MAT 461 – Tópicos de Matemática II
Aula 4: Resumo de Probabilidade

Edson de Faria

Departamento de Matemática
IME-USP

19 de Agosto, 2013

Esta aula está, como de praxe, dividida em duas partes:

- Na primeira parte, examinaremos alguns exemplos do uso de probabilidade condicional, introduzido na aula anterior. Também apresentaremos a importantíssima fórmula de Bayes, bem como o conceito de independência, que ocupa um papel central na teoria de probabilidades.
- Na segunda parte, introduziremos o conceito de variável aleatória, e examinaremos alguns exemplos de aplicações. Também introduziremos a noção de distribuição de uma variável aleatória, bem como o conceito de esperança ou valor esperado de uma variável aleatória.

- Suponhamos que dispomos inicialmente de uma distribuição de probabilidades para os eventos de um espaço amostral dado.
- Como vimos na aula anterior, o conceito de probabilidade condicional é a chave para respondermos de maneira quantitativa à seguinte pergunta:
- Como a informação de que um determinado evento E ocorreu afeta os valores das probabilidades dos demais eventos do espaço amostral?
- Nesta aula, continuaremos a examinar a noção de probabilidade condicional mediante exemplos, e exploraremos algumas de suas propriedades, em especial a chamada fórmula de Bayes.

Exemplo 1

Exemplo

De acordo com dados do censo americano de 1990, sabe-se que de uma população de 100 000 *mulheres americanas, cerca de* 89 800 *vivem pelo menos até os* 60 *anos, enquanto cerca de* 57 060 *vivem pelo menos até os* 80 *anos. Se uma mulher dessa população tem* 60 *anos hoje, qual é a probabilidade de que ela viva pelo menos até os* 80 *anos?*

Solução: Trata-se, é claro, de um problema de probabilidade condicional.

- Considere como espaço amostral do problema a população de 100 000 mulheres.
- Sejam A e B os eventos desse espaço amostral que correspondem aos conjuntos de mulheres que vivem pelo menos até os 60 anos, e pelo menos até os 80 anos, respectivamente.

Exemplo 1

- Observe que B é um sub-evento, isto é um sub-conjunto, de A.
- Note também que $P(A) = 89\,800/100\,000 = 0.8980$ e que $P(B) = 57\,060/100\,000 = 0.5706$.
- Portanto, a probabilidade condicional desejada é:

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B)}{P(A)} = \frac{0.5706}{0.8980} \simeq 0.6354 \; .$$

- Em outras palavras, uma mulher dessa população que tenha 60 anos hoje tem uma probabilidade de 63.54% de viver pelo menos mais 20 anos.

Exemplo 2

Mesmo quando o problema que enfrentamos não envolve diretamente o conceito de probabilidade condicional, o uso de tal conceito pode simplificar as coisas.

Exemplo

Uma urna contém 9 *bolas vermelhas e* 5 *bolas azuis. Exceto por diferenças de cor, tais bolas são idênticas. Duas bolas são retiradas da urna sucessivamente, sem reposição. Qual é a probabilidade de que ambas as bolas sejam vermelhas?*

Solução:

- Já sabemos resolver esse problema de maneira puramente combinatória.
- Vamos mostrar como usar probabilidade condicional para resolvê-lo.

Exemplo 2

- Seja V_1 o evento: "a primeira bola retirada da urna é vermelha". E seja V_2 o evento: "a segunda bola retirada da urna é vermelha".
- O que queremos é calcular $P(V_1 \cap V_2)$.
- Como inicialmente existem $9 + 5 = 14$ bolas na urna, temos claramente $P(V_1) = \frac{9}{14}$.
- Por outro lado, dado que a primeira bola selecionada seja vermelha, restam na urna 8 bolas vermelhas e 5 azuis – um total de 13 bolas – e portanto a probabilidade *condicional* de que a segunda bola retirada da urna seja vermelha, *dado que a primeira seja vermelha*, é igual a $\frac{8}{13}$.
- Ou seja, $P(V_2|V_1) = \frac{8}{13}$.
- Portanto, temos $P(V_1 \cap V_2) = P(V_1)P(V_2|V_1) = \frac{9}{14} \cdot \frac{8}{13} \simeq 0.3956$.

Fórmula de Bayes

- Apresentaremos a seguir uma das ferramentas probabilísticas mais populares entre os estatísticos, a chamada *fórmula de Bayes*.
- O principal motivo pelo qual esta fórmula é tão popular é o fato de permitir ao estatístico *testar hipóteses*. A fórmula de Bayes está na raiz do que se costuma denominar *teoria da confiabilidade*.
- Antes de apresentarmos a fórmula de Bayes propriamente dita, destacamos uma propriedade de aditividade simples que probabilidade condicional possui.
- Observe inicialmente que se $A, B \in \mathcal{A}$ são dois eventos quaisquer, então

$$P(A) \; = \; P(A \cap B) + P(A \cap B^c) \; .$$

Portanto, se $0 < P(B) < 1$, temos

$$P(A) \; = \; P(A|B)P(B) + P(A|B^c)P(B^c) \; . \tag{1}$$

Fórmula de Bayes: Exemplo 1

- Como $P(B^c) = 1 - P(B)$, igualdade acima nos diz que $P(A)$ é a média ponderada das probabilidades condicionais de A em relação a B e B^c, com pesos proporcionais às probabilidades individuais $P(B)$ e $P(B^c)$.
- A igualdade (1) é útil porque nos permite calcular a probabilidade de um evento A condicionando-o à ocorrência ou não-ocorrência de um outro evento conhecido, B. O exemplo a seguir ilustra esta técnica.

Exemplo

A direção de uma escola acredita, de modo um tanto maniqueísta, que seus alunos podem ser divididos em dois grupos, os bons alunos e os maus alunos. As estatísticas da escola mostram que um mau aluno é reprovado em pelo menos uma disciplina ao final do ano letivo com probabilidade 0.7*, enquanto um bom aluno é reprovado com probabilidade* 0.1*. Se* 80% *dos alunos da escola são bons alunos, qual é a probabilidade de que um aluno escolhido ao acaso tenha sido reprovado ao final do ano letivo?*

Solução: Seja R o evento de que o aluno selecionado tenha sido reprovado ao final do ano letivo, e sejam B e M os eventos de que o aluno selecionado é bom ou mau, respectivamente.

- Observe que B e M são eventos complementares.
- Sabemos que $P(R|B) = 0.1$ e $P(R|M) = 0.7$.
- Também sabemos que $P(B) = 0.8$, de modo que $P(M) = P(B^c)$.
- Logo,

$$\begin{aligned} P(R) &= P(R|B)P(B) + P(R|M)P(M) \\ &= 0.1 \times 0.8 + 0.7 \times 0.2 \;=\; 0.22 \;, \end{aligned}$$

ou seja, a probabilidade solicitada é 22%.

Fórmula de Bayes: Exemplo 2

Exemplo

No exemplo acima, suponha que o aluno escolhido ao acaso tenha sido reprovado numa disciplina ao final do ano letivo. Qual é a probabilidade de que ele seja um mau aluno?

Solução: Na notação do exemplo anterior, trata-se agora de calcular a probabilidade condicional de M dado R.

- Temos

$$\begin{aligned} P(M|R) &= \frac{P(MR)}{P(R)} \\ &= \frac{P(R|M)P(M)}{P(R)} \\ &= \frac{0.7 \times 0.2}{0.22} \simeq 0.63 \equiv 63\% \,. \end{aligned}$$

Fórmula de Bayes: Caso geral

- Suponhamos agora que $A, B_1, B_2, \ldots, B_n \in \mathcal{A}$ são eventos tais que $A \subseteq \bigcup_{i=1}^n B_i$ e tais que $B_i \cap B_j \neq \emptyset$ for all $i \neq j$.
- Podemos imaginar que $B_1, B_2, \ldots, B_n$ representam hipóteses estatísticas cuja "plausibilidade" gostaríamos de testar.
- Intuitivamente, se a hipótese B_j torna o evento A bastante provável, então a ocorrência de A torna a hipótese B_j bastante plausível.
- A fórmula de Bayes dá um sentido quantitativamente mais preciso para essa expectativa.

Fórmula de Bayes: Caso geral

Proposição

Dados eventos $A, B_1, B_2, \ldots, B_n \in \mathcal{A}$ *com probabilidades positivas, e satisfazendo as condições acima, temos*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} . \tag{2}$$

Demonstração:

- Como os eventos B_i são *dois a dois disjuntos e cobrem* A, temos

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i) .$$

Fórmula de Bayes: Caso geral

- Por outro lado, por definição de probabilidade condicional, temos

$$P(B_j|A) = \frac{P(AB_j)}{P(A)} ,$$

- Além disso, temos $P(AB_j) = P(A|B_j)P(B_j)$.
- Colocando essas três igualdades juntas obtemos (2).

Fórmula de Bayes: Exemplo 3

Eis aqui uma típica aplicação da fórmula de Bayes.

Exemplo

Temos três cartas essencialmente idênticas, exceto pelo fato de que a primeira tem ambas as faces pintadas de azul, a segunda tem uma face azul e a outra branca, e a terceira tem ambas as faces brancas. As três cartas são colocadas num chapéu, e uma delas é sorteada e colocada sobre uma mesa, exibindo uma face azul. Qual é a probabilidade de que a outra face da carta sorteda seja branca?

Solução:

- Seja A o evento "a face superior da carta sorteada é azul".
- Sejam também C_{aa}, C_{ab}, C_{bb} os eventos correspondentes à carta sorteada ser *azul-azul*, *azul-branca*, *branca-branca*, respectivamente.
- O que o problema pede é a probabilidade condicional $P(C_{ab}|A)$.

Fórmula de Bayes: Exemplo 3

- Temos claramente

$$P(A|C_{aa}) = 1 \ , \quad P(A|C_{ab}) = \frac{1}{2} \ , \quad P(A|C_{bb}) = 0 \ .$$

- Além disso, cada uma das três cartas tem a mesma probabilidade de ser sorteada, de modo que

$$P(C_{aa}) = P(C_{ab}) = P(C_{bb}) = \frac{1}{3} \ .$$

- Portanto, pela fórmula de Bayes, temos

$$P(C_{ab}|A) = \frac{P(A|C_{ab})P(C_{ab})}{P(A|C_{aa})P(C_{aa}) + P(A|C_{ab})P(C_{ab}) + P(A|C_{bb})P(C_{bb})}$$

$$= \frac{\frac{1}{2}\cdot\frac{1}{3}}{1\cdot\frac{1}{3} + \frac{1}{2}\cdot\frac{1}{3} + 0\cdot\frac{1}{3}} = \frac{1}{3} \ .$$

Fórmula de Bayes: Exemplo 4

Um campo em que a noção de probabilidade condicional e a fórmula de Bayes servem como importantes ferramentas é a medicina diagnóstica. Vejamos um exemplo simples.

Exemplo

Um médico suspeita que seu paciente possa ter um tipo particular de câncer, e solicita um exame. Antes do exame, o médico sabe apenas que há 1 *chance em* 1000 *de que uma pessoa tenha esse tipo de câncer. A experiência mostra que o exame dá positivo em* 99% *dos casos em que o câncer está presente; e dá negativo em* 95% *dos casos em que o câncer não está presente. Se o resutado do exame é positivo, qual é a probabilidade de que o paciente tenha a doença?*

Fórmula de Bayes: Exemplo 4

Solução:

- Seja C o evento: "o paciente tem a doença", e seja E o evento: "o resultado do exame é positivo".
- O problema é calcular $P(C|E)$, a probabilidade *condicional* de que o paciente tenha câncer *dado que* o resultado do exame seja positivo.
- Sabemos que $P(C) = 1/1000$, pois a chance de uma pessoa escolhida aleatoriamente ter a doença é de 1 em 1000.
- Além disso, temos $P(E|C) = 0.99$, pois a chance de o resultado do exame ser positivo *dado que* (isto é, na hipótese de que) o paciente tenha a doença é de 99%.
- Também sabemos que $P(E^c|C^c) = 0.95$, pois a chance de o resultado do exame ser negativo *dado que* o paciente não tenha a doença é de 95%.

Fórmula de Bayes: Exemplo 4

- Agora, aplicando a fórmula de Bayes, podemos escrever:

$$P(C|E) = \frac{P(E|C) \cdot P(C)}{P(E|C) \cdot P(C) + P(E|C^c) \cdot P(C^c)} \tag{3}$$

- Nesta fórmula, as probabilidades que ainda não conhecemos são as duas últimas que aparecem no denominador.
- Uma delas é imediata: $P(C^c) = 1 - P(C) = 0.999$, pois C^c é o evento complementar a C (isto é, "o paciente não tem câncer").
- Que dizer de $P(E|C^c)$?
- Uma das propriedades básicas de probabilidade condicional nos diz que

$$P(E|C^c) = 1 - P(E^c|C^c) = 0.05$$

Fórmula de Bayes: Exemplo 4

- Substituindo todas essas informações na fórmula (3), obtemos

$$P(C|E) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.05 \times 0.999} \simeq 0.0194$$

- Ou seja, a probabilidade de que o paciente tenha câncer *dado que o resultado do exame seja positivo* é de apenas 1.94%, o que talvez surpreenda algumas pessoas.
- Observe no entanto que o real significado é este: antes da informação obtida ("o resultado do exame é positivo"), a chance estimada de o paciente ter câncer era de 0.1%, mas depois da informação, essa chance subiu para 1.94%, ou seja, ficou quase 20 vezes maior!

Fórmula de Bayes: Exemplo 5

Uma outra aplicação interessante do conceito de probabilidade condicional diz respeito a questões jurídicas ou criminais. Eis aqui um exemplo simples.

Exemplo

Durante a investigação criminal de um assassinato, o delegado responsável está 60% *convencido de que o principal suspeito é culpado. Num dado momento, uma nova pista surge da análise do crime, mostrando que o assassino tem uma característica: é canhoto. Se apenas* 20% *da população da cidade possui essa característica, e se de fato o principal suspeito é canhoto, quão convencido deve agora estar o delegado de que tal suspeito é de fato o assassino?*

Solução:

- Seja S o evento: "o suspeito é culpado", e seja C o evento "o suspeito é canhoto".
- Temos $P(S) = 0.6$, pois de acordo com o delegado há uma chance de 60% de que o suspeito seja culpado.

Fórmula de Bayes: Exemplo 5

- Além disso, se o suspeito *não é culpado*, então a probabilidade condicional de que ele seja canhoto é simplesmente a probabilidade de que uma pessoa escolhida ao caso dentre os habitantes da cidade seja canhota, que é de 20%. Ou seja, $P(C|S^c) = 0.2$.
- Note também que $P(C|S) = 1$, pois se o suspeito é realmente o assassino, então é certamente canhoto.
- O que o problema pede, no entanto, é $P(S|C)$, a probabilidade condicional de que o suspeito seja culpado dado que ele seja canhoto.
- Para calcularmos $P(S|C)$, utilizamos a fórmula de Bayes:

$$P(S|C) = \frac{P(SC)}{P(C)} = \frac{P(C|S)P(S)}{P(C|S)P(S) + P(C|S^c)P(S^c)}$$

$$\Rightarrow P(S|C) = \frac{1 \times 0.6}{1 \times 0.6 + 0.2 \times 0.4} \simeq 0.882$$

Eventos Independentes

Chegamos agora a um dos conceitos mais fundamentais da teoria de probabilidades, que é o conceito de *independência*.

- Intuitivamente, dois eventos A e B são *independentes* se a ocorrência de um deles não exerce qualquer influência sobre a ocorrência ou não do outro.
- Matematicamente, isto significa que a probabilidade condicional de A dado B é igual à probabilidade de A, ou seja, $P(A|B) = P(A)$.
- Esta definição parece assimétrica em A e B, mas observe que $P(A|B) = P(A)$ se e só se $P(AB) = P(A)P(B)$, o que por sua vez é equivalente a $P(B|A) = P(B)$.

Eventos Independentes

De modo bem mais geral, podemos definir independência de uma coleção qualquer de eventos da seguinte forma.

Definição

Dizemos que os eventos de uma coleção $\mathcal{C} \subseteq \mathcal{A}$ são independentes se, para quaisquer $A_1, A_2, \ldots, A_n \in \mathcal{C}$ temos

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) \;=\; P(A_1)P(A_2)\cdots P(A_n)\ .$$

Segue-se um exemplo concreto do uso do conceito de independência.

Eventos Independentes: Exemplo 1

Exemplo

Suponha que a cada dia, ao final do pregão da bolsa, o preço de fechamento de um ativo seja no mínimo igual ao preço de fechamento do dia anterior com probabilidade 0.55*, e que os resultados dos pregões sejam todos independentes. Qual é a probabilidade de que o preço do ativo caia somente uma vez nos próximos quatro dias?*

Solução:

- Seja A_i o evento de que o preço de fechamento do ativo cai no dia i, para $i = 1, 2, 3, 4$.
- Se E denota o evento de que o preço do ativo cai somente uma vez nos próximos quatro dias, então temos

$$E = A_1 A_2^c A_3^c A_4^c \cup A_1^c A_2 A_3^c A_4^c \cup A_1^c A_2^c A_3 A_4^c \cup A_1^c A_2^c A_3^c A_4 \,. \qquad (4)$$

Eventos Independentes: Exemplo 1

- Note que os eventos no segundo membro são dois a dois disjuntos. Portanto a probabilidade de E é a soma das probabilidades das interseções que aparecem no segundo membro de (4).
- Como A_1, A_2^c, A_3^c, A_4^c são independentes, temos

$$\begin{aligned} P(A_1 A_2^c A_3^c A_4^c) &= P(A_1)P(A_2^c)P(A_3^c)P(A_4^c) \\ &= 0.45 \cdot 0.55 \cdot 0.55 \cdot 0.55 \simeq 0.075 \, . \end{aligned}$$

- O mesmo se aplica para os demais termos em (4), de modo que $P(E) \simeq 4 \times 0.075 = 0.3$. Ou seja, há aproximadamente 30% de chances de que o preço do ativo caia exatamente uma vez nos próximos cinco dias.

Variáveis aleatórias: Definição e exemplos

- Frequentemente, ao realizarmos um experimento probabilístico, não estamos propriamente interessados no evento elementar que ocorre, mas em algum valor numérico resultante dessa ocorrência.
- Por exemplo, considere o experimento que consiste em lançar dois dados honestos – digamos um dado azul e outro vermelho – e observar as faces superiores dos dois dados.
- Suponha que estamos interessados no valor correspondente à soma dos valores dessas duas faces.
- Neste caso, a soma em questão é uma função numérica do evento elementar ocorrido.
- Isto nos conduz à noção central de variável aleatória, que formulamos da seguinte maneira.

Variáveis aleatórias: Definição e exemplos

Definição

Uma variável aleatória *num espaço de probabilidade* $(\Omega, \mathcal{A}, P)$ *é uma função* $X : \Omega \to \mathbb{R}$ *que associa a cada evento elementar* $\omega \in \Omega$ *um número real* $X(\omega)$.

- Por exemplo, no experimento do par de dados que acabamos de descrever, o espaço amostral é o conjunto de todos os pares $\omega = (\omega_1, \omega_2)$ com $\omega_i \in \{1, 2, \ldots, 6\}$.
- A variável aleatória do exemplo é a função $X : \Omega \to \mathbb{R}$ dada por $X(\omega) = \omega_1 + \omega_2$ (a soma dos valores das faces superiores dos dois dados).
- Neste mesmo experimento, no mesmo espaço amostral, podemos realizar várias outras observações. Por exemplo, ao lançarmos os dois dados, podemos anotar apenas o maior dos dois números obtidos nas faces superiores.
- Neste caso, temos uma outra variável aleatória $Y : \Omega \to \mathbb{R}$ que pode ser definida assim: $Y(\omega) = \max\{\omega_1, \omega_2\}$

Variáveis aleatórias: Definição e exemplos

- Vemos assim que, num mesmo espaço de probabilidade, podemos definir muitas variáveis aleatórias.
- Podemos efetuar vários tipos de operação com uma coleção dada de variáveis aleatórias: somá-las, multiplicá-las, etc.
- Em outras palavras, o conjunto de todas as possíveis variáveis aleatórias num espaço de probabilidade é o que chamamos de uma *álgebra*.

Distribuição de uma variável aleatória

- Suponha que X é uma variável aleatória definida sobre um espaço amostral Ω.
- Seja $a \in \mathbb{R}$ um dos valores assumidos por X. Escrevemos $[X = a]$ para denotar o evento "X assume o valor a".
- Em outras palavras:

$$[X = a] \ = \ \{\omega \in \Omega : \ X(\omega) = a\}$$

- Denotamos por $P[X = a]$ a probabilidade deste evento.
- Por exemplo, se Ω é o espaço amostral do experimento anterior (lançamento simultâneo de dois dados honestos), e X é a variável aleatória ali introduzida (a soma dos valores das faces superiores dos dados), então:

$$[X = 1] = \emptyset \ ; \ [X = 2] \ = \ \{(1,1)\} \ ; [X = 4] \ = \ \{(1,3);(2,2);(3,1)\} \ ;$$

Em particular, $P[X = 1] = 0$, $P[X = 2] = 1/36$, $P[X = 4] = 1/12$, etc.

Distribuição de uma variável aleatória

Podemos agora apresentar uma definição mais formal do conceito de distribuição de uma variável aleatória, pelo menos no caso *discreto*.

Definição

Seja $X : \Omega \to \mathbb{R}$ *uma variável aleatória sobre o espaço de probabilidade* $(\Omega, \mathcal{A}, P)$. *Suponha que* X *é* *discreta*, *isto é, assume apenas um conjunto discreto de valores* $x_1, x_2, \ldots, x_i, \ldots$ *O conjunto de valores*

$$p_i \; = \; P[X = x_i] \; , \quad i = 1, 2, \ldots$$

é chamado de *distribuição de probabilidade* *da variável aleatória* X.

- Note que $p_1 + p_2 + \cdots + p_i + \cdots = 1$.

Exemplo de distribuição

Eis a distribuição da variável aleatória X dada anteriormente.

i	x_i	$p_i = P[X = x_i]$
1	2	$\frac{1}{36}$
2	3	$\frac{2}{36}$
3	4	$\frac{3}{36}$
4	5	$\frac{4}{36}$
5	6	$\frac{5}{36}$
6	7	$\frac{6}{36}$
7	8	$\frac{5}{36}$
8	9	$\frac{4}{36}$
9	10	$\frac{3}{36}$
10	11	$\frac{2}{36}$
11	12	$\frac{1}{36}$