Probabilidade (PPGECD000000001)

Programa de Pós-Graduação em Estatística e Ciência de Dados (PGECD)

Sessão 2

Raydonal Ospina

Departamento de Estatística
Universidade Federal da Bahia
Salvador/BA

Experimento Aleatório

Um *experimento* é qualquer processo de observação.

- Incerteza ou chance. Implica em falta de capacidade para de predizer seu comportamento em futuras realizações.
- Razões.
  - não sabemos todas as causas envolvidas;
  - não temos dados suficientes sobre as condições iniciais do experimento;
  - as causas podem ser tão complexas que o cálculo do seu efeito combinado não é possível;
  - existe alguma aleatoriedade fundamental no experimento, por exemplo, mecânica quântica;

Os seguintes traços caracterizam um experimento aleatório:

(a) Se for possível repetir as mesmas condições do experimento, os resultados do experimento em diferentes realizações podem ser diferentes.

(b) Pode-se descrever o conjunto de todos os possíveis resultados do experimento.

(c) quando o experimento for repetido um grande número de vezes, uma configuração definida ou regularidade surgirá. É esta regularidade que torna possível construir um modelo probabilístico.

Caracterização

Os resultados de um experimento aleatório são caracterizados pelos seguintes componentes:

1. o conjunto de resultados possíveis Ω.
2. a coleção de conjuntos de resultados de interesse $\mathcal{A}$;
3. um valor numérico P da verosimilhança ou probabilidade de ocorrência de cada um dos conjuntos de resultados de interesse.

Espaço Amostral

O conjunto de possíveis resultados de um experimento aleatório é chamado de *espaço amostral* (SS = Sample Space). Existem quatro pontos que são desejáveis da especificação de um espaço amostral:

SS1. listar os possíveis resultados do experimento;

SS2. fazê-lo sem duplicação;

SS3. fazê-lo em um nível de detalhamento suficiente para os interesses desejados;

SS4. especificar essa lista completamente em um sentido prático, embora usualmente não completa no que se refere a todos os resultados logicamente ou fisicamente possíveis.

(Este conjunto não necessariamente é único e sua determinação depende do interesse do observador ou pessoa que realiza o experimento aleatório).

Espaços amostrais podem ser enumeráveis ou não enumeráveis; se os elementos do espaço amostral podem ser colocados em uma correspondência 1-1 com um subconjunto dos inteiros, o espaço amostral é enumerável.

Exemplo 1

Por exemplo, em uma única jogada de uma moeda poderíamos ter:

- $\Omega_1 = \{cara, coroa\}$; $\Omega_2 = \{cara, coroa, borda\}$; ou $\Omega_3 = \{(x, y) \in \mathbb{R}^2\}$, onde (x, y) são as coordenadas do centro da moeda após parar.

Exemplo 2

Suponhamos que é selecionado ao acaso um estudante de uma sala de aula e é medida a sua altura. O resultado do experimento é a estatura, medida em metros.

$$\Omega_1 = \{x \in \mathbb{R} : x = 1 + a_1 10^{-1} + a_2 10^{-2}, \text{ com } a_1, a_2 \in \mathbb{N}_0 \text{ e } a_1, a_2 \leq 10\},$$

em que $\mathbb{N}_0$ é o conjunto dos números naturais incluindo o zero. Ω_1 indica que a estatura está medida em metros, com estatura mínima de um metro ($a_1 = a_2 = 0$) e estatura máxima de dois metros e dez centímetros ($a_1 = a_2 = 10$). Este espaço amostral pode ser adequado para a população Brasileira sendo que um metro como estatura minima pode ser muito pequena. Em princípio, e de forma conceitual, a estatura é uma variável contínua. Assim, podemos propor como espaço amostral

$$\Omega_2 = [1, \quad 2,1] = \{x \in \mathbb{R} : 1 \leq x \leq 2,1\},$$

que reflete a possibilidade de mensurar a estatura com uma precisão absoluta, com estatura mínima de exatamente um metro e máxima de dois metros e dez centímetros.

