UNIVERSIDADE DE SÃO PAULO
Faculdade de Zootecnia e Engenharia de Alimentos

ZAB1111 – Estatística Básica

Prof. César Gonçalves de Lima   cegdlima@usp.br

Aula 8 – PROBABILIDADE (2)

Teorema de Bayes + Variáveis aleatórias discretas

Exemplo 4.2. Consideremos três baias da granja de suínos com as características:

Baia 1: tem 10 leitões, 4 dos quais já foram vacinados

Baia 2: tem 6 leitões, 1 dos quais já foi vacinado

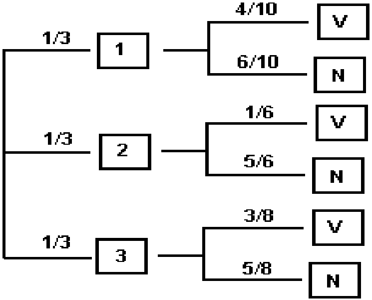
Baia 3: tem 8 leitões, 3 dos quais já foram vacinados

O experimento consiste de duas etapas: sortear uma das três baias e desta baia escolhida sortear um leitão. Pergunta-se:

a) Qual é a probabilidade deste leitão sorteado já estar vacinado?

b) Qual é a probabilidade deste leitão sorteado ser da baia 1, sabendo-se que ele já foi vacinado? E da baia 2? E da baia 3?

$Dica$: Construir um diagrama de árvore.



Evento	Probabilidade
$1 \cap V$	$(1/3)(4/10) = \ 48/360 \cong 0{,}1333$
$1 \cap N$	$(1/3)(6/10) = \ 72/360 \cong 0{,}2000$
$2 \cap V$	$(1/3)(1/6) \ = \ 20/360 \cong 0{,}0556$
$2 \cap N$	$(1/3)(5/6) \ = 100/360 \cong 0{,}2778$
$3 \cap V$	$(1/3)(3/8) \ = \ 45/360 \cong 0{,}1250$
$3 \cap N$	$(1/3)(5/8) \ = \ 75/360 \cong 0{,}2083$

a) Já sabemos que $P(V) \cong 0{,}3139$ e $P(N) = 1 - P(V) \cong 0{,}6861$

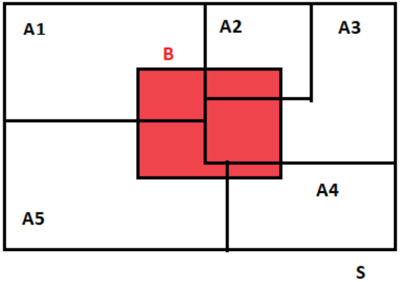
Para resolver o item b) vamos usar o Teorema de Bayes, que será apresentado a seguir.

4.4. TEOREMA DE BAYES

Suponhamos que os eventos $A_1, A_2, \ldots, A_k$ formem uma partição do espaço amostral S, isto é, $A_1 \cup A_2 \cup \ldots \cup A_k = S$ e $A_i \cap A_j = \emptyset$ para $\forall\, i \neq j$. Os eventos A_i são mutuamente exclusivos e exaustivos. Seja B outro evento qualquer.

Então:

$$
\begin{aligned}
B &= B \cap S \\
&= B \cap (A_1 \cup A_2 \cup \ldots \cup A_k) \\
&= (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_k)
\end{aligned}
$$



Note que os eventos $(B \cap A_i)$, para $i = 1,\ 2, \ldots, k$, também são mutuamente exclusivos.

Consequentemente, temos que:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_k)$$

Pelo Teorema da Multiplicação $P(B \cap A_i) = P(A_i)P(B|A_i)$ então:

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_k)P(B|A_k)$$

A probabilidade condicional de um evento A_i ocorrer, dado que o evento B já ocorreu, é calculada pelo **Teorema de Bayes** como:

$$P(A_i|B) = \frac{P(B \cap A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)}$$

ou seja, para calcularmos $P(A_i|B)$ dividimos a probabilidade do caminho $A_i \Rightarrow B$ pela probabilidade do espaço amostral reduzido B, formado por todos os caminhos que levam a este evento.

No Exemplo 4.2 vamos calcular $P(1|V)$, a probabilidade de o leitão sorteado ser da baia 1, sabendo que ele já foi vacinado:

$$P(1|V) = \frac{P(1 \cap V)}{P(V)} = \frac{P(1)P(V|1)}{P(V)} = \frac{(1/3)(4/10)}{113/360} = \frac{48}{113} \cong 0{,}4248$$

Portanto, sabendo que o leitão escolhido está vacinado, a probabilidade dele ter sido escolhido da baia 1 é igual a 0,4248.

De maneira análoga, calculamos também as probabilidades de o leitão ser da baia 2 ou da baia 3, já sabendo que ele está vacinado:

$$P(2|V) = \frac{20}{113} \cong 0{,}1770 \qquad P(3|V) = \frac{45}{113} \cong 0{,}3982$$

Vale lembrar que inicialmente, antes de sabermos que o leitão estava vacinado, $P(1) = P(2) = P(3) = 1/3$.

5. VARIÁVEIS ALEATÓRIAS DISCRETAS

Os recursos disponíveis para o estudo e análise das variáveis quantitativas são mais ricos e numerosos. Inicialmente, vamos estudar algumas características importantes de uma $v.a.$ discreta.

Definição 5.4. Dada uma variável aleatória discreta X, assumindo os valores $x_1, \dots, x_n$, com as respectivas probabilidades $p_1, \dots, p_n$, chamamos de <u>média</u> ou <u>esperança matemática</u> da v.a. X, o valor numérico calculado por:

$$E(X) = \sum_{i=1}^{n} x_i P(\mathrm{X} = x_i) = \sum_{i=1}^{n} x_i p_i = \mu$$

Chamamos de <u>variância</u> da v.a. X o valor calculado pela fórmula:

$$var(X) = \sum_{i=1}^{n} [x_i - E(X)]^2 \, p_i = \sigma^2$$

A variância (σ^2) também pode ser calculada como:

$$var(X) = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

em que $E(X^2) = \sum_{i=1}^{n} x_i^2 P(\mathrm{X} = x_i) = \sum_{i=1}^{n} x_i^2\, p_i$.

O <u>desvio padrão</u> da v.a. X é calculado como a raiz quadra da variância, ou seja:

$$DP(\mathrm{X}) = \sqrt{var(X)} = \sigma$$

A tabela formada pelos valores da v.a. X e suas respectivas probabilidades é chamada **distribuição de probabilidades da v.a. $\boldsymbol{X}$**:

x_i	x_1	x_2	x_3	$\cdots$	x_n
$P(X = x_i)$	p_1	p_2	p_3	$\cdots$	p_n

Note que numa distribuição de probabilidades $\sum_{i=1}^{n} p_i = 1$.

5.2. Algumas propriedades da esperança matemática

Se somarmos uma constante (k) a todos os valores da v.a. X, sua média fica aumentada pela constante, mas a variância e o desvio padrão não serão alterados.

$$i)\ E(X+k)=E(X)+k \qquad ii)\ var(k+X)=var(X)$$

$$iii)\ DP(k+X)=DP(X)$$

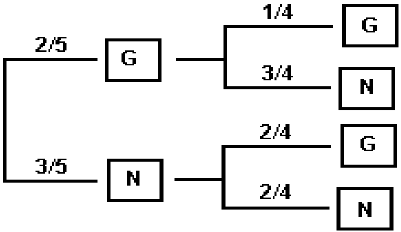
Se multiplicarmos por uma constante (k) todos os valores da v.a. X, sua média e seu desvio padrão ficarão multiplicados pela constante e sua variância, pelo quadrado da constante.

$$iv)\ E(kX)=kE(X) \qquad v)\ var(kX)=k^2var(X)$$

$$vi)\ DP(kX)=k\ DP(X)$$

Exemplo 5.1. Em um piquete com dois bezerros Gir (G) e três Nelores (N) foram sorteados <u>sem reposição</u>, dois animais para serem submetidos a um tratamento com carrapaticida.

- Espaço amostral é S = {GG, GN, NG, NN}
- Variável aleatória: $X =$ "número se bezerros Gir na amostra"



<u>Evento</u>	<u>Probabilidade</u>
G∩G	$(2/5)(1/4) = 1/10$
G∩N	$(2/5)(3/4) = 3/10$
N∩G	$(3/5)(2/4) = 3/10$
N∩N	$(3/5)(2/4) = 3/10$

A distribuição de probabilidades de X: "número se bezerros Gir na amostra" fica:

x	0	1	2
$P(X = x)$	0,3	0,6	0,1

Neste caso:

- $E(X) = (0)(3/10)+(1)(6/10)+(2)(1/10) = 8/10 = 0{,}8$ bezerros
- $E(X^2) = (0^2)(3/10) + (1^2)(6/10) + (2^2)(1/10) = 1{,}0$
- $Var(X) = E(X^2) - [E(X)]^2 = 1{,}0 - (0{,}8)^2 = 0{,}36$ bezerros^2
- $DP(X) = \sqrt{Var(X)} = \sqrt{0{,}36} = 0{,}6$ bezerros.

Exemplo 5.2 [MAGALHÃES, M.A. e LIMA, A.C.P., 2008] Na construção de certo prédio, as fundações devem atingir 15 metros de profundidade e, para cada 5 metros de estacas colocadas, o operador anota se houve alteração no ritmo de perfuração previamente estabelecido. Essa alteração resulta de mudanças (para mais ou para menos) na resistência do subsolo e quando ocorrem, atingem a perfuração de todas as estacas. Se ocorrer alguma alteração aos 5 ou a 10 metros de profundidade, serão necessárias medidas corretivas, que encarecerão a obra.

Com base em avaliações geológicas feitas no terreno admite-se que a probabilidade de ocorrência de alterações é de 0,10 para cada 5 metros.

O custo básico inicial da obra é de 100 UPC (unidade padrão de construção) e será acrescido de $50k$, em que k representa o número de alterações observadas.

Admitindo que as alterações na perfuração ocorrem de forma independente entre cada um dos intervalos de 5 metros, pergunta-se:

- Como se comporta a variável "<u>custo final da fundação</u>"?
- Qual é o custo esperado da fundação?

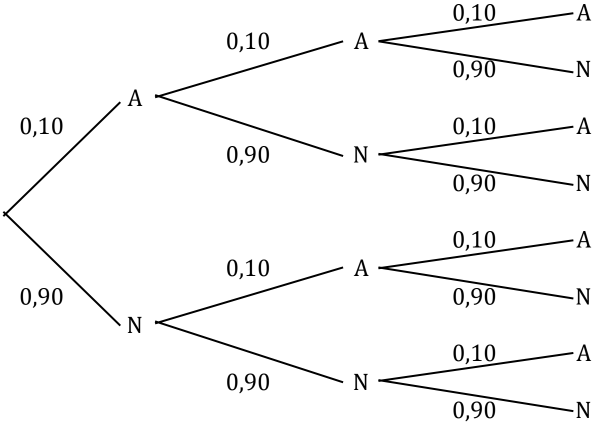
Resolução: Vamos trabalhar com 2 eventos:

A: "ocorrência de alteração em cada intervalo de 5 metros"

N: "não ocorrência de alteração",

Note que $N = A^c$, é o evento complementar de A. Como $P(A) = 0{,}10$ $\Rightarrow P(N) = 1 - P(A) = 0{,}90$.

Como o problema envolve o estudo de ocorrência de alteração em três etapas (início, 5 m e 10 m), vamos organizá-lo num diagrama de árvore de probabilidades.



Evento	Probabilidade
AAA	$0{,}10^3 = 0{,}001$
AAN	$0{,}10^2 \times 0{,}90 = 0{,}009$
ANA	$0{,}10^2 \times 0{,}90 = 0{,}009$
ANN	$0{,}10 \times 0{,}90^2 = 0{,}081$
NAA	$0{,}10^2 \times 0{,}90 = 0{,}009$
NAN	$0{,}10 \times 0{,}90^2 = 0{,}081$
NNA	$0{,}10 \times 0{,}90^2 = 0{,}081$
NNN	$0{,}90^3 = 0{,}729$

Resumindo, temos:

Número de alterações	0	1	2	3
Custo	100	150	200	250
Probabilidade	0,729	0,243	0,027	0,001

A distribuição de probabilidades da variável $C =$ "custo final da obra de fundação" pode ser apresentada como:

c_i	100	150	200	250
$P(C = c_i)$	0,729	0,243	0,027	0,001

O custo esperado (custo médio) da obra de fundação é igual a:

$$E(C) = \sum_i c_i P(C = c_i) = 100(0{,}729) + \ldots + 250(0{,}001) = 115 \text{ UPC´s}$$

Esse valor pode ser útil na elaboração de futuros orçamentos.

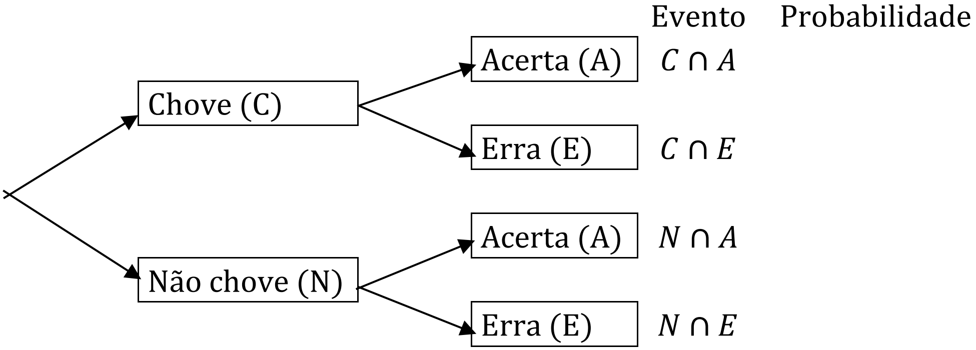
EXERCÍCIOS

1) Sejam A e B dois eventos em um espaço amostral, com $P(A) = 0{,}2$, $P(B) = p$, $P(A \cup B) = 0{,}5$ e $P(A \cap B) = 0{,}1$. Determine o valor de p.

2) Uma moeda é viciada de modo que a probabilidade de sair cara é 4 vezes maior que a de sair coroa. Para dois lançamentos independentes dessa moeda, determinar:

a) O espaço amostral do experimento.
b) A probabilidade de sair somente uma cara nos dois lançamentos.
c) A probabilidade de sair pelo menos uma cara nos dois lançamentos.
d) A probabilidade de sair dois resultados iguais nos dois lançamentos.

3) Em uma região, a probabilidade de chuva em um dia qualquer de primavera é 0,10. Um meteorologista da rádio local acerta suas previsões em 80% dos dias em que chove e em 90% dos dias em que não chove. *Dica*: diagrama de árvore.



Pede-se:

a) Qual é a probabilidade de o meteorologista acertar uma previsão em um dia qualquer de primavera? E de errar?

b) Qual é a probabilidade de ter sido um dia de chuva sabendo-se que a previsão feita pelo meteorologista se confirmou?

4) Historicamente sabe-se que um grande time paulista tem probabilidade 0,55 de vitória em jogos do segundo turno do Campeonato Brasileiro realizados aos sábados. Se este time atuar 4 vezes aos sábados com a mesma escalação. Pede-se:

a) A distribuição de probabilidades da v.a. V: "número de vitórias aos sábados". Calcular $E(V)$ e $DP(V)$.

b) Calcule a probabilidade de que o time vença: i) todas as partidas; ii) mais de duas partidas e iii) no máximo uma partida.