
Mariana Perna

MPEI

Aleatório: acontecimento cujo resultado não possa ser determinado com certeza absoluta. Diz-se de uma grandeza que pode tomar certo número de valores, a cada um dos quais está ligada uma probabilidade.
No entanto, na maioria das aplicações existe algum tipo de regularidade que se manifesta se o número de observações/experiências for elevado

Slide 5

0 1 2 3 4 5 6 7 8 9

$$\underline{\quad} \; \underline{\quad} \; \underline{\quad} \; \underline{\quad}$$
$$10 \times 10 \times 10 \times 10$$

a) $P = \dfrac{1}{10000}$

b) $P = \dfrac{1}{10^{20}}$

Experiência Aleatória

é associado

- espaço de amostragem
- conjunto de acontecimentos
- lei da probabilidade

procedimento que deve produzir um resultado mas com um resultado imprevisível e diferente

Espaço de amostragem: é o conjunto de todos os resultados possíveis de uma experiência aleatória. (S) Elementos deste conjunto são considerados resultados elementares.

- S constitui o evento certo (1)
- ∅ (conjunto vazio) representa o evento impossível (0)

A **lei de probabilidade** atribui probabilidade aos vários eventos.

PROBABILIDADE: número associado a um evento que indica a "verosimilhança" de esse evento ocorrer quando se efetua a experiência. Valor entre 0 e 1.

Como se obtêm as probabilidades associadas a eventos?

Há diferentes abordagens:

- **Teoria clássica (de Laplace)** probabilidades teóricas

$$P(\text{evento}) = \frac{\text{n}^\circ \text{ de casos favoráveis}}{\text{n}^\circ \text{ de casos possíveis}}$$

REGRAS: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

equiprobabilidade para os eventos elementares

$P(\bar{A}) = 1 - P(A)$

$\boxed{P(A \cap B) = P(A) \times P(B)}$ só para Acontecimentos Independentes

slide 25

família com 2 filhos
- nascimento de um menino, M
- nascimento de uma menina, F

Qual é a probabilidade de haver pelo menos um rapaz numa família com 2 filhos?

casos possíveis
- MM
- MF
- FM
- FF

casos favoráveis
- MM
- MF
- FM

$$P(\text{pedida}) = \frac{3}{4}$$

Simples mas não perfeita

• Abordagem frequencista

Repete-se a experiência um certo número de vezes (N). Seja K o número de vezes que ocorre o acontecimento que nos interessa. Determina-se a frequência relativa de ocorrência, $f = \frac{K}{N}$

medida empírica de probabilidade

Frequência Relativa

- Se uma experiência for repetida N vezes nas mesmas condições a frequência relativa do evento A é

$$f(A) = \frac{\text{nº de ocorrências do evento A}}{N}$$

$$0 \leq f(A) \leq 1$$

• Se a frequência relativa convergir quando N aumentar, então o limite da frequência relativa é a probabilidade de A.

$$p(A) = \lim_{N \to +\infty} \frac{\text{nº de ocorrências de } A}{N}$$

probabilidade do evento A acontecer

exemplo em MATLAB

Probabilidade de sair 2 caras em 3 lançamentos

```
% simular os 3 lançamentos
lan_3 = rand (3,1) < 0,5

% repetir N vezes
N = 1e6
lancamentos = rand(3,N);
```

cria uma matriz 3×N de valores aleatórios entre 0 e 1
cada coluna representa um ensaio

```
% contar o número de ocorrências de "2 caras"
% contar o número de caras (1's) em cada experiência
numCarasExperiencia = sum(lancamentos);

% contar vezes em que esse número de caras é 2
numOcorrencias = sum(numCarasNaExperiencia == 2)

% calcular a frequência relativa
fr = numOcorrencias / N

% usar como estimativa da probabilidade
pA = fr
```

Problema de aprovações a MPEI
Simulação em MATLAB

```
p = 0.85;    % prob aproximada de aluno avaliado passar
n = 165;     % nº de alunos de MPEI
N = 1e6;     % nº de experiências
K = fix(0.8*n);   % 80% dos alunos
```

fix: remove a vírgula flutuante, arredondando os números inferiormente

```
aprovados = rand(n,N) < p;   % 1 indica a aprovação do aluno

numOcorrencias = 0;
for K1 = K:n
    sucessos = sum(aprovados) == K1;
    fprintf('%d aprovados -> %8d, p=%.5f\n', K1, sum(sucessos), sum(sucessos)/N);
    numOcorrencias = numOcorrencias + sum(sucessos);
end

probSimulacao = numOcorrencias/N;
fprintf('prob de %d ou mais em %d passarem é de %4f\n', K, n, probSimulacao);
```

