pode obter em (β,A) também se pode obter em (α,A) .

PROPOSIÇÃO 3.1.2 - Suponhamos que os espaços amostrais considerados sejam finitos. Para um dado N

i) o experimento menos informativo é aquêle correspondente a uma matriz de Markov com linhas iguais,

ii) o experimento mais informativo é aquêle correspondente a matriz identidade I_N .

Prova: Seja β o experimento correspondente a matriz de Markov $P = (p_{ij})_{1\leq i\leq N\,;\,1\leq j\leq n}$, onde $(p_{i1},p_{i2},\ldots,p_{in})=(p_{11},p_{12},\ldots,p_{1n})$, $i=1,2,\ldots,N$. Seja A um subconjunto convexo, fechado, limitado de R^N e $D = (d_{ji})_{1\leq j\leq n\,;\,1\leq i\leq N}$ uma matriz real correspondente a uma função de decisão d para (β,A) . Como o vetor -

risco associado a d ,

$$v(d) = \left(\sum_{j=1}^{n} p_{1j} d_{j1} \,,\, \sum_{j=1}^{n} p_{1j} d_{j2}, \ldots, \sum_{j=1}^{n} p_{1j} d_{jN} \right) =$$

$$= p_{11}(d_{11}, d_{12}, \ldots, d_{1N}) + p_{12}(d_{21}, d_{22}, \ldots, d_{2N}) + \ldots +$$

$$+ p_{1n}(d_{n1}, d_{n2}, \ldots, d_{nN}) \quad \text{e} \quad (d_{j1}, d_{j2}, \ldots, d_{jN}) \in A \,,$$

$$j = 1, 2, \ldots, n \,,$$

segue, pela convexidade de A , que $v(d) \in A$. Desde que $A \subset R(\alpha, A)$ qualquer que seja o experimento α (ver prop. 3.1.1), então, $v(d) \in R(\alpha, A)$ e $\alpha \supset \beta$, o que prova i) .

Seja $P = (p_{ij})_{1 \le i \le N \,;\, 1 \le j \le n}$ uma matriz de Markov correspondente a um experimento α qualquer, A um subconjunto convexo, fechado, limitado de R^N e $D = (d_{ji})_{1 \le j \le n \,;\, 1 \le i \le N}$ uma matriz real correspondente a uma função de decisão d para (α, A) . Notando que PD é uma matriz real correspondente a uma função de decisão d' para (γ, A), onde γ é o experimento correspondente a matriz I_N , tem-se $v(d) = v(d')$, o que prova ii) .

†

É imediato que dois experimentos podem não ser comparáveis .

3.2 - Experimento padrão e medida padrão associados a um experimento.

Seja $\alpha = \{u_1, u_2, \ldots, u_N\}$ um experimento qualquer e $Nu_o = \sum_{i=1}^{N} u_i$. Como $u_i << Nu_o$, $i = 1, 2, \ldots, N$, segue, pelo teorema de Radon-Nykodim, que existem funções $p_i : X \longrightarrow R$, $\mathcal{B}(X)$-mensuráveis tais que

a) $p_i \ge 0$ q.c. $[Nu_o]$ e para todo $S \in \mathcal{B}(X)$, $u_i(S) = \int_S p_i(x)(Nu_o)(dx)$, $i=1,2,\ldots,N$. Afirmamos também que

b) $\sum_{i=1}^{N} p_i(.) = 1$ q.c. $[Nu_o]$.

Com efeito, para todo $S \in \mathcal{B}(X)$,

$$\int_S \sum_{i=1}^{N} p_i(x)(Nu_o)(dx) = \sum_{i=1}^{N} \int_S p_i(x)(Nu_o)(dx) = \sum_{i=1}^{N} u_i(S) =$$

$$= (Nu_o)(S) = \int_S (Nu_o)(dx)$$

o que implica b) . Sejam $P = \{ c=(c_1,c_2,\ldots,c_N) \in R^N : c_i \geq 0,$ $i=1,2,\ldots,N$ e $\sum_{i=1}^{N} c_i=1 \}$ e $G = \{ x \in X : p_i(x) \geq 0,\ i=1,2,\ldots,N$ e $\sum_{i=1}^{N} p_i(x) = 1 \}$. Definamos as funções $\overline{p}_i : X \longrightarrow R$, $i = 1,2,\ldots,N$, do seguinte modo :

$$p_1(x) = \begin{cases} p_1(x) & \text{se } x \in G \\ 1 & \text{se } x \in X-G \end{cases}$$

e

$$\overline{p}_i(x) = \begin{cases} p_i(x) & \text{se } x \in G \\ 0 & \text{se } x \in X-G \ ; \quad i = 2,3,\ldots,N \end{cases}$$

Consideremos a seguinte aplicação $\mathcal{B}(X)$-mensurável :

$$x \in X \longrightarrow p(x) = (\overline{p}_1(x) , \overline{p}_2(x),\ldots, \overline{p}_N(x)) \in P$$

e para todo $A \in \mathcal{B}(P)$: σ-álgebra de Borel de subconjuntos de P, definamos

$$(3.2.1) \qquad m_i(A) = u_i([p \in A]) \ ; \ i = 1,2,\ldots,N .$$

DEFINIÇÃO 3.2.1 - O experimento $\alpha^* = \{ m_1,m_2,\ldots,m_N \}$, onde m_i , $i=1,2,\ldots,N$, são medidas de probabilidade sôbre $(P,\mathcal{B}(P))$ definidas em (3.2.1) se diz o experimento padrão associado ao experimento α .