Exemplo 3

Outras possibilidades poderiam ser

$$\Omega_3 = (0, \quad 100) = \{x \in \mathbb{R} : 0 < x < 100\},$$

e

$$\Omega_4 = (0, \quad \infty) = \{x > 0\},$$

Podemos pensar em Ω_3 e Ω_4 como exageros para o espaço amostral, não pelo fato da medição ser absoluta mas pela própria escala de variação dos valores. Contudo, estes dois espaços contêm os possíveis resultados de interesse.

Exemplo 4

Observamos a posição de uma partícula que se move aleatoriamente sobre o eixo real. Aqui, $\Omega = \mathbb{R}$.

Eventos e Coleção de Eventos

- o elemento $\omega \in \Omega$ é chamado de *evento elementar*, e todo subconjunto do espaço amostral $A \subset \Omega$ é dito de *evento*.
- Se ao realizarmos um experimento aleatório, o resultado pertence a um dado evento A, dizemos que A *ocorreu*.
- Utilizaremos as operações Booleanas de conjuntos (complementar, união, intersecção, diferença) para expressar eventos combinados de interesse.

Eventos Disjuntos

Os eventos A e B são *disjuntos* ou *mutuamente excludentes* ou *mutuamente exclusivos* se não puderem ocorrer juntos.

Exercício 1

Sejam A, B, e C eventos em um mesmo espaço amostral Ω. Expresse os seguintes eventos em função de A, B, e C e operações Booleanas de conjuntos (uniões, interseções, etc).

(a) Pelo menos um deles ocorre:

(b) Exatamente um deles ocorre:

(c) Apenas A ocorre:

(d) Pelo menos dois ocorrem:

(e) No máximo dois deles ocorrem:

Definição 1 (Partição)

Dado um espaço amostral Ω*, uma partição* $\Pi = \{A_\alpha, \alpha \in \mathcal{I}\}$ *de* Ω *é uma coleção de eventos que satisfaz:*

P1. *Para todo* $\alpha \neq \beta$*,* $A_\alpha \cap A_\beta = \emptyset$*;*

P2. $\cup_{\alpha \in \mathcal{I}} A_\alpha = \Omega$.

Portanto, cada elemento $\omega \in \Omega$ pertence a um, e somente um, dos eventos A_α de uma partição.

Exemplo 5

A coleção de intervalos $\{(n, n+1] : n \in \mathbb{Z}\}$ é uma partição dos números reais $\mathbb{R}$.

Álgebra de Eventos

Por que não analisar todos os subconjuntos de um dado espaço amostral?

- o espaço amostral pode conter um grau de detalhamento superior ao que estamos interessados no momento;
- quereremos associar cada evento A com uma probabilidade numérica $P(A)$, porém nosso conhecimento sobre P pode não estender para todos os subconjuntos de Ω;
- Axiomas de Kolmogorov podem implicar que não exista P definida em todos os subconjuntos de Ω (não enumerável).

Estaremos interessados em uma coleção especial $\mathcal{A}$ de subconjuntos do espaço amostral chamada de uma *álgebra de eventos*.

Álgebra de Eventos

Definition 1

Uma álgebra de eventos $\mathcal{A}$ é uma coleção de subconjuntos do espaço amostral Ω que satisfaz:

1. não é vazia;
2. fechada com respeito a complementos (se $A \in \mathcal{A}$, então $A^c \in \mathcal{A}$);
3. fechada com respeito a uniões finitas (se $A, B \in \mathcal{A}$, então $A \cup B \in \mathcal{A}$).

Observação: Pelas Leis de De Morgan, vemos que $\mathcal{A}$ é fechada com respeito a intersecções finitas também.

Exemplos de Álgebra de Eventos

1. A menor álgebra de eventos é $\mathcal{A} = \{\emptyset, \Omega\}$;
2. A maior álgebra de eventos é o conjunto das partes de Ω;
3. Um exemplo intermediário, temos:

$$\Omega = \{1, 2, 3\}, \mathcal{A} = \{\Omega, \emptyset, \{1\}, \{2, 3\}\}.$$

4. A álgebra de eventos finitos e co-finitos (conjuntos cujo complemento é finito). Seja $\Omega = \mathbb{R}$ e

$$\mathcal{A} = \{A \subseteq \mathbb{R} : A \text{ é finito}\} \cup \{A \subseteq \mathbb{R} : A^c \text{ é finito}\},$$

ou seja, $\mathcal{A}$ consiste dos subconjuntos de $\mathbb{R}$ que ou são finitos ou têm complementos finitos. $\mathcal{A}$ é uma álgebra de eventos.

Propriedades de Álgebra de Eventos

Lema 1

Se $\mathcal{A}$ é uma álgebra, então $\Omega \in \mathcal{A}$

Demonstração.

Como $\mathcal{A}$ é não vazia, seja A um elemento qualquer seu. Pela segunda propriedade de álgebras, temos que $A^c \in \mathcal{A}$, e pela terceira propriedade temos que $\Omega = A \cup A^c \in \mathcal{A}$. □

Teorema 1

Sejam $\mathcal{A}_1$ e $\mathcal{A}_2$ álgebras de subconjuntos de Ω e seja $\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2$ a coleção de subconjuntos comuns as duas álgebras. Então, $\mathcal{A}$ é uma álgebra.

Demonstração.

Como $\mathcal{A}_1$ e $\mathcal{A}_2$ são álgebras, ambos contém Ω. Então, $\Omega \in \mathcal{A}$. Se $A \in \mathcal{A}$, então A está em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$. Logo, A^c está em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$, e portanto na sua intersecção $\mathcal{A}$. Se $A, B \in \mathcal{A}$, então eles estão em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$. Consequentemente, $A \cup B$ está em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$ e, portanto, em $\mathcal{A}$. Como $\mathcal{A}$ satisfaz as três condições da definição de álgebra de eventos, $\mathcal{A}$ é uma álgebra de eventos. □

É fácil ver que a prova do teorema anterior pode ser estendida para o caso de uma intersecção de um número arbitrário de álgebras. O seguinte corolário usa este fato para provar que sempre existe uma menor álgebra contendo uma família qualquer de eventos.

Corolário 1

Existe uma menor (no sentido de inclusão) álgebra contendo qualquer família dada de subconjuntos de Ω.

Demonstração.

Seja $\mathcal{C}$ uma coleção qualquer de subconjuntos de Ω, defina $\mathcal{A}(\mathcal{C})$ como sendo o conjunto que é igual a intercessão de todas as álgebras de eventos que contém $\mathcal{C}$, isto é:

$$\mathcal{A}(\mathcal{C}) = \bigcap_{\mathcal{A} \supseteq \mathcal{C} : \mathcal{A} \text{ é uma álgebra de eventos}} \mathcal{A}.$$

Pelo teorema anterior, $\mathcal{A}(\mathcal{C})$ é uma álgebra de eventos, e consequentemente é a menor álgebra de eventos contendo $\mathcal{C}$. $\mathcal{A}(\mathcal{C})$ é conhecida como a *álgebra de eventos gerada por* $\mathcal{C}$. □

Teorema 2

Se $\mathcal{A}$ é uma álgebra de eventos, então

$$A_i \in \mathcal{A}, i = 1, 2, \ldots, n \Rightarrow \cup_{i=1}^{n} A_i \in \mathcal{A}$$

Demonstração.

Para $n = 1$, o resultado é óbvio. Para $n = 2$, o resultado segue diretamente da terceira propriedade na definição de álgebra de eventos. Vamos agora provar o passo indutivo, suponha que

$$A_i \in \mathcal{A}, i = 1, 2, \ldots, k \Rightarrow \cup_{i=1}^{k} A_i \in \mathcal{A}.$$

Vamos agora provar que o caso $n = k + 1$ é verdadeiro. Suponha que $A_i, i = 1, 2, \ldots, k + 1 \in \mathcal{A}$, então como

$$\cup_{i=1}^{k+1} A_i = (\cup_{i=1}^{k} A_i) \cup A_{k+1},$$

temos que utilizando o caso $n = k$, $\cup_{i=1}^{k} A_i \in \mathcal{A}$. Como $\cup_{i=1}^{k} A_i \in \mathcal{A}$ e $A_{k+1} \in \mathcal{A}$, temos que utilizando o caso $n = 2$, $(\cup_{i=1}^{k} A_i) \cup A_{k+1} \in \mathcal{A}$. $\square$

Construindo uma Álgebra de Eventos

Nota 1

Uma maneira de construir uma álgebra de eventos, é primeiro particionar Ω em um número finito subconjuntos e depois considerar álgebra que consiste dos eventos que são uniões finitas dos subconjuntos da partição.

Exemplo 6

Por exemplo, $\Omega = \{a, b, c, d\}$. Considere a partição, $\{\{a, c\}, \{b, d\}\}$, então considere a coleção de eventos que consiste de uniões finitas dos eventos desta partição: $\mathcal{A} = \{\emptyset, \Omega, \{a, c\}, \{b, d\}\}$. É fácil ver que $\mathcal{A}$ é uma álgebra de eventos.

Definição 2

Dada uma coleção finita eventos $\mathcal{C} = \{A_1, A_2, \ldots, A_n\}$, define-se um **átomo** *de $\mathcal{C}$ como sendo qualquer evento B da seguinte forma: $B = B_1 \cap B_2 \cap \ldots \cap B_n$, onde $B_i = A_i$ ou $B_i = A_i^c$ para $i = 1, 2, \ldots, n$.*

Exercício 2

1. Note que existem no máximo $2^{|\mathcal{C}|}$ átomos diferentes e que eles formam uma partição de Ω (verifique!).
2. Quando $\mathcal{C}$ for uma coleção finita de eventos, um evento pertencerá a $\mathcal{A}(\mathcal{C})$, se e somente se, for igual a uma união finita de átomos de $\mathcal{C}$. Note que $\mathcal{A}(\mathcal{C})$ terá no máximo $2^{2^{|\mathcal{C}|}}$ elementos (verifique!).

Exemplo 7

Se $\Omega = \{a, b, c, d, e, f\}$, encontre a álgebra gerada por $\mathcal{C} = \{\{a, b, d\}, \{b, d, f\}\}$.
Note que Os átomos de $\mathcal{C}$ são $\{\{a\}, \{f\}, \{c, e\}, \{b, d\}\}$. Logo,

$$\mathcal{A}(\mathcal{C}) = \{\emptyset, \Omega, \{a\}, \{f\}, \{c, e\}, \{b, d\}, \{a, f\}, \{a, c, e\},$$
$$\{a, b, d\}, \{c, e, f\}, \{b, d, f\}, \{b, c, d, e\}, \{a, f, c, e\},$$
$$\{a, f, b, d\}, \{a, b, c, d, e\}, \{b, c, e, d, f\}\}.$$

Definição 3 (Função Indicadora)

Seja Ω o espaço amostral e $\mathcal{A}$ uma coleção de subconjuntos de Ω. A função indicadora de um conjunto $A \in \mathcal{A}$ é denotada por $\mathbf{1}_A(x)$ e é definida por

$$\mathbf{1}_A(x) = \begin{cases} 1 & se\ x \in A, \\ 0 & se\ x \notin A. \end{cases} = \begin{cases} 1 & se\ x \in A, \\ 0 & se\ x \in A^c = X - A. \end{cases}$$

Note que podemos determinar A a partir de sua função indicadora:

$$A = \{\omega : \mathbf{1}_A(\omega) = 1\}.$$

Exemplo 8

Se $\mathbf{1}_A(\omega)$ for identicamente igual a 1, ou seja, $\mathbf{1}_A(\omega) = 1, \forall \omega \in \Omega$, então A é igual ao espaço amostral Ω. Se $\mathbf{1}_A(\omega)$ for identicamente igual a 0, então A é igual ao conjunto vazio $\emptyset$. Se $\mathbf{1}_A(\omega)$ for igual a 1 somente quando $\omega = \omega_0$, então A é o evento $\{\omega_0\}$ que contém somente o elemento ω_0.