%d, arredonda o nº para inteiro
\n, acrescenta uma nova linha
%.5f, nº de dígitos à direita da vírgula
%8d, o nº irá ocupar 8 dígitos e será inteiro

prob de 132 ou mais em 165 passarem é de 0.98

• Teoria Axiomática de Probabilidade

com base nas propriedades das frequências relativas e das operações sobre conjuntos

O que é uma axiomática?

Em determinado ponto da evolução de uma teoria de pensamento matemático, torna-se imperioso ordenar, sistematizar e relacionar todos os conhecimentos entretanto nela reconhecidos, isto é, proceder à sua AXIOMATIZAÇÃO.

Axioma 1: probabilidades são não negativas

$$P(A) \geq 0$$

Axioma 2: normalização (S tem probabilidade 1)

$$P(S) = 1$$

Axioma 3a: se A e B forem mutuamente exclusivos

$$P(A \cup B) = P(A) + P(B)$$

Axioma 3b: se $A_1, A_2, \ldots$ for uma sequência infinita de acontecimentos mutuamente exclusivos

$$P\left(\bigcup_{K=1}^{\infty} A_K\right) = \sum_{i=1}^{\infty} P(A_K)$$

TEOREMAS

- Probabilidade do acontecimento complementar

$$P(\bar{A}) = 1 - P(A)$$

Demonstração: como $A \cap \bar{A} = \emptyset$ e $A \cup \bar{A} = S$

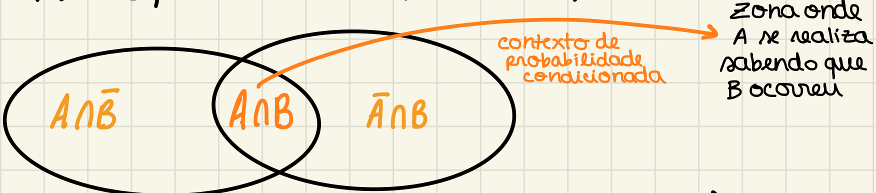
pelos axiomas 2 e 3 $P(A \cup \bar{A}) = P(A) + P(\bar{A}) = P(S) \ldots$

$$P(A) + P(\bar{A}) = 1$$

$$\Leftrightarrow P(\bar{A}) = 1 - P(A) \text{, c.q.d}$$

- Probabilidade da união

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

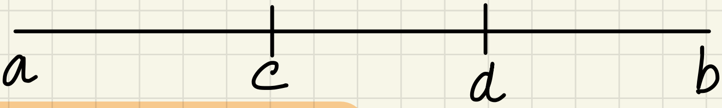


Demonstração: $P(A \cup B) = \left(P(A\bar{B}) + P(AB)\right) + \left(P(AB) + P(\bar{A}B)\right) - P(AB)$

$$= P(A) + P(B) - P(AB) \text{, c.q.d.}$$

• Probabilidade em espaço de amostragem n/ contáveis

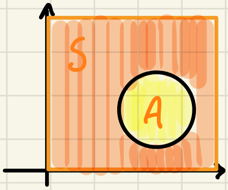
exemplo 1: Escolha de um número real no intervalo $[a,b]$.
Seja o acontecimento A "número pertencer a $[c,d]$"



$$P(A) = \frac{(d-c)}{(b-a)}$$

exemplo 2: No caso de um par de números reais x,y entre 0 e 1

$$S = \{x,y : x \in [0,1] \cap y \in [0,1]\}$$



$$P(A) = \frac{\text{Área}(A)}{\text{Área}(S)}$$

• A axiomática é compatível com a definição frequentista!
(PROVADO nos slides 57, 58 e 59)

• A axiomática é compatível com a Lei de La Place!
(PROVADO nos slides 60, 61, 62 e 63)

NOTA: complementar usando o condicional

$$P(\neg A|B) + P(A|B) = 1$$
$$P(\neg A|B) = 1 - P(A|B)$$

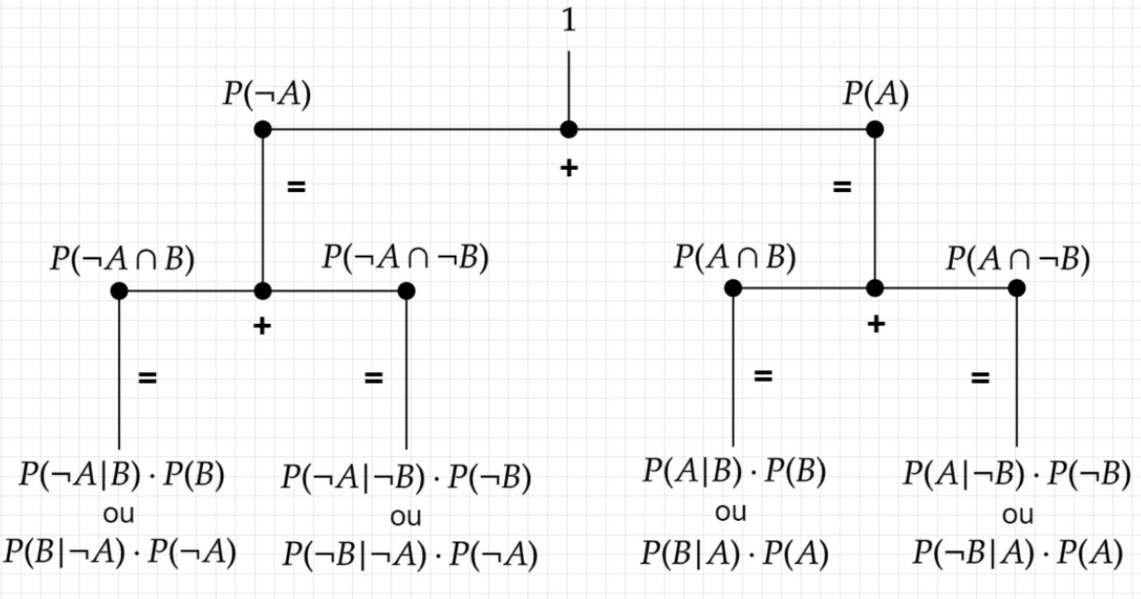
Probabilidade condicional

↳ quando dois acontecimentos estaõ relacionados. A ocorrência de um depende ou faz depender a ocorrência do outro.

$$P(A|B) = \frac{P(AB)}{P(B)} \quad \text{se } P(B) \neq 0$$

Cadeia condicional da probabilidade

é a possibilidade de dividir a probabilidade de um evento usando a probabilidade condicional



Podemos decompor a probabilidade de acordo com sua condional

$$P(\neg A) = P(\neg A \cap B) + P(\neg A \cap \neg B) \qquad P(A) = P(A \cap B) + P(A \cap \neg B)$$

$$1 = P(\neg A) + P(A)$$

$$P(A \cap B) = P(A|B) \times P(B)$$
$$P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$$

Lei da probabilidade total

- dividir para conquistar
- partição do espaço de amostragem A_1, A_2, A_3
- ter $P(B|A_i)$, para todos os i

$$P(B) = \sum_j P(B|A_j) P(A_j)$$

Condicionamento inverso

↳ é simplesmente pensar no problema inverso

Regra de Bayes

A ideia principal do Teorema de Bayes é inverter as probabilidades condicionais. Por exemplo, queremos saber a probabilidade $P(X|Y)$ mas conhecemos apenas a probabilidade de $P(Y|X)$.

Antes, vamos reescrever a nossa fórmula da probabilidade condicional para utilizarmos de uma forma melhor. Vamos passar o denominador multiplicando para o outro lado da equação:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A|B) \cdot P(B) = P(A \cap B)$$

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

e também

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \Leftrightarrow P(B|A) \cdot P(A) = P(B \cap A)$$

então podemos concluir que:

$$\boxed{P(B \cap A) = P(A \cap B) = P(B|A) \cdot P(A) = P(A|B) \cdot P(B)}$$

variação da regra da probabilidade condicional

montando o Teorema de Bayes
sabendo que

- $P(A \cap B) = P(B|A) \cdot P(A)$
- $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

e assim obtemos o Teorema de Bayes

$$\boxed{P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}}$$

Vamos a um típico problema que o Teorema de Bayes é usado: Uma certa doença afeta 1 a cada 10,000 pessoas. Existe um teste para verificar se a pessoa possui a doença ou não.

Desse problema sabemos que:

- A probabilidade do teste dar **positivo**, dado que a pessoa **não tem a doença** é de **2%.**
- A probabilidade do teste dar negativo, dado que a pessoa **possui a doença** é de apenas **1%.**

Se uma pessoa aleatória fizer o teste da doença e der positivo, qual é a probabilidade de que a pessoa tenha a doença ?

Vamos dizer que D é o evento em que a pessoa tem a doença, e que T seja o evento em que o teste tenha dado positivo. Logo queremos descobrir $P(D \setminus T)$.

sabemos

- $P(D) = 0{,}0001$ • $P(T \mid \neg D) = 0{,}02$ • $P(\neg T \mid D) = 0{,}01$

$$P(D \cap T) = P(D \setminus T) \, . \, P(T)$$
$$= P(T \setminus D) \, . \, P(D)$$
$$= (1 - P(\neg T \setminus D)) \, . \, P(D)$$
$$= (1 - 0{,}01) \, . \, 0{,}0001$$
$$= 0{,}000099$$

$$P(T) = P(T \cap D) + P(T \cap \neg D)$$
$$= (P(T \setminus D) \, . \, P(D)) + (P(T \setminus \neg D) \, . \, P(\neg D))$$
$$= ((1 - P(\neg T \setminus D)) \, . \, P(D)) + (P(T \setminus \neg D) \, . \, P(\neg D))$$
$$= ((1 - 0{,}1) \, . \, 0{,}0001) + ((0{,}02) \, . \, 0{,}9999)$$
$$= 0{,}020097$$

$$P(D \setminus T) = \frac{P(D \cap T)}{P(T)} = \frac{0{,}000099}{0{,}020097} \approx 0{,}0049$$

Indepência

dois acontecimentos são independentes se e só se
$$P(AB) = P(A) . P(B)$$

aplica-se mesmo que $P(A) = 0$

implica $\boxed{P(A|B) = P(A)}$ ocorrência de B não fornece informações sobre ocorrências de A

Experiências de Bernoulli

consiste em realizar uma experiência e registar se um dado acontecimento se verifica (sucesso) ou não (falha).

Qual a probabilidade de K sucessos em n ensaios independentes?

- seja p a probabilidade de sucesso e $(1-p)$ a de falha
- a probabilidade de K sucessos e $(n-K)$ falhas é:

$$p^K (1-p)^{n-K}$$

- K sucessos em n experiências podem ocorrer C^n_K maneiras

então a probabilidade pedida é:

$$\boxed{P_n(K) = C^n_K \; p^K \; (1-p)^{n-K}}$$

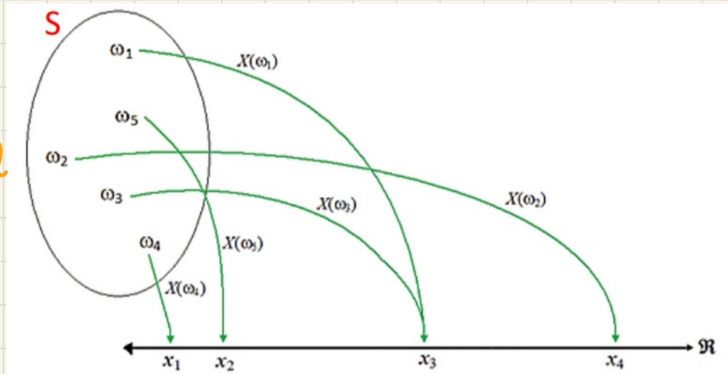
Lei Binomial

Variáveis aleatórias

Uma variável aleatória é uma função que mapeia o espaço de amostragem na reta real. Por outras palavras, é o resultado numérico das nossas experiências (aleatórias).

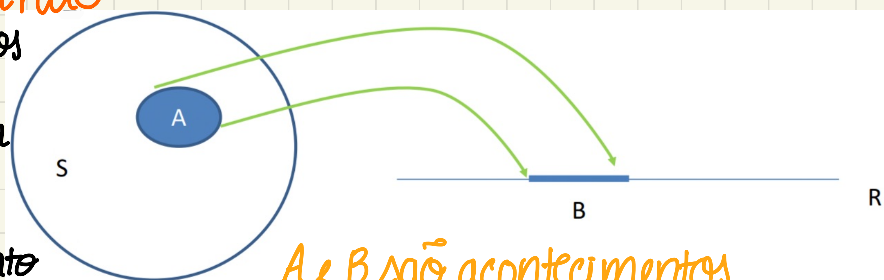
DEFINIÇÃO: Uma variável aleatória escalar X é formalmente definida como sendo um mapeamento de um espaço amostral S para a reta real.

a qq elemento w de S associa-se uma imagem X(w) na reta real



Caso contínuo

se os conjuntos que representam os eventos forem contínuos, o mapeamento é para um segmento da reta real



A e B são acontecimentos equivalentes

Tipos de variáveis aleatórias:

- Discreta, se os valores que a variável aleatória pode assumir forem finitos (ou infinitos mas contáveis, por exemplo: o nº de acessos por minuto a uma página web);

- contínua, se os valores que pode assumir formarem um ou vários intervalos disjuntos, por exemplo a duração de uma aula no zoom;

- mista, onde se verificam os atributos que definem os 2 tipos anteriores.

Caracterização das variáveis aleatórias

PARTE 1

• Distribuição de probabilidades

as variáveis aleatórias são caracterizáveis pelo conjunto de valores que podem assumir e as probabilidades associadas

• Função (massa) de probabilidade

Uma variável aleatória discreta escalar X é especificada por:

1. conjunto de valores que pode assumir $x_i, i = 1, 2, \dots$
2. probabilidade associada a cada um desses valores $p_x(x_i)$ → denominada de função massa de probabilidade

$$p_x(x_i) = P(X = x_i) = P(w : X(w) = x_i)$$

AXIOMAS:

$$p_x(x_i) \geq 0$$

$$\sum_i p_x(x_i) = 1$$

EXEMPLO: Lançamento de um dado equilibrado e X igual ao nº que sai

X: variável aleatória discreta

função de probabilidade: $x_i = \{1,2,3,4,5,6\}$

$$p_x(x_i) = \frac{1}{6}$$

```
%% Matlab
xi = 1:6;  p = ones(1,6)/6;
stem(xi,p); xlabel('x'); ylabel('px(x)');
```

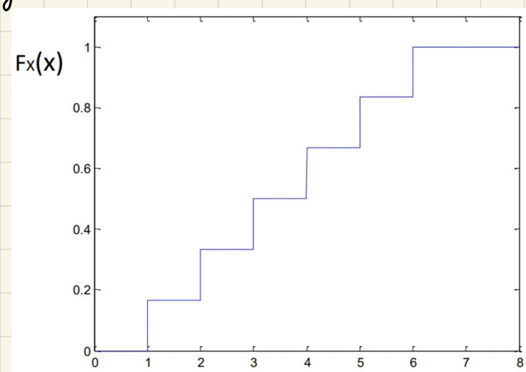
Função distribuição acumulada (discreta), fda

$$F_x(x) = p_x(X \le x) = \sum_{i:\, x_i \le x} p_x(x_i)$$

↓ é uma função não decrescente

- $\lim\limits_{x \to -\infty} F_x(x) = 0$
- $\lim\limits_{x \to \infty} F_x(x) = 1$

EXEMPLO: Para uma variável aleatória discreta a função distribuição acumulada é uma função em escada.



Variáveis aleatórias contínuas

Tb pode ser especificada pela sua função distribuição acumulada

$$F_x(x) = Prob(X \le x)$$

PROPRIEDADES:

$$0 \le F_x(x) \le 1$$

$$\lim_{x \to \infty} F_x(x) = 1$$

$$a < b \Rightarrow F_x(a) \le F_x(b)$$

$$\lim_{x \to -\infty} F_x(x) = 0$$

$$P[a < X \le b] = F_x(b) - F_x(a)$$

Tb podem ser especificadas pela sua função de densidade de probabilidade

$$\frac{prob}{dx} \equiv f_X(x) = \frac{d F_X(x)}{dx}$$

obtém-se derivando a função de distribuição

Não é uma probabilidade, apenas define os valores de probabilidade quando integrada num intervalo

$$p(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$$

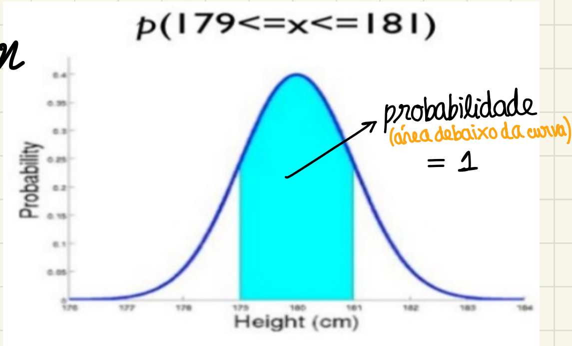
é a probabilidade da variável x pertencer ao intervalo $(x, x+dx)$, sendo dx um acréscimo infinitesimal

Relações entre funções de densidade e de distribuição (caso contínuo)

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$$

Probabilidades e função de densidade

$$P(a < X \leq b) = \int_a^b f_X(x)\,dx$$



Caracterização das variáveis aleatórias

PARTE 2

em muitos casos não necessitamos de toda a informação. No caso dos "bugs" em módulos de código saber o valor médio pode ser suficiente !!!

consideremos N lançamentos de um dado

Média: $\frac{numDe1s}{N} \times 1 + \frac{numDe2s}{N} \times 2 + \ldots + \frac{numDe6s}{N} \times 6$

Assumindo que N tende para infinito

Média $= p(1)\times 1 + p(2)\times 2 + p(3)\times 3 + \dots + p(6)\times 6$

$= \sum_i p(x_i)\, x_i$ c/ $x_i = 1, 2, \dots, 6$

Valor esperado

mas não é algo que devemos esperar que aconteça ...

Consideremos as experiências na origem da variável aleatória X:

assim

o valor esperado de X é o valor médio de X ao repetirmos as experiências indefinidamente

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \dots + X_n}{n}$$

só existe valor esperado se existir o limite

(o limite existe se X_i tiver limite inferior e superior finitos)

X_i é o valor de X na experiência i

Representando por x_i os m diferentes valores que X_i pode assumir e por $K_{i,n}$ o número de vezes que ocorre cada x_i

$$\lim_{n\to\infty} \frac{x_1 K_{1,n} + x_2 K_{2,n} + \dots + x_m K_{m,n}}{n}$$

$$\sum_{i=1}^{m} x_i \lim_{n\to\infty} \frac{K_{i,n}}{n} = \sum_{i=1}^{m} x_i P(X = x_i)$$

PROPRIEDADES:

como $E[x]$ é um operador linear

então:

- $E[aX] = aE[X]$
- $E[X+Y] = E[X] + E[Y]$
- $E[X+c] = E[X] + c$

o Valor esperado de uma variável designa-se por $E(X)$

- no caso DISCRETO $E[X] = \sum_i x_i\, p(x_i)$
- no caso CONTÍNUO $E[X] = \int_{-\infty}^{+\infty} x\, f_x(x)\, dx$

a Média pode não ser suficiente

Por exemplo, se pretendermos comparar as classificações de duas turmas práticas de MPEI, posso ter a mesma média e ter turmas muito diferentes.

Uma medida dessa "dispersão" é dada pela Variância

ideia base

usar a diferença dos valores da variável para a média (valor esperado) e fazer a sua média

e para evitar nºs negativos, usaremos o seu valor quadrático

$$Var(X) = E[(X - E(X))^2]$$
$$= \sigma^2$$
$$= \sum_i [x_i - E(X)]^2 \, p(x_i)$$

PROPRIEDADE

$Var(X) = E[X^2] - E^2[X]$

X → variável aleatória; c → constante

$Var(X+c) = Var(X)$

$Var(cX) = c^2 Var(X)$

Desvio padrão, raiz quadrada da variância (σ)

Interpretação:

– Média ou valor esperado

$E(X)$ pode ser interpretado como:

- Valor médio de X
- Centro de gravidade da função massa de probabilidade (caso discreto) ou função de densidade de probabilidade

– Desvio padrão/variância

Dá uma medida da dispersão da variável aleatória

- pequenos valores indicam variável aleatória muito concentrada em torno da média (se for 0 não temos variável aleatória, todos os valores são iguais à média).

Momento de ordem n (caso discreto)

$$m_n = E[X^n] = \sum x_i^n p_x(x_i)$$

Momento centrado de ordem n (resulta da generalização da variância)

$$E[(X - E(X))^n] = \sum_i (x_i - E[X])^n p_x(x_i)$$

A variância é o momento centrado de 2ª ordem

Variáveis aleatórias - DISTRIBUIÇÕES

As funções de massa de probabilidade e de densidade de probabilidade podem assumir as mais variadas formas. Mas existe um conjunto de "formas" (distribuições) que aparecem repetidamente em mts problemas

Discretas
- Bernoulli
- Binomial
- Poisson
- Geométrica

[...]

Contínuas
- Uniforme
- Normal
- Exponencial

[...]

Distribuições Discretas

• Distribuição de Bernoulli

(diretamente relacionada c/ as experiências de Bernoulli)

Seja A um acontecimento relacionado c/ o resultado de uma experiência aleatória. A variável de Bernoulli define-se como

$$I_A(w) = \begin{cases} 1 & \text{se } w \in A \\ 0 & \text{caso contrário} \end{cases}$$

qd o evento ocorre a variável aleatória I assume o valor 1, caso contrário o valor 0

espaço de amostragem

$S_I = \{0,1\}$

$p = Pr(A)$

$\cdot\, p_I(1) = p$

$\cdot\, p_I(0) = 1-p$

$$E(I) = \sum_i x_i\, p(x_i)$$
$$= 0 \times (1-p) + 1 \times p$$
$$= p$$

$E[I^2] = 0^2 \times (1-p) + 1^2 \times p = p$

$$Var(I) = E[I^2] - (E[I])^2$$
$$= p - p^2$$
$$= p(1-p)$$

· Distribuição Binomial

(diretamente relacionada c/ a Lei Binomial)

Seja X o número de vezes que um acontecimento A ocorre em n experiências de Bernoulli, isto é, X representa o nº de sucessos em n experiências (observações)

$$X = \sum_{j=1}^{n} I_j \longrightarrow \underbrace{S_X = \{0, 1, 2, \dots, n\}}_{\text{espaço de amostragem}}$$

- $p_x(k) = Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
- $F_x(x) = \sum_{k=0}^{[x]} \binom{n}{k} p^k (1-p)^{n-k}$
- $E[X] = np$
- $var(X) = np(1-p)$

Têm distribuição Binomial, p. ex.:

- nº de peças defeituosas num lote de um determinado tamanho
- nº de respostas certas num exame de V/F
- nº de clientes que efetuaram compras em 100 que entraram numa loja

exemplo de aplicação 2: segurança e aviões

- P, probabilidade de falhar
- um avião cai qd + de 50% dos motores falham

avião de 4 motores vs avião de 2 motores

(este avião despenha-se qd 3 ou 4 motores falhem) (este avião despenha-se qd 2 motores falham)

$$p_4 = p_X(3, n=4) + p_X(4, n=4)$$
$$= \binom{4}{3} p^3 (1-p)^{4-3} + \binom{4}{4} p^4 (1-p)^{4-4}$$
$$= 4p^3 - 3p^4$$

$$p_2 = p_X(2, n=2)$$
$$= \binom{2}{2} p^2 (1-p)^{2-2}$$
$$= p^2$$

Relação entre p_2 e p_4

$$\frac{p_4}{p_2} = 4p - 3p^2 = p(4-3p)$$

depende de p

testar para valores de p

• Distribuição geométrica

Seja X o nº de vezes que é necessário repetir uma experiência de Bernoulli até obter um sucesso

prob. sucesso $= p$ prob. falha $= 1-p$

- $p_X(k) = p(1-p)^{k-1}$, $k = 1, 2, 3, \dots$ pq teremos $k-1$ insucessos e depois sucesso
- $F_X(x) = \sum_{k=0}^{[x]} p(1-p)^{k-1}$

exemplo: Supondo que a probabilidade de conseguir contactar o suporte é $p = 0{,}1$ (só ao fim de 10 tentativas). Determine a probabilidade de necessitar de menos de 3 chamadas até conseguir expor o seu problema.

$$Pr(\text{nº chamadas} < 3) = Pr(1 \text{ chamada} \vee 2 \text{ chamadas})$$
$$= p(1-p)^{1-1} + p(1-p)^{2-1}$$
$$= 0{,}1(1-0{,}1)^0 + 0{,}1(1-0{,}1)^1$$
$$= 0{,}19$$

qt + provável atenderem - chamadas teremos que fazer (em média)

- $E[X] = \dfrac{1}{p}$
- $var(X) = \dfrac{(1-p)}{p^2}$

Distribuição de Poisson

Considere-se que temos uma variável Binomial, n cresce e p decresce por forma a $np \to \lambda > 0$

para n grande pode fazer-se as seguintes aproximações:

- $p \cong \dfrac{\lambda}{n}$
- $1 - p \cong 1 - \dfrac{\lambda}{n}$

- $p_n(k) = \dfrac{\lambda^k e^{-\lambda}}{k!} = Pr(X = k)$

, tem apenas um parâmetro, o lambda λ

- $E[X] = \lambda$ → é aproximadamente np

- $Var(X) = \lambda$

A distribuição de Poisson foca-se apenas no nº de ocorrências (discreto) num intervalo de tempo contínuo (ou região do espaço)

Esta distribuição não tem um número de experiências(n) como na Binomial

As ocorrências são independentes das outras ocorrências

Aproximação de Poisson à distribuição Binomial

Problemas envolvendo a distribuição Binomial em que n é grande e o valor de p é pequeno, gerando desta forma eventos raros, são os candidatos à utilização da distribuição de Poisson

REGRA: se $n > 20$ e $np \leq 7$ a aproximação de Poisson é suficientemente próxima para ser usada em vez da Binomial

PROCEDIMENTO:

1. Calcular a média da Binomial $\mu = np$;

2. Como μ é o valor esperado da Binomial, passa a ser o λ ($= E[X]$) de Poisson

3. Usar a fórmula de Poisson

$$P_X(K) = \frac{\lambda^K}{K!} e^{-\lambda}$$

exemplos da aplicação da distribuição de Poisson:

- o nº de resultados que ocorrem num determinado intervalo de tempo ou região é independente do nº que ocorre em qq outro intervalo temporal ou região espacial disjunta;
- a probabilidade que um resultado ocorra durante um intervalo ou região infinitesimal é proporcional ao comprimento do intervalo ou dimensão da região e não depende das ocorrências fora desse intervalo ou região;
- a probabilidade de haver mais que um resultado numa região infinitesimal é desprezável;

Distribuições contínuas

• Distribuição uniforme

$U(a,b)$ é definida por:

• $$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases}$$

• $$F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \end{cases}$$

EXEMPLO:

$P(42 \leq X \leq 45)$ com $U(41,47)$

$$P(x_1 \leq X \leq X_2) = \frac{x_2 - x_1}{b-a}$$

• $$E[X] = \frac{a+b}{2}$$

• $$var(X) = \frac{(b-a)^2}{12}$$

A Função rand() do Matlab gera números obedecendo a uma distribuição uniforme com $a=0$ e $b=1$ então $U(a,b) = a + rand()*(b-a)$

• Distribuição Normal (ou Gaussiana)

• $$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$$

• notação: $N(m,\sigma^2)$

• curva em forma de sino, simétrica em torno da média (m) e com alargamento σ

• $$F_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

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

é mt comum utilizar-se μ em vez de m para representar a média

Guassiana normalizada

como existe um nº infinito de combinaçoẽs para m e σ pode gerar-se uma família infinita de curvas

$\Downarrow$ entaõ

Foi desenvolvido um mecanismo pelo qual qualquer distribuiçaõ normal pode ser convertida numa distribuiçãо única, a Gaussiana normalizada $N(0,1)$.

A fórmula de conversaõ é:

$$\boxed{Z = \frac{x - m}{\sigma}}$$

, subtrair a média e dividir pelo desvio padraõ

- Funçaõ densidade de probabilidade

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x)^2}{2}}$$

- Funçaõ de distribuiçaõ acumulada

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t)^2}{2}} dt$$

valor tabelado

$$F_X(x) = \Phi\left(\frac{x - m}{\sigma}\right)$$

A **distribuição normal** adequa-se / ajusta-se a mts características humanas e a mts outras coisas da natureza. Surge quando vários efeitos acumