Estatística

Daniel Vital

November 2024

1 Probabilidades Finitas dos Espaços Amostrais Finitos

Seja S um espaço amostral finito $S = \{a_1\ , a_2\ , ...a_n\}$, considere o evento de resultado simples $A = \{a_i\}$. Para cada evento simples $\{a_i\}$, existe uma probabilidade associada P_i, que satisfaz as aseguintes condições:

- $p_i \geq 0 \qquad i = 1, 2, ..., n$
- $p_1 + p_2 + p_3 + ... + p_n = 1$

A probabilidade $P(A)$ de um evento com mais de um elemento (evento composto), é definida pela soma das probabilidades de seus pontos.
Os eventos não tem, necessariamente, a mesma probabilidade.
Não há interseção entre os eventos

Exemplo: Sejam três cavalos de corrida "A", "B" e "C", onde:
"A" tem 2x mais probabilidade de ganhar que "B"
"B" tem 2x mais probabilidade de ganhar que "C"

a) Qual a probabilidade de vitória de cada um?
$P(C) = P \qquad P(B) = 2P \qquad P(A) = 2 \cdot (2P) = 4P$

$P + 2P + 4P = 1$
$7P = 1$
$P = \frac{1}{7}$

$P(C) = \frac{1}{7} \qquad P(B) = \frac{2}{7} \qquad P(A) = \frac{4}{7}$

b) Qual a probabilidade de B **ou** C ganharem?
$P(B \cup P) = P(B) + P(C)$
$P(A \cup B) = \frac{2}{7} + \frac{1}{7} \qquad \therefore \ P(A \cup B) = \frac{3}{7}$

ATENÇÃO: Para calcular a probabilidade nesse caso, é usada soma e não multiplicação, pois são eventos mutuamente exclusivos, não podem ocorrer ao mesmo tempo.

2 Espaços Amostrais Finitos e Equiprováveis

Espaço amostral Equiprovável ou Uniforme: Quando cada ponto amostral tem a mesma probabilidade. Se S possui n pontos, a probabilidade de cada ponto é

$$p = \frac{1}{n}$$

Se um evento A contem r pontos, então:

$$P(A) = r \cdot \frac{1}{n} = \boxed{\frac{r}{n}}$$

Essa fórmula pode, também, ser escrita como:

$$P(A) = \frac{\text{Nº de vezes que “A” pode ocorrer}}{\text{Nº de vezes que o espaço amostral “S” ocorre}}$$

Ou

$$P(A) = \frac{\text{NCF (Nº de casos favoráveis)}}{\text{NTC (Nº total de casos)}} \qquad = \qquad \frac{r}{n}$$

Exemplo: David randomly (“Aleatório” indica que o espaço é equiprovável!) chose one card from a deck of 52 cards. What is the probability of “A”? What is the probability of “B”?
$A = \{\text{A carta é de ouros}\}$
$B = \{\text{A carta é uma figura}\}$

$$P(A) = \frac{\text{Nº de ouros}}{\text{Nº de cartas}} = \frac{13}{52} = \frac{1}{4} \qquad ; \qquad P(B) = \frac{\text{Nº de figuras}}{\text{Nº de cartas}} = \frac{12}{52} = \frac{3}{13}$$

Como se observa, o cálculo de probabilidade está intimamente ligado a um problema de contagem, então usaremos análise combinatória para contabilizar casos totais e favoráveis.

Análise Combinatória

$$C_{n,p} = \binom{n}{p} = \frac{n!}{p!(n-p)!}$$

Exemplo: Quantas comissões de 3 pessoas são possíveis formar a partir de um grupo de 10 pessoas?

$$C_{10,3} = \binom{10}{3} = \frac{10!}{3!(10-3)!} = 120$$

Espaços Amostrais Finitos e equiprováveis - Exercício:

Num lote de 12 peças, 4 são defeituosas. Duas peças são escolhidas aleatoriamente
a) Qual a probabilidade de ambas serem defeituosas?

A={ambas são defeituosas}

A pode ocorrer $\binom{4}{2} = C_{4,2} = \frac{4!}{2! \cdot 2!} = 6$ vezes $\qquad \therefore \boxed{P(A) = \frac{NCF}{NTC} = \frac{6}{66} = \frac{1}{11}}$

S pode ocorrer $\binom{12}{2} = C_{12,2} = \frac{12!}{2! \cdot 10!} = 66$ vezes

b) Probabilidade de ambas serem defeituosas

B={Ambas não possuírem defeitos defeituosas}

B pode ocorrer $\binom{8}{2} = C_{8,2} = \frac{8!}{2! \cdot 6!} = 28$ $\qquad \therefore \boxed{P(B) = \frac{NCF}{NTC} = \frac{28}{66} = \frac{14}{33}}$

S pode ocorrer $\binom{12}{2} = C_{12,2} = \frac{12!}{2! \cdot 10!} = 66$ vezes

c) Probabilidade de ao menos uma ser defeituosa

C={ao menos uma é defeituosa}

C é o complemento de B, ou seja, C=$\overline{B}$, pois:
Se tirarmos os pontos em que ambas as peças não possuem defeitos (B), no espaço amostral só sobrarão os pontos em que uma ou mais são defeituosas (que é equivalente a $\overline{B}$, que satisfaz C)

$\therefore P(C) = 1 - P(B) = 1 - \dfrac{14}{33} = \boxed{\dfrac{19}{22}}$

3 Probabilidade Condicional

A probabilidade de ocorrer um evento A, dado que já ocorreu um evento B
A fórmula é:

$$P(A/B) = \frac{\overbrace{P(A \cap B)}^{\text{Prob. da Interseção}}}{P(B)} = \frac{\frac{\text{NTCF}A\cap B}{\cancel{\text{NTC (S)}}}}{\frac{\text{NTCF}B}{\cancel{\text{NTC (S)}}}} = \frac{\text{NTCF}(A \cap B)}{\text{NTCF}(B)}$$

Ex: Uma urna tem 15 bolas numeradas de 1 a 15. retira-se uma ao acaso, e verifica-se que é maior que 6. Qual a probabilidade de ser um múltiplo de 3?

$B = \{7, 8, 9, 10, 11, 12, 13, 14, 15\}$
$A = \{3, 6, 9, 12, 15\}$
$A \cap B = \{9, 12, 15\}$ $\qquad P(A/B) = \frac{\frac{3}{15}}{\frac{9}{15}} = \frac{3}{9}$

$P(B) = \frac{9}{15}$
$P(A \cap B) = \frac{3}{15}$

Há outro método de se fazer o mesmo cálculo - Reduzindo os espaços amostrais:

Dado que B ocorreu, o espaço amostral se reduz a B

$$B = \{7, 8, 9, 10, 11, 12, 13, 14, 15\}$$

Partindo desse novo espaço amostral, quais números da bola são múltiplos de 3?

$$\text{Múltiplos de 3 no espaço "B"} = \{9, 12, 15\}$$

Portanto, a probabilidade de uma bola ter um número múltiplo de 3, dado que é um número superior a 6, é:

$$P(A/B) = \frac{3}{9} = \frac{\text{NTCF}(A \cap B)}{\text{NTCF}(B)}$$

4 Teorema do Produto

"A probabilidade da ocorrência <u>simultânea</u> de dois eventos A e B pertencentes ao mesmo espaço-amostral é igual ao produto da probabilidade de um deles, pela prob. condicional do outro, dada a ocorrência do primeiro":

$$P(A/B) = \frac{P(A \cap B)}{P(A)} \qquad \longrightarrow \qquad \boxed{P(A \cap B) = P(B)(A/B)}$$

5 Independência Estatística

Um evento A é independente de B se a probabilidade de A for igual a probabilidade condicional de (A/B). Ou seja, a probabilidade de A é a mesma se B ocorrer ou não.
Se A é independente de B, B é independente de A.

$$P(A) = P(A/B) \qquad ; \qquad P(B) = P(B/A)$$

considerando o teorema do produto, podemos afirmar que, se A e B são independentes, então:

$$P(A \cap B) = P(A) \cdot P(B)$$

6 Teorema de Bayes - Bayes' Theorem

Sejam $A1, A2, A3...A_n$, "n" eventos mutuamente exclusivos, tais que $S = \{A_1 \cup A_2 \cup A_3 ... \cup A_n\}$ sejam as probabilidades $P(A_i)$ conhecidas, e "B" um evento da "S" tal que são conhecidas todas as probabilidades condicionais $P(B/A)$:
Então, para cada evento "i", tem se:

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + ...P(A_n) \cdot P(B/A_n)}$$

Exemplo: Escolhendo-se uma urna ao acaso e dela extraiu-se uma bola branca, também ao acaso. Qual a probabilidade da bola branca ter vindo da urna "2"? E da "3"?

cor/urna	U1	U2	U3
Preta	3	4	2
Branca	1	3	3
Vermelha.	5	2	3

Observa-se que são 3 eventos:

- **Evento A:** Escolher a urna ao acaso
- **Evento B:** Escolher uma bola branca aleatoriamente da urna escolhida
- **Evento C:** A bola branca ter vindo da urna 2 ou 3
- Vamos calcular a probabilidade do evento C, dado que B ocorreu. Para B ocorrer, A precisa ter acontecido.

$P(Br/U1) = 1/9 \qquad P(Br/U2) = 3/9 \qquad P(Br/U3) = 3/8$

$\underbrace{P(U1) = 1/3}_{\text{prob. de eu escolher U1}} \qquad P(U2) = 1/3 \qquad P(U3) = 1/3$

Teorema de Bayes:

$$P(U2/Br) = \frac{P(U2) \cdot P(Br/U2)}{P(U1) \cdot P(Br/U1) + P(U2) \cdot P(Br/U2) + P(U3) \cdot P(Br/U3)}$$

$$P(U2/Br) = \frac{1/3 \cdot 1/3}{1/3 \cdot 1/9 + 1/3 \cdot 1/3 + 1/3 \cdot 3/8}$$

$$P(U2/Br) = \frac{1/9}{1583/5832} = \frac{1}{9} \cdot \frac{5832}{1583} \approx \boxed{0,41}$$

$$P(U3/Br) = \frac{P(U3) \cdot P(Br/U3)}{P(U1) \cdot P(Br/U1) + P(U2) \cdot P(Br/U2) + P(U3) \cdot P(Br/U3)}$$

$$P(U2/Br) = \frac{1/3 \cdot 3/8}{1/3 \cdot 1/9 + 1/3 \cdot 1/3 + 1/3 \cdot 3/8} = \frac{1}{8} \cdot \frac{5832}{1593} \approx \boxed{0,45}$$

Para conferir o resultado, devemos somar todas as probabilidades e obter 1 como resultado

$$P(U1/Br) = \frac{1}{27} \cdot \frac{5832}{1593} \approx 0,13 \qquad \text{ou} \qquad P(U1/Br) = 1 - [P(U2/Br) + P(U3/Br)] \approx 0,13$$

$$\therefore 0,45 + 0,41 + 0,13 \approx 1$$

7 Variável Aleatória Discreta

Sejam E um experimento e S um espaço associado ao experimento.
Uma função X, que associe a cada elemento **s** $\in S$ um número real X(s) é denominada variável aleatória
$\sum P(X = i) = 1$

Exemplo: Duas moedas são arremessadas, independentemente e sequencialmente (esse é o evento E).Qual a probabilidade de sair pelo menos uma vez cara?

E = Lançar duas moedas em sequência, de maneira independente

X = Núm. de caras obtidas nas moedas

$S = \{(c,c);(c,k);(k,k);(k,c)\}$

Ou seja, a questão pede a probabilidade de um X maior igual a 1 $\longrightarrow \quad \therefore \quad P(X \geq 1)$

$$\begin{array}{lcl} (c,c) & \rightarrow & \text{X=2} \\ (c,k) & \rightarrow & \text{X=1} \\ (k,c) & \rightarrow & \text{X=1} \\ (k,k) & \rightarrow & \text{X=0} \end{array}$$

Então, $X = \{0, 1, 2\}$ $\qquad P(X = 0) = \dfrac{1}{4} \quad ; \quad P(X = 1) = \dfrac{2}{4} \quad ; \quad P(X = 2) = \dfrac{1}{4}$

Respondendo a questão: $P(X \geq 1) = \frac{2}{4} + \frac{1}{4} = \boxed{\dfrac{3}{4}}$

A variável aleatória:

Nesse caso, X é uma variável aleatória.
A variável aleatória assume todos os valores possíveis que um número de acontecimentos pode apresentar.
Na questão, o X assumiu todas as quantidades de caras obtidas nas moedas.

Por que discreta? Pois o X só assume valores finitos, ou infinitos numeráveis.

8 Função de probabilidade

Seguindo o mesmo exemplo da variável aleatória:

$$P(X = 0) = \frac{1}{4} = 0,25 \quad ; \quad P(X = 1) = \frac{2}{4} = 0,5 \quad ; \quad P(X = 2) = \frac{1}{4} = 0,25$$

Distribuição de probabilidades de X

$P(X_i)$	X
0,25	0
0,50	1
0,25	2

Temos uma tabela que associa a probabilidade da variável aleatória, com o seu respectivo valor, logo, temos uma função de probabilidade.
Qualquer função de uma V.A também é uma V.A. Exemplo:
$Y = X + X \longrightarrow$ V.A soma dos pontos de dois lançamentos

É possível representar essa relação graficamente

Notação: $P(x_i) = P(X = x_i)$

9 Função de Repartição ou Função de Distribuição Acumulada

Seja X uma variável aleatória discreta

A função repartição da Variável Aleatória X, no ponto x, é igual probabilidade de que X assuma um valor menor igual que x:

$$F(x) = P(X \leq x)$$

Propriedades (tem muito mais):

$F(x) = \sum_{x_i \leq x} \cdot P(x_i)$

$F(-\infty) = 0$

$F(+\infty) = 1$

10 Variável Aleatória Contínua

Quando a variável aleatória pode assumir qualquer valor real entre os valores de mínimo e máximo. Os valores possíveis de X não são numeráveis nesse caso, logo, não é possível calcular uma probabilidade. Por isso, é necessário definir outro conceito:

10.1 Função densidade de probabilidade

Sempre que calcularmos variáveis contínuas, vamos determinar um intervalo.

A função que utilizaremos para o cálculo já não é representada por uma tabela, como na variável aleatória discreta, mas nesse caso, por um gráfico. Onde a área entre os dois valores abaixo da curva, é igual a probabilidade.

$$P(a < X < b) = \int_a^b f(x)dx$$

$f(X) \geq 0$ para qualquer valor de X

$\int_{-\infty}^{+\infty} f(X)dx = 1$ A área total abaixo da curva é a probabilidade máxima, 1.

Exemplo: O tempo de vida útil de um equipamento pode ser expresso por uma V.A contínua X, cuja função densidade é:

$f(X) = \frac{1}{2} \cdot e^{\frac{-x}{2}}$ Para valores de X ≥ 0 e 0 para valores de X < 0

Determine a probabilidade do equipamento durar entre 6 e 18 meses.

vamos verificar se é de fato uma função de densidade, para isso, testaremos se essa sentença:
$\int_{-\infty}^{+\infty} f(X)dx = 1$ é verdadeira nesse caso.

$\therefore \int_{-\infty}^{0} 0dx + \int_{0}^{+\infty} \frac{1}{2} \cdot e^{\frac{-x}{2}} = 1 \longrightarrow$ É igual a 1, portanto, podemos resolver como função densidade

Visto que a sentença é verdadeira, agora podemos calcular a probabilidade.

$$P(0 \leq X \leq 0) = \int_{0,5}^{1,5} \frac{1}{2} \cdot e^{\frac{-x}{2}} = \boxed{0,3064}$$

11 Modelos de Distribuições Discretas de Probabilidade

11.1 Distribuição Bernoulli

Dada a realização de um experimento E, onde sucesso é a ocorrência do evento desejado e fracasso a não ocorrência do mesmo.
Onde x é a variável aleatória sucesso ou fracasso. $\rightarrow x_1 = 1$ (sucesso) ; $x_2 = 0$(fracasso)

Notação:
Esperança: $\mu(x) = p$
Variância: $\sigma^2_{(x)} = pq$

11.2 Distribuição binomial

É uma distribuição de probabilidade adequada aos experimentos que apresentam apenas dois resultados: sucesso ou fracasso, utilizada para descrever uma V.A.
Porém, diferentemente da distribuição de Bernoulli, nesse caso, há repetições de mais de um evento de Bernoulli, sendo as repetições independentes.

$$P(Y = y) = \binom{n}{y} p^y \underbrace{(1-p)}_{q}^{n-y}$$

$Y =$ número de acertos totais

$y =$ número de sucessos que queremos

$n =$ número de repetições (núm. de eventos)

$p =$ probabilidade de sucesso de um evento

$q = (1 - p)$

Notação :

Esperança: $\mu(y) = n \cdot p$

Variância: $\sigma^2_{(y)} = npq$

Exemplo: Supondo 4 questões de 5 alternativas cada. Qual a probabilidade de acertar 3 no chute?

$$P(Y = 3) = \binom{4}{3} \cdot \left(\frac{1}{5}\right)^3 \cdot \left(1 - \frac{1}{5}\right)^{4-3}$$

$$\therefore P(Y = 3) = \frac{4!}{3!(4-3)!} \cdot 0,2^3 \cdot 0,8 = \boxed{0,0256}$$

Qual a probabilidade de acertar 3 ou mais?

$$P(Y \geq 3) = P(X = 3) + P(X = 4)$$

11.3 Distribuição de Poisson

Distribuição de Poisson é aplicada para calcular probabilidades em situações em que o evento ocorre a uma taxa específica, por exemplo:
Números de chamada que o SAMU recebe em 1 hora
Número de falhas em tecido por m^2
Número de relatórios de acidentes enviados a uma seguradora em 1 semana

Portanto, sendo x uma V.A discreta, tal que:
x = núm. de ocorrências em um determinado intervalo

A sua distribuição de probabilidade é dada por:

$$P(x,t) = \frac{(\lambda t)^x}{x!} \cdot e^{-\lambda t}$$

λ = Taxa média de ocorrência
t = Intervalo de tempo

Notação:
$\mu(x) = \sigma^2_{(x)} = \lambda t$

12 Modelos de Distribuições Contínuas de Probabilidade

12.1 Distribuição uniforme ou retangular

X é uma V.A uniformemente distribuída no intervalo [a, b] se a função densidade for dada por:

$\boxed{f(x) = 0}$ para x fora de [a, b]

$\boxed{f(x) = \frac{1}{b-a}}$ para $a \leq x \leq b$

Sua função repartição\Função de Distribuição acumulada é:

$F(x) = 0$ para x fora de [a, b]
$F(x) = \frac{x-a}{b-a}$ para $a \leq x \leq b$
$F(x) = 1$ para $x \geq b$

Notação:

$\mu(x) = \frac{a+b}{2}$ $\sigma^2_{(x)} = \frac{(b-a)^2}{12}$

Exemplo: Um ônibus passa no ponto entre 7h e 7:30h, Se um passageiro chega de 7:24, qual a probabilidade dele conseguir pegar esse ônibus?

X= instante da chegada do ônibus $\longrightarrow$ $0 \leq x \leq 30$
O que queremos é saber a $P(24 \leq x \leq 30)$ Isso não é a mesma coisa que $(a \leq x \leq b)$
Lembrando que para $a \leq x \leq b$ temos como função densidade: $F(x) = \frac{1}{b-a}$

logo, dados os limites superior e inferior, podemos calcular:

$$P(24 \leq x \leq 30) = \int_{24}^{30} f(x) = \int_{24}^{30} \frac{1}{b-a} = \int_{24}^{30} \frac{1}{30-0}$$

$$\int_{24}^{30} \frac{1}{30-0} dx = \frac{x}{30}\Big|_{24}^{30} = \frac{30}{30} - \frac{24}{30} = \boxed{\frac{1}{5}}$$

Também é possível obter esse resultado calculando a área do retângulo observado em um plano cartesiano através das informações da questão. Basta fazer B x H.

13 Distribuição Normal, ou de Gauss, laplace ou Laplace-Gauss

A mais importante distribuição de probabilidade.
É representada pela letra N:

$$X \sim N(\mu, \sigma^2)$$

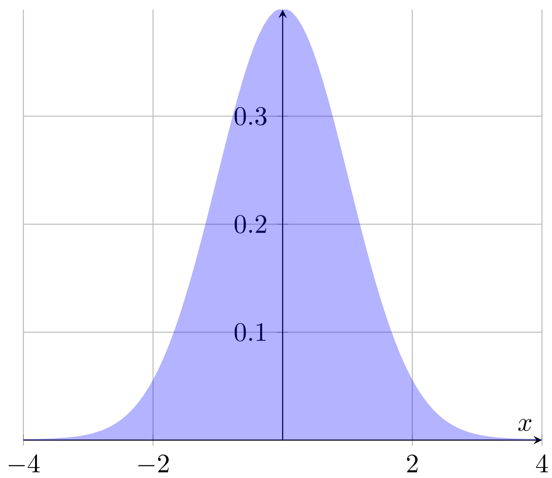
significa que: X é uma variável aleatória que segue uma distribuição normal, com média μ e variância σ^2

Seja X uma variável aleatória contínua, X terá distribuição normal se:

$$f(x) = \frac{1}{\sigma\sqrt{2\Pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

$\mu =$ média da distribuição
$\sigma =$ desvio padrão da distribuição
$f(x) =$ função da densidade de probabilidade

A distribuição é simétrica, o que significa que: $\mu = \tilde{X} =$ Mo
Logo, a média divide o gráfico ao meio, com 50% para cada lado



Quando X segue uma distribuição normal, podemos calcular a probabilidade de X por integral ou pela tabela (tabela de distribuição normal padrão):