Note que existe uma correspondência 1-1 entre eventos e suas funções indicadoras:

$$A = B \Leftrightarrow (\forall \omega \in \Omega)\mathbf{1}_A(\omega) = \mathbf{1}_B(\omega).$$

Nota 2

O fato que eventos são iguais se, e somente se, suas funções indicadoras forem idênticas nos permitem explorar a aritmética de funções indicadoras, ou seja, para construir argumentos rigorosos no que se refere a relação entre eventos. nós transformamos proposições sobre eventos em proposições sobre funções indicadoras e podemos então utilizar nossa familiaridade com álgebra para resolver perguntas menos familiares sobre eventos.

Sejam A e B subconjuntos de Ω então:

1. $A \subseteq B \Leftrightarrow \mathbf{1}_A \leq \mathbf{1}_B$
2. $\mathbf{1}_{A^{\complement}} = 1 - \mathbf{1}_A$. (complemento de um conjunto)
3. $\mathbf{1}_{A\cap B} = \min\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A \cdot \mathbf{1}_B$, (interseção de conjuntos)
4. $\mathbf{1}_{A\cup B} = \max\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \cdot \mathbf{1}_B$, (união de conjuntos)
5. $\mathbf{1}_{A\Delta B} = \max\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A + \mathbf{1}_B - 2 \cdot \mathbf{1}_A \cdot \mathbf{1}_B$,(diferença simétrica de conjuntos)
6. $\mathbf{1}_{A-B} = \max\{\mathbf{1}_A - \mathbf{1}_B, 0\} = \mathbf{1}_A \mathbf{1}_{B^c}$
7. Suponhamos que $A_1, \ldots, A_n$ é uma coleção de subconjuntos de $\mathcal{A}$ e seja $\mathbf{1}_n = \{1, 2, 3, ..., n\}$ como o conjunto de índices, então

$$\prod_{k\in\mathbf{1}_n} (1 - \mathbf{1}_{A_k}(\omega)),$$

para todo $\omega \in \Omega$. é claramente um produto de 0's e 1's. Este produto vale 1 precisamente para os $\omega \in \Omega$ que não pertencem a nenhum dos conjuntos A_k e 0 em caso contrário. Isto é,

$$\prod_{k\in\mathbf{1}_n} (1 - \mathbf{1}_{A_k}) = \mathbf{1}_{\Omega - \bigcup_k A_k} = 1 - \mathbf{1}_{\bigcup_k A_k}.$$

Exemplo 9

Utilizando funções indicadoras, verifique que $A \subseteq B \Leftrightarrow B^c \subseteq A^c$.
Solução: Temos que

$$A \subseteq B \Leftrightarrow \mathbf{1}_A \leq \mathbf{1}_B \Leftrightarrow 1 - \mathbf{1}_A \geq 1 - \mathbf{1}_B$$
$$\Leftrightarrow \mathbf{1}_{A^c} \geq \mathbf{1}_{B^c} \Leftrightarrow B^c \subseteq A^c.$$

Exercício 3

Provar as propriedades da função indicadora do slide anterior.

Exemplo 10

Utilizando funções indicadoras, determine se

$$(A - C) \cup (B - C) = (A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c).$$

(Sugestão: Faça um Diagrama de Venn.)

Solução:

Seja $\omega \in A \cap B \cap C^c$. Então, $\mathbf{1}_A(\omega) = \mathbf{1}_B(\omega) = \mathbf{1}_{C^c}(\omega) = 1$. Portanto, temos

$$\begin{aligned}\mathbf{1}_{(A-C)\cup(B-C)} &= \mathbf{1}_{A-C} + \mathbf{1}_{B-C} - \mathbf{1}_{A-C}\mathbf{1}_{B-C}\\ &= \mathbf{1}_A\mathbf{1}_{C^c} + \mathbf{1}_B\mathbf{1}_{C^c} - \mathbf{1}_A\mathbf{1}_{C^c}\mathbf{1}_B\mathbf{1}_{C^c}.\end{aligned}$$

De onde conclui-se que $\mathbf{1}_{(A-C)\cup(B-C)}(\omega) = 1$. Por outro lado,

$$\begin{aligned}&\mathbf{1}_{(A\cap B^c\cap C^c)\cup(A^c\cap B\cap C^c)} = \mathbf{1}_{(A\cap B^c\cap C^c)}\\ &+\mathbf{1}_{(A^c\cap B\cap C^c)} - \mathbf{1}_{(A\cap B^c\cap C^c)}\mathbf{1}_{(A^c\cap B\cap C^c)}\\ &= \mathbf{1}_A\mathbf{1}_{B^c}\mathbf{1}_{C^c} + \mathbf{1}_{A^c}\mathbf{1}_B\mathbf{1}_{C^c} - \mathbf{1}_A\mathbf{1}_{B^c}\mathbf{1}_{C^c}\mathbf{1}_{A^c}\mathbf{1}_B\mathbf{1}_{C^c}\end{aligned}$$

De onde conclui-se que

$$\mathbf{1}_{(A\cap B^c\cap C^c)\cup(A^c\cap B\cap C^c)}(\omega) = 0.$$

Logo,

$$\mathbf{1}_{(A-C)\cup(B-C)} \neq \mathbf{1}_{(A\cap B^c\cap C^c)\cup(A^c\cap B\cap C^c)},$$

o que implica que

$$(A-C)\cup(B-C) \neq (A\cap B^c\cap C^c)\cup(A^c\cap B\cap C^c).$$

Exemplo 11

As seguintes questões não estão relacionadas umas com as outras.

a. Se $\mathbf{1}_A\mathbf{1}_B$ for identicamente igual a zero, o que sabemos a respeito da relação entre A e B?

b. Se $A \cap B^c = B \cap A^c$, o que sabemos a respeito da relação entre A e B?

c. Se $\mathbf{1}_A^2 + \mathbf{1}_B^2$ for identicamente igual a 1, o que podemos concluir sobre A e B?

d. Se $\mathbf{1}_A - \mathbf{1}_B$ for identicamente igual a 1, o que podemos concluir sobre A e B?

e. Se $A \cap B = B \cup A$, o que podemos concluir sobre A e B?

Exemplo 12 (Definindo funções)

Funções indicadoras servem para ajudar a definir funções. Por exemplo,

$$f(x) = \begin{cases} 0\,, & x < -1 \\ 1 + x\,, & -1 \leq x < 0 \\ 1 - x\,, & 0 \leq x < 1 \\ 0\,, & x \geq 1 \end{cases}$$

Note que $f(x)$ pode ser escrita como

$$f(x) = (1 + x)\mathbf{1}_{[-1,0)}(x) + (1 - x)\mathbf{1}_{[0,1)}(x),$$

ou de forma mais compacta como

$$f(x) = (1 - |x|)\mathbf{1}_{[-1,1]}(x)$$

Fundamentos de Probabilidade

O que entendemos por probabilidade?

Raciocínio probabilístico.

- comum em nosso dia-dia;
- precisa ser levado seriamente em conta se desejamos tomar decisões racionais.

Em geral, precisamos incorporar conhecimento probabilístico que seja tanto qualitativo como também quantitativo. Antes de focarmos em uma teoria probabilística, vamos explorar o espaço de alternativas. Pode-se classificar as formas de raciocínio probabilístico nas seguintes dimensões:

- grau de precisão – o conceito estrutural
- o significado, ou interpretação a ser dada a probabilidade
- estrutura matemática formal de probabilidade dada por um conjunto de axiomas

Compreensão de fundamentos de probabilidade é importante, pois aplicações de teoria da probabilidade dependem fortemente de seus fundamentos. Por exemplo, os fundamentos influem na escolha dos métodos estatísticos a serem utilizados (Frequentistas, Bayesianos, ...) e na interpretação dos resultados obtidos.

Hierarquia de Conceitos Estruturais de Probabilidade

Possivelmente

"Possivelmente A" é o conceito mais rudimentar e menos preciso, e o usado pelos antigos Gregos para distinguir entre o que era necessário e o que era contingente. Existe um número de conceitos de possibilidade que incluem os seguintes:

possibilidade lógica, no sentido que não se contradiz logicamente;

possibilidade epistêmica, segundo a qual ocorrência de A não contradiz nosso conhecimento, que inclui, mas estende mais que mera lógica;

possibilidade física, a ocorrência de A é compatível com leis físicas, contudo ela pode ser extremamente improvável — por exemplo, uma moeda parando e ficando equilibrada na borda em uma superfície rígida;

possibilidade prática, a noção do dia-a-dia segundo a qual A é praticamente possível se ele tem pelo menos uma verosimilhança não tão pequena de ocorrer.

Provavelmente

Provavelmente. Provavelmente A é um fortalecimento da noção de possibilidade que significa mais que provável que não.

Probabilidade Comparativa. "A é pelo menos tão provável quanto B". A probabilidade comparativa inclui "provavelmente A" através de "A é pelo menos tão provável quanto A^c". Pode ser relacionada com probabilidade numérica através de $P(A) \geq P(B)$; embora como nos dois exemplos anteriores, probabilidade comparativa não requer nenhum comprometimento com probabilidade numérica.

Probabilidade Intervalar. "A tem probabilidade intervalar, ou probabilidade inferior e superior
$(\underline{P}(A), \overline{P}(A))$". Isto permite um grau de indeterminação variável sem nenhum comprometimento com que exista um "verdadeiro" valor no intervalo.

Probabilidade Numérica. "A probabilidade de A é o número real $P(A)$." Este é o conceito usual e com o qual nos ocuparemos neste curso,porém não é o único conceito utilizado em linguagem ordinária e no raciocínio probabilístico do dia-a-dia. É provável que ela tenha inibido o desenvolvimento de teorias matemáticas apropriadas para outros fenômenos aleatórios.

Interpretações de Probabilidade

Parece não ser possível reduzir probabilidade a outros conceitos; ela é uma noção em si mesma. O melhor que podemos fazer é relacionar probabilidade a outros conceitos através de uma interpretação. Os cinco mais comuns grupos de interpretação são os seguintes:

1. **Lógica:** grau de confirmação da hipótese de uma proposição que "A ocorre" dada uma evidência através da proposição que "B ocorreu". Esta interpretação está ligada a um sistema lógico formal e não, digamos, ao mundo físico. Ela é usada para tornar o raciocínio indutivo quantitativo. Quando as evidências ou premissas são insuficientes para deduzir logicamente a hipótese ou conclusão, podemos ainda medir quantitativamente o grau de suporte que uma evidência da a uma hipótese através de probabilidade lógica.
2. **Subjetiva:** se refere ao grau de crença pessoal na ocorrência do evento A e é medida através da interpretação comportamental de disposição a apostar ou agir.

Interpretações de Probabilidade

3. **Frequentista:** se refere ao limite da frequência relativa de ocorrência do evento A em repetidas realizações não relacionadas do experimento aleatório $\mathcal{E}$. Note que limites de frequência relativas são uma idealização, pois não se pode realizar infinitas realizações de um experimento.
4. **Propensidade:** tendência, propensidade, ou disposição para um evento A ocorrer. Por exemplo, considerações de simetria, podem levar a conclusão que um dado tem a mesma propensão ou tendência a cair em qualquer uma de suas faces.
5. **Clássica:** baseada em uma enumeração de *casos igualmente prováveis*.

Frequências Relativas

Resta-nos discutir o terceiro elemento para modelagem do raciocínio probabilístico, a associação de uma medida numérica a eventos que representam a verosimilhança com que eles ocorrem. As propriedades desta associação são motivadas em grande parte pelas propriedades de frequência relativas.
Considere uma coleção de experimentos aleatórios $\mathcal{E}_i$ que possuem a mesma álgebra de eventos $\mathcal{A}$ e tem resultados individuais não necessariamente numéricos $\{\omega_i\}$.
Seja $X(\omega)$ uma função real dos resultados, com $X_i = X(\omega_i)$ sendo o valor associado com o resultado ω_i do i-ésimo experimento.

Seja $Av_nX = \frac{1}{n}\sum_{i=1}^{n} X_i$ (Av_n de Average (média) que depende de n) a média dos resultados dos n primeiros experimentos. Por simplicidade matemática, assumiremos que a função X é escolhida de uma família $\mathcal{F}$ de funções que podem assumir apenas um **número finito** de valores numéricos. Fixando uma dada sequência de resultados $\{\omega_i\}$, é fácil verificar as seguintes propriedades de Av_n:

Av0. $Av_n : \mathcal{F} \to \mathbb{R}$.

Av1. Se para todo ω, $X(\omega) \geq 0$, então $Av_n \geq 0$.

Av2. Se X é uma função constante, então $Av_nX = X$.

Av3. Para todo $X, Y \in \mathcal{F}$, para todo $\alpha, \beta \in \mathbb{R}$,

$$Av_n(\alpha X + \beta Y) = \alpha Av_nX + \beta Av_nY.$$

Em particular, se estamos interessados em um dado evento A e escolhemos $X(\omega) = \mathbf{1}_A(\omega)$, uma função binária, então a média é conhecida como a frequência relativa de A.

Definição 4 (Frequências Relativas)

A frequência relativa de um evento A, determinada pelos resultados $\{\omega_1, \ldots, \omega_n\}$ *de n experimentos aleatórios, é*

$$r_n(A) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_A(\omega_i) = \frac{N_n(A)}{n}.$$

Exemplo 13

Suponha que lança-se um dado 10 vezes e obtém-se a seguinte sequência de resultados: $\{1, 2, 2, 6, 5, 4, 4, 4, 6, 1\}$. A frequência relativa do evento $A = \{2, 4\}$ é igual a $r_{10}(A) = 5/10$, a frequência relativa do evento $B = \{3, 5\}$ é igual a $r_{10}(B) = 1/10$ e a frequência relativa de $A \cup B$ é igual a $r_{10}(A \cup B) = 6/10$.

Propriedades

Propriedades chaves da frequência relativa são:

FR0. $r_n : \mathcal{A} \to \mathbb{R}$.

FR1. $r_n(A) \geq 0$.

FR2. $r_n(\Omega) = 1$.

FR3. Se A e B são disjuntos, então $r_n(A \cup B) = r_n(A) + r_n(B)$.

FR4. Se $A_1, A_2, \cdots A_n, \cdots$ é uma sequência de eventos disjuntos dois a dois, então $r_n(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} r_n(A_i)$.

Sejam A e B dois eventos do espaço amostral associado Ω. Sejam $r_n(A)$ e $r_n(B)$ as frequências relativas de A e B respectivamente. Então,

1. $0 \leq r_n(A) \leq 1$, isto é, a frequência relativa do evento A é um número que varia entre 0 e 1.
2. $r_n(A) = 1$, se e somente se, A ocorre em todas as n repetições do experimento aleatório.
3. $r_n(A) = 0$, se e somente se, A nunca ocorre nas n repetições do experimento aleatório.

Note que, podemos expressar Av_n em termos de r_n. De fato, dada uma função X que assume valores no conjunto finito $\{x_1, x_2, \ldots, x_k\}$, considere os k eventos $\{A_i = \{\omega : X(\omega) = x_i\}, i = 1, 2, \ldots, k\}$. Podemos rearranjar os termos em Av_nX e reescrevê-la da seguinte forma:

$$Av_nX = \sum_{i=1}^{k} x_i r_n(A_i) = \sum_{i=1}^{k} x_i r_n(X = x_i).$$

Em particular, se para cada i, temos convergência da sequência $r_1(X = x_i), r_2(X = x_i), \ldots, r_n(X = x_i)$ para um limite p_i, então também temos convergência da média Av_nX,

$$\lim_n Av_nX = \sum_{i=1}^{n} x_i p_i.$$

Este limite das médias, quando existe, serve como interpretação para o conceito essencial de esperança ou média de uma quantidade aleatória numérica X.

Nós prosseguiremos como se existisse alguma base empírica ou metafísica que garanta que $r_n(A) \to P(A)$ (mesmo útil está é uma das principais críticas ao uso r_n), embora que o sentido de convergência quando n cresce só será explicado pela Lei dos Grandes Números. Esta tendência da frequência relativa de estabilizar em um certo valor é conhecida como *regularidade estatística*. Deste modo, P herdará propriedades da frequência relativa r_n.