DEFINIÇÃO 3.2.2 - A medida de probabilidade $m_\alpha = \frac{\sum_{i=1}^{N} m_i}{N}$ sôbre $(P, \mathcal{B}(P))$ se diz a medida padrão associada ao experimento α. †

A resultante de m_α é o vetor $(\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N})$. Com efeito, sejam $\{e_1, e_2, \ldots, e_N\}$ a base canônica de R^N e f uma forma-linear qualquer definida em R^N. Então,

$$\int_P f(p) m_\alpha(dp) = \int_X f(p(x))\Big(\sum_{i=1}^{N} u_i \,/\, N\Big)(dx) =$$

$$= \frac{1}{N}\int_X f(p(x))(Nu_o)(dx) =$$

$$= \frac{1}{N}\int_X f\Big(\sum_{i=1}^{N} \overline{p}_i(x) e_i\Big)(Nu_o)(dx) =$$

$$= \frac{1}{N}\int_X \sum_{i=1}^{N} \overline{p}_i(x) f(e_i)(Nu_o)(dx) =$$

$$= \frac{1}{N}\sum_{i=1}^{N} \Big(f(e_i)\int_X \overline{p}_i(x)(Nu_o)(dx)\Big) =$$

$$= \frac{1}{N}\sum_{i=1}^{N} f(e_i) u_i(X) = \frac{1}{N}\sum_{i=1}^{N} f(e_i) =$$

$$= f\Big(\sum_{i=1}^{N} \frac{1}{N} e_i\Big) = f\Big(\Big(\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}\Big)\Big)$$

e portanto,

$$r(m_\alpha) = \Big(\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}\Big).$$

A seguinte proposição mostra que todo experimento é tão informativo quanto o experimento padrão que lhe é associado.

PROPOSIÇÃO 3.2.1 - Para todo subconjunto convexo, fechado, limitado A de R^N tem-se $R(\alpha, A) = R(\alpha^*, A)$

<u>Prova</u>: Seja d^* uma função de decisão para (α^*, A). A fun-

ção $d = d^* op$ é uma função de decisão para (α,A) e para $i = 1,2,\ldots N$, $E[L(i,d^*(.))] = \int_P L(i,d^*(q))dm_i(q) = \int_X L(i,d^*(p(x)))du_i(x) = E[L(i,d(.))]$. Então, $v(d^*) = v(d)$ e $R(\alpha^*,A) \subset R(\alpha,A)$. Para mostrar que $R(\alpha,A) \subset R(\alpha^*,A)$ sejam $(X,\mathcal{B}(X), u_o)$ o espaço de probabilidade básico considerado e d uma função de decisão para (α,A).

Consideremos a aplicação

$$q \in P \longrightarrow g^*(q) = (E(L(1,d(.))|p=q),\ E(L(2,d(.))|p=q),\ \ldots,\ E(L(N,d(.))|p=q)) \in R^N$$

e seja $f^*: X \longrightarrow R^N$ definida por

$$f^*(x) = (E(L(1,d(.))|p)(x),\ E(L(2,d(.))|p)(x),\ldots,\ E(L(N,d(.))|p)(x)).$$

Notemos que $f^*(x) = g^*(p(x))$ e, para toda função $g:P \longrightarrow R$, $\mathcal{B}(P)$-mensurável,

$$\text{(a)} \quad \int L(i,d(x))g(p(x))du_o(x) = \int E(L(i,d(.))|p=q)g(q)dm_\alpha(q)$$

para $i=1,2,\ldots,N$, pela definição de esperança condicional. Para $i=1,2,\ldots,N$, escolhendo $g(q) = g((q_1,\ldots,q_N)) = q_i$ temos

$$\text{(b)} \quad \int L(i,d(x))\ \bar{p}_i(x)du_o(x) = \frac{1}{N}\int L(i,d(x))du_i(x) \quad \text{e}$$

$$\text{(c)} \quad \int E(L(i,d(.))|p=q)\ q_i dm_\alpha(q) = \frac{1}{N}\int E(L(i,d(.))|p=q)dm_i(q)$$

Por (a), (b) e (c),

$$\int L(i,d(x))du_i(x) = \int E(L(i,d(.))|p=q)dm_i(q),\ i=1,2,\ldots,N.$$

Se provarmos que $f^*(x) \in A$, então, g^* é uma função de decisão para (α^*,A) e como para $i=1,2,\ldots,N$,

$$E[L(i,g^*(.))] = \int L(i,g^*(q))dm_i(q) =$$

$$= \int E(L(i,d(.))|p=q)dm_i(q) =$$

$$= \int L(i,d(x))du_i(x) = E[\,L(i,d(.))] \ ,$$

segue que o vetor risco associado a g^* é igual ao vetor risco associado a d, o que implica $R(\alpha,A) \subset R(\alpha^*,A)$. Vamos provar que $f^*(x) \in A$, para todo $x \in X$, por absurdo. Supondo que exista $x \in X$ t.q. $f^*(x) \notin A$, existe uma forma linear u definida em R^N t.q. $u(a) \leq 0$, para todo $a \in A$, e $u_o(\{x : x \in X \text{ e } u(f^*(x)) > 0\}) > 0$. Desde que $d(x) \in A$, para todo $x \in X$, então, $u(d(x)) \leq 0$ e

$$\int_S u(d(x))du_o(x) \leq 0 \ ,$$

onde

$$S = \{\, x : x \in X \ \text{ e } \ u(f^*(x)) > 0 \,\} \ .$$

Como

$$\int_S u(f^*(x))du_o(x) > 0$$

temos

$$\int_S u(d(x))du_o(x) \neq \int_S u(f^*(x))du_o(x) \ ,$$

o que é um absurdo, pois

$$\int_S u(d(x))du_o(x) = \int_S u(f^*(x))du_o(x) \ ,$$

pela definição de esperança condicional .$\dagger$

Observamos que, dado um experimento qualquer $\alpha = \{u_1, u_2, \ldots, u_N\}$ as medidas de probabilidade $m_1, m_2, \ldots, m_N$ do experimento padrão associado a α estão completamente determinadas por m_α, pois a densidade de m_i em relação a m_α é Nh_i, onde $h_i : P \longrightarrow R$ definida por

$$h_i(q) = h_i((q_1, \ldots, q_N)) = q_i \ ; \quad i = 1, 2, \ldots, N \ .$$

De fato, para todo $S \in \mathcal{B}(P)$,

$$m_i(S) = u_i([p \in S]) = N \int_{[p\in S]} \bar{p}_i(x)du_o(x) =$$

$$= N \int_{[p\in S]} h_i(p(x))du_o(x) = N \int_S h_i(q)dm_\alpha(q) =$$

$$= \int_S Nh_i(q)dm_\alpha(q) \;\; ; \;\; i=1,2,\ldots,N \;.$$

Já demonstramos que a resultante da medida padrão associada a todo experimento é o vetor $(\frac{1}{N},\ldots,\frac{1}{N})$. Reciprocamente, toda medida de probabilidade m sôbre $(P,\mathcal{B}(P))$ tal que $r(m) = (\frac{1}{N},\ldots,\frac{1}{N})$ é medida padrão do experimento $\gamma = \{m_1,\ldots, m_N\}$, onde

$$m_i(S) = N \int_S h_i(q)dm(q) \;\; ; \;\; i=1,2,\ldots,N \;,$$

com $S \in \mathcal{B}(P)$. Com efeito, pela definição de resultante da medida de probabilidade m ,

$$\int h_i(q)dm(q) = h_i((\frac{1}{N},\ldots,\frac{1}{N})) = \frac{1}{N} \;, \;\; i=1,2,\ldots,N \;;$$

logo, $m_1,m_2,\ldots,m_N$ são medidas de probabilidade sôbre $(P,\mathcal{B}(P))$. Notando que $\gamma^* = \gamma$ e, para todo $S \in \mathcal{B}(P)$,

$$m_\gamma(S) = \frac{1}{N} \sum_{i=1}^{N} m_i(S) = \frac{1}{N} \sum_{i=1}^{N} (N \int_S h_i(q)dm(q)) =$$

$$= \sum_{i=1}^{N} \int_S h_i(q)dm(q) = \int_S \sum_{i=1}^{N} h_i(q)dm(q) = \int_S dm(q) =$$

$$= m(S) \quad \text{temos} \quad m = m_\gamma \;.$$

Dados os experimentos α, β indicaremos o fato $\alpha \supset \beta$ pela notação $m_\alpha \supset m_\beta$.

A seguinte proposição dá uma condição necessária e suficiente -

para que um experimento seja mais informativo do que outro.

PROPOSIÇÃO 3.2.2 - Sejam α, β experimentos. Então, $m_\alpha \supset m_\beta$ se e sòmente se

$$\int f(q)dm_\alpha(q) \geq \int f(q)dm_\beta(q) ,$$

para toda função $f: P \longrightarrow R$, convexa, contínua.

<u>Prova</u>: Seja A a envoltória convexa do conjunto finito $\{a_1, a_2, \ldots, a_K\}$, onde $a_i = (a_{i1}, a_{i2}, \ldots, a_{iN})$; $i = 1,2,\ldots,K$. Consideremos as funções ℓ_i , $i=1,2,\ldots,K$; ℓ de P em R definidas por

$$\ell_i(q) = \ell_i((q_1,\ldots,q_N)) = \sum_{j=1}^{N} a_{ij}q_j = a_i.q \text{ (.:produto escalar) ; } i=1,2,\ldots,K ,$$

$$\ell(q) = \min_{1 \leq i \leq K} \ell_i(q)$$

e $g : P \longrightarrow A$ definida por

$$g(q) = a_i \quad \text{quando} \quad \ell_j(q) > \ell(q) \text{ , } j<i \text{ e } \ell_i(q) = \ell(q) .$$

A função g é uma função de decisão para (γ, A) , para todo experimento padrão γ e, para qualquer função de decisão g^* para (γ, A) , vale

$$(1) \quad \sum_{j=1}^{N} L(j, g^*(q))q_j \geq \sum_{j=1}^{N} L(j, g(q))q_j$$

para todo $q = (q_1,\ldots,q_N) \in P$. De fato, se $q \in P$ então, $g^*(q) = \sum_{j=1}^{K} \lambda_j a_j$ onde $\lambda_j \geq 0$; $j=1,2,\ldots,K$ e $\sum_{j=1}^{K} \lambda_j = 1$, pois $g^*(q) \in A$. Como $g(q).q = a_r.q = \ell_r(q) = \ell(q) = \min_{j \leq i \leq K} \ell_i(q) \leq \ell_i(q) = a_i.q$; $i=1,2,\ldots,K$, segue que

$$\sum_{j=1}^{N} L(j, g^*(q))q_j = g^*(q).q = (\sum_{j=1}^{K} \lambda_j a_j).q = \sum_{j=1}^{K} \lambda_j a_j.q \geq$$

$$\geq \sum_{j=1}^{K} \lambda_j g(q).q = \left(\sum_{j=1}^{K} \lambda_j\right) g(q).q =$$

$$= g(q).q = \sum_{j=1}^{N} L(j,g(q))q_j \quad .$$

Por (1), se $r^* = (r_1^*, r_2^*, \ldots, r_N^*)$ é o vetor risco associado a g^* e $r = (r_1, r_2, \ldots, r_N)$ é o vetor risco associado a g, temos

$$\sum_{j=1}^{N} r_j^* = N \sum_{j=1}^{N} \int L(j,g^*(q))h_j(q)dm_\gamma \geq$$

$$\geq N \sum_{j=1}^{N} \int L(j,g(q))h_j(q)dm_\gamma = \sum_{j=1}^{N} r_j = N \int \ell(q)dm_\gamma ,$$

ou seja,

$$(2) \qquad \min_{r \in R(\gamma,A)} \sum_{j=1}^{N} r_j = N \int \ell(q)dm_\gamma .$$

Em [4], teorema 2, pg. 94 é demonstrado o seguinte resultado: "Dados dois experimentos α, β, $\alpha \supset \beta$ se e sòmente se, para todo A:convexo, fechado, limitado de R^N,

$$(3) \qquad \min_{r \in R(\alpha,A)} \sum_{i=1}^{N} r_i \leq \min_{r \in R(\beta,A)} \sum_{i=1}^{N} r_i ."$$

De (2) segue que, para experimentos padrões γ, θ com medidas padrões m_γ, m_θ a condição (3) vale, para todo A: envoltória convexa de um conjunto finito se e sòmente se $\int \ell(q)dm_\gamma \leq$ $\leq \int \ell(q)dm_\theta$ para toda ℓ que é o mínimo de um número finito de funções lineares, isto é, se e sòmente se $\int h(q)dm_\gamma \geq \int h(q)dm_\theta$ para toda h, que é o máximo de um número finito de funções lineares. Como a condição (3) vale para todo A: envoltória convexa de um conjunto finito, então, (3) vale para todo A: convexo, fechado, limitado de R^N e como $\int h(q)dm_\gamma \geq \int h(q)dm_\theta$ para toda h, que é o máximo de um número finito de funções lineares implica que $\int g(q)dm_\gamma \geq \int g(q)dm_\theta$, para toda função convexa, -

contínua g, a proposição está demonstrada. $\dagger$

PROPOSIÇÃO 3.2.3 - Sejam $\alpha = \{u_1, u_2, \ldots, u_N\}$ e $\beta = \{v_1, \ldots, v_N\}$ experimentos; $m_\alpha = m_\beta$ se e sòmente se $R(\alpha, A) = R(\beta, A)$, para - todo subconjunto convexo, fechado, limitado A de R^N.

Prova: $\Longrightarrow$

Sejam $\alpha^* = \{m_1, m_2, \ldots, m_N\}$ e $\beta^* = \{M_1, M_2, \ldots, M_N\}$ os experimentos padrões associados a α e β, respectivamente. Se d é uma função de decisão para (α^*, A) $((\beta^*, A))$, então, d é uma função de decisão para (β^*, A) $((\alpha^*, A))$ e para $i = 1, 2, \ldots, N$,

$$\int L(i, d(q))dm_i(q) = N \int L(i, d(q))h_i(q)dm_\alpha(q) =$$

$$= N \int L(i, d(q))h_i(q)dm_\beta =$$

$$= \int L(i, d(q))dM_i(q) \; ,$$

o que implica

$$R(\alpha^*, A) = R(\beta^*, A) \; .$$

Notando que $R(\alpha, A) = R(\alpha^*, A)$ e $R(\beta, A) = R(\beta^*, A)$ (ver proposição 3.2.1), segue que $R(\alpha, A) = R(\beta, A)$.

$\Longleftarrow$

Pela proposição anterior,

$$\int \phi(q)dm_\alpha(q) = \int \phi(q)dm_\beta(q)$$

para toda função ϕ convexa, contínua sôbre P. Observando - que o conjunto das funções convexas, contínuas sôbre P é total em $C(P)$: conjunto das funções contínuas a valores reais - definidas em P (1), segue que

$$\int f(q)dm_\alpha(q) = \int f(q)dm_\beta(q)$$

para toda $f \in C(P)$. Logo, $m_\alpha = m_\beta$.

PROPOSIÇÃO 3.2.4 - Sejam α, β experimentos com $N = 2$. $m_\alpha \supset m_\beta$ se e sòmente se $\int_0^y F_{m_\alpha}(x)dx \geq \int_0^y F_{m_\beta}(x)dx$, para todo y, $0 \leq y \leq 1$, onde

$$F_{m_\alpha}(x) = m_\alpha(\ \{p=(p_1,p_2) : p \in P \ \text{e} \ 0 \leq P_1 \leq x\}\ ) ,$$

$$F_{m_\beta}(x) = m_\beta(\ \{p=(p_1,p_2) : p \in P \ \text{e} \ 0 \leq p_1 \leq x\}\ ) .$$

Prova: Consideremos a medida de probabilidade μ definida sôbre ($[0,1]$, $\mathcal{B}$ ($[0,1]$)) por $\mu(B) = m_\alpha$ ($[h_1 \in B]$) , para todo $B \in \mathcal{B}([0,1])$, onde h_1 é a projeção de P sôbre $[0,1]$ definida por $h_1(p) = h_1((p_1,p_2)) = p_1$. A função

$$f(y,u) = \begin{cases} y-u , & 0 \leq u \leq y \leq 1 \\ 0 , & y \leq u \leq 1 \end{cases}$$

é, para cada y , uma função convexa, contínua de u . Logo, a função $\phi_y : P \longrightarrow R$ definida por $\phi_y(p) = \phi_y((p_1,p_2)) = f(y,p_1)$ é,para cada y , convexa, contínua e vale

$$\text{(i)} \int_P \phi_y(p)dm_\alpha(p) = \int_0^y F_{m_\alpha}(x)dx .$$

Com efeito,

$$\int_P \phi_y(p)dm_\alpha(p) = \int_0^y (y-x)d\mu(x) = \int_0^y yd\mu(x) - \int_0^y xd\mu(x) =$$

$$= y\mu([0,y]) - \int_0^y xd\mu(x)$$

e

$$\int_0^y F_{m_\alpha}(x)dx = xF_{m_\alpha}(x)\ \Big|_0^y - \int_0^y xdF_{m_\alpha}(x) = yF_{m_\alpha}(y) -$$

$$- \int_0^y xdF_{m_\alpha}(x) = y\mu([0,y]) - \int_0^y xd\mu(x) ,$$

o que prova (i) .

De modo análogo

$$\int_P \phi_y(p)dm_\beta(p) = \int_o^y F_{m_\beta}(x)dx .$$

Então, se $m_\alpha \supset m_\beta$, pela proposição (3.2.2) ,

$$\int_o^y F_{m_\alpha}(x)dx \geq \int_o^y F_{m_\beta}(x)dx ,$$

para cada y , $0 \leq y \leq 1$.

Reciprocamente, suponhamos que

$$\int_o^y F_{m_\alpha}(x)dx \geq \int_o^y F_{m_\beta}(x)dx ,$$

para todo y, $0 \leq y \leq 1$. Notando que toda função ϕ convexa, contínua sôbre P pode ser uniformemente aproximada por funções da forma

$$\sum_{i=1}^{K} a_i\phi_{y_i} + \ell ,$$

onde $a_i \geq 0$ e ℓ é linear e que

$$\int_P \ell(p)dm_\alpha(p) = \ell((\frac{1}{2},\frac{1}{2})) = \int_P \ell(p)dm_\beta(p)$$

(A resultante de m_α é o vetor $(\frac{1}{2},\frac{1}{2}))$, temos

$$\int_P \phi(p)dm_\alpha(p) \geq \int_P \phi(p)dm_\beta(p)$$

o que implica, pela proposição (3.2.2) , $m_\alpha \supset m_\beta$. †

Utilizando as mesmas notações da proposição (3.2.4) , indiquemos por

$$c_{m_\alpha}(y) = \int_o^y F_{m_\alpha}(x)dx , \quad 0 \leq y \leq 1 .$$

PROPOSIÇÃO 3.2.5 - A função c_{m_α} é não decrescente, $c_{m_\alpha}(0) = 0$, $c_{m_\alpha}(1) = \frac{1}{2}$ e $m_\alpha \supset m_\beta$ se e sòmente se $c_{m_\alpha}(y) \geq c_{m_\beta}(y)$, para todo y, $0 \leq y \leq 1$.

Prova: F_{m_α} é não decrescente; logo, c_{m_α} é não decrescente. É imediato que $c_{m_\alpha}(0) = 0$ e, pela proposição anterior, $m_\alpha \supset m_\beta$ se e sòmente se $c_{m_\alpha}(y) \geq c_{m_\beta}(y)$, $0 \leq y \leq 1$.
Finalmente,

$$c_{m_\alpha}(1) = \int_0^1 F_{m_\alpha}(x)dx = xF_{m_\alpha}(x)\Big|_0^1 - \int_0^1 x dF_{m_\alpha}(x) =$$

$$= F_{m_\alpha}(1) - \int_0^1 x dF_{m_\alpha}(x) = m_\alpha(P) - \int h_1(p)dm_\alpha(p) =$$

$$= 1 - h_1\left(\left(\frac{1}{2}, \frac{1}{2}\right)\right) = 1 - \frac{1}{2} = \frac{1}{2}$$

(h_1 é linear e $r(m_\alpha) = \left(\frac{1}{2}, \frac{1}{2}\right)$). †

3.3. Um método introduzido por David Blackwell para comparar experimentos.

3.3.1 - Suficiência: um método de comparar experimentos

Sejam $\alpha = \{u_1, u_2, \ldots, u_N\}$, $\beta = \{v_1, \ldots, v_N\}$ experimentos, com u_i; $i=1,2,\ldots,N$, medidas de probabilidade definidas sôbre $(X, \mathcal{B}(X))$ e v_i; $i=1,2,\ldots,N$, medidas de probabilidade - definidas sôbre $(Y, \mathcal{B}(Y))$. É devido a Blackwell a seguinte

DEFINIÇÃO 3.3.1 - O experimento α é suficiente para o experimento β e indicamos $\alpha \succ \beta$, se existe um kernel de Markov T de X em Y tal que $Tu_i = v_i$; $i=1,2,\ldots,N$.

Então, $\alpha \succ \beta$ significa que, se um ponto amostral $x \in X$ é observado segundo uma das distribuições u_i; $1 \leq i \leq N$, podemos selecionar um ponto $y \in Y$ segundo a distribuição $T(x,.)$, obtendo um resultado equivalente ao resultado de selecionar y, - segundo a distribuição v_i.

A próxima proposição relaciona o conceito de suficiência com as medidas padrões associadas aos experimentos

PROPOSIÇÃO 3.3.1 - $\alpha > \beta$ se e sòmente se existe uma dilatação T sôbre P tal que $Tm_\beta = m_\alpha$.

Ver demonstração desta proposição em [4], pg. 97, teorema 6.

Observação: Quando os espaços amostrais considerados forem finitos, os experimentos serão caracterizados por matrizes de Markov, conforme foi visto. Neste caso, se P e Q são matrizes de Markov correspondentes, respectivamente, a dois experimentos α e β, com o mesmo N, indicaremos por $P > Q$ ($P \supset Q$) o fato $\alpha > \beta$ ($\alpha \supset \beta$). Então, se P é uma matriz de Markov Nxn e Q é uma matriz de Markov Nxm, $P > Q$ se e sòmente se existe uma matriz de Markov nxm, M tal que $PM = Q$.

3.3.2 - Equivalência entre $>$ e $\supset$.

O nosso propósito neste parágrafo é mostrar que o método $>$, introduzido por Blackwell para comparar experimentos é equivalente ao método $\supset$.

No caso em que os espaços amostrais considerados são finitos o seguinte teorema dá a equivalência entre $>$ e $\supset$:

Teorema 1 : Sejam P, Q matrizes de Markov Nxn, Nxm, respectivamente, associadas a dois experimentos. Então, $P > Q$ se e sòmente se $P \supset Q$. (ver demonstração no livro de Blackwell e Girshick, Theory of Games and Statistical Decisions, pg. 328, - teorema 12.2.2).

Consideremos agora os experimentos $\alpha = \{u_1, u_2, \ldots, u_N\}$, $\beta = \{v_1, v_2, \ldots, v_N\}$ onde u_i, v_i; $1 \leq i \leq N$ são medidas de probabili dade sôbre $(X,\mathcal{B}(X))$, $(Y,\mathcal{B}(Y))$, respectivamente .

Teorema 3.3.1 - $\alpha > \beta$ se e sòmente se $\alpha \supset \beta$.

Prova: $\alpha > \beta \overset{(1)}{\Longleftrightarrow}$ existe uma dilatação T sôbre P tal que $Tm_\beta = m_\alpha \overset{(2)}{\Longleftrightarrow}$ para toda função $f:P \longrightarrow R$, convexa, contínua, $\int f(q)dm_\alpha(q) \geq \int f(q)dm_\beta(q) \underset{(3)}{\Longleftrightarrow} \alpha \supset \beta$

(1) : Proposição (3.3.1) ;

(2) : Teorema (2.3.1) ;

(3) : Proposição (3.2.2) .

3.4 - Experimentos Binomiais

3.4.1 - Introdução

Suponhamos que o espaço amostral X consista de dois elementos, digamos, $X = \{0,1\}$. Neste caso, o experimento $\alpha = \{u_1,u_2,..,u_N\}$ será especificado pelo vetor $(a_1,a_2,...,a_N)$, $0 \leq a_i \leq 1$, onde $a_i = u_i(\{1\})$; $i=1,2,...,N$. Se $N = 2$ o experimento padrão $\alpha^* = \{m_1,m_2\}$ associado a $\alpha = (a_1,a_2)$ é tal que, se

$$P_1 = \frac{a_1}{a_1+a_2} \quad ;$$

$$P_2 = \frac{1-a_1}{2-(a_1+a_2)}$$

e

$$d = \frac{a_1+a_2}{2} \quad ;$$

então,

$$m_1(\{(p_1,1-p_1)\}) = a_1 \;;\; m_1(\{(p_2,1-p_2)\}) = 1-a_1$$

$$m_2(\{(p_1,1-p_1)\}) = a_2 \;;\; m_2(\{(p_2,1-p_2)\}) = 1-a_2 .$$

Logo,

$$m_\alpha(\{(p_1,1-p_1)\}) = \frac{a_1+a_2}{2} = d \;;$$

$$m_\alpha(\{(p_2,1-p_2)\}) = \frac{(1-a_1)+(1-a_2)}{2} = 1-\frac{a_1+a_2}{2} = 1-d\,.$$

Observação: O experimento $\alpha = (1,1)$ é tal que

$$m_1(\{(\tfrac{1}{2},\tfrac{1}{2})\}) = m_2(\{(\tfrac{1}{2},\tfrac{1}{2})\}) = m_\alpha(\{(\tfrac{1}{2},\tfrac{1}{2})\}) = 1.$$

Determinemos, em seguida, as funções F_{m_α} e c_{m_α}.

Se $a_1 \leq a_2$,

$$\begin{aligned} F_{m_\alpha}(x) &= 0 \;, & 0 \leq x < p_1 \\ &= d \;, & p_1 \leq x < p_2 \\ &= 1 \;, & p_2 \leq x \leq 1 \end{aligned}$$

e

$$\begin{aligned} c_{m_\alpha}(x) = \int_0^x F_{m_\alpha}(t)dt &= 0 & , \; 0 \leq x \leq p_1 \\ &= d(x-p_1) & , \; p_1 \leq x \leq p_2 \\ &= d(p_2-p_1)+(x-p_2) & , \; p_2 \leq x \leq 1 \;; \end{aligned}$$

se $a_2 \leq a_1$,

$$\begin{aligned} F_{m_\alpha}(x) &= 0 \;, & 0 \leq x < p_2 \\ &= 1-d \;, & p_2 \leq x < p_1 \\ &= 1 \;, & p_1 \leq x \leq 1 \end{aligned}$$

e

$$\begin{aligned} c_{m_\alpha}(x) = \int_0^x F_{m_\alpha}(t)dt &= 0 & , \; 0 \leq x \leq p_2 \\ &= (1-d)(x-p_2) & , \; p_2 \leq x \leq p_1 \\ &= (1-d)(p_1-p_2)+(x-p_1) & , \; p_1 \leq x \leq 1 \,. \end{aligned}$$

Seja

$$M(p_1,p_2) = \max\,\{p_1,p_2\}$$

e

$$m(p_1,p_2) = \min\,\{p_1,p_2\}\,.$$

Pela forma da função c_{m_α} , se $\alpha = (a_1, a_2)$ e $\beta = (b_1, b_2)$ são experimentos binomiais, então, $c_{m_\alpha} \geq c_{m_\beta}$ se e sòmente se

$$m(p_1(\alpha), p_2(\alpha)) \leq m(p_1(\beta), p_2(\beta))$$

e

$$M(p_1(\alpha), p_2(\alpha)) \geq M(p_1(\beta), p_2(\beta)) .$$

Logo, pela proposição (3.2.5) e pelo teorema (3.3.1)

(1) $\alpha \succ \beta$ se e sòmente se

$$m(p_1(\alpha), p_2(\alpha)) \leq m(p_1(\beta), p_2(\beta))$$

e

$$M(p_1(\alpha), p_2(\alpha)) \geq M(p_1(\beta), p_2(\beta)) .$$

3.4.2 - Exemplo

Consideremos uma população na qual cada indivíduo possue ou não cada uma das características H,S . Suponhamos que as proporções h, s de indivíduos com características H, S sejam conhecidas e que a proporção δ dos indivíduos tendo ambas características H e S é tal que $\delta = hs$ ou $\delta = c \neq hs$. Um experimento que o estatístico pode realizar é, por exemplo, a observação de um indivíduo com a característica H, outro, é a observação de um indivíduo com a característica S . Denotemos por α_H , α_S, α_{NH} e α_{NS} os experimentos binomiais correspondentes as observações de um indivíduo com a característica H, S , não H e não S, respectivamente. Então,

$$\alpha_H = (s, \frac{c}{h}) \quad ; \quad \alpha_S = (h , \frac{c}{s})$$

$$\alpha_{NH} = (s , \frac{s-c}{1-h}) \quad e \quad \alpha_{NS} = (h, \frac{h-c}{1-s}) .$$

Calculando p_1, p_2 para cada um desses quatro experimentos e usando a condição (1) tem-se:

$\alpha_H \succ \alpha_S$ se e sòmente se $h \leq s$;

$\alpha_H \succ \alpha_{NH}$ se e sòmente se $h \leq s$, $h+s \leq 1$;

$\alpha_H \succ \alpha_{NS}$ se e sòmente se $h+s \leq 1$;

$\alpha_S \succ \alpha_{NS}$ se e sòmente se $s \leq h$, $s+h \leq 1$;

$\alpha_S \succ \alpha_{NH}$ se e sòmente se $h+s \leq 1$;

$\alpha_{NS} \succ \alpha_{NH}$ se e sòmente se $h \leq s$.

Supondo

$$h = \min \{h, s, 1-h, 1-s\} ,$$

então,

$$\alpha_H \succ \alpha_S \succ \alpha_{NH} \; ; \; \alpha_H \succ \alpha_{NS} \succ \alpha_{NH} \quad \text{e} \quad \alpha_S \; , \; \alpha_{NS}$$

são não comparáveis até que $h=s$ ou $h=1-s$.

Logo, o procedimento que sempre seleciona a característica mais rara da população é mais informativo do que qualquer outro procedimento da classe considerada. O experimento α_{NH} é o menos informativo dos experimentos α_H , α_S, α_{NH} e α_{NS} enquanto que α_S , α_{NS} são intermediários.

NOTAS E REFERÊNCIAS

(1) Ver [6], volume I, pg. 277, definição 15.8 e volume II, pg. 118, definição 26.5.

O capítulo 3 foi baseado no capítulo XII do livro do Blackwell e Girshick, "Theory of Games and Statistical Decisions" e em [4] .

ERRATA

página	linha	onde está	leia-se
9	-12	$\|\|x\|\| < \delta$	$\|\|x\|\| \leq \delta$
9	-9	seque	segue
12	2	$f \not\geq p$	$f \not\geq p$
12	9	Família	Famílias
25	-11	$x^*_\omega(x)=\tilde{Q}(\tilde{X})$	$x^*(x)=\tilde{Q}(\tilde{X})$
25	-4	$x^* \in E$	$x^* \in E^*$
32	-9	II-4-5	I-4-5
37	-5	$d''= \quad +(1-\lambda)d'$	$d''=\lambda d+(1-\lambda)d'$
42	9	$p_1(x) =$	$\overline{p}_1(x) =$
47	10	$\ell_i(q_1,\ldots,q_N))$	$\ell_i((q_1,\ldots,q_N))$

BIBLIOGRAFIA

[1] BARTLE, R.G., *The Elements of Integration*, John Wiley & Sons (1966).

[2] BILLINGSLEY, P., *Convergence of Probability Measures*, John Wiley & Sons (1968).

[3] BLACKWELL, D., *The Range of Certain Vector Integrals*, Proc. Amer. Math. Soc., Vol. 2 (1951).

[4] BLACKWELL, D., *Comparison of Experiments*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press (1951), pg. 93-102.

[5] BOURBAKI, N., *Éléments de Mathématique, Livre VI, Intégration, Actualités Scientifiques et Industrielles*, 1175, Hermann, Paris, (1952).

[6] CHOQUET, G., *Lectures on Analysis*, W.A. Benjamin (1969).

[7] DIEUDONNÉ, J., *Fundamentos de Análisis Moderno*, Editorial Reverté, S.A. (1966).

[8] LIMA, E.L., *Elementos de Topologia Geral*, Instituto de Matemática Pura e Aplicada, Rio de Janeiro (1969).

[9] MEYER, P.A., *Probabilités et Potentiel*, publicação da Universidade de Strasbourg (1966).

[10] NEVEU, J., *Bases Mathématiques du Calcul des Probabilités*, Masson et Cie (1964).

[11] ROBERTSON, A.P. and ROBERTSON, W.J., *Topological Vector Spaces*, Cambridge University Press (1964).

[12] STRASSEN, V., *The existence of Probability Measures with Given Marginals*, Annals of Math. Stat., Vol. 36, (1965), pg. 423-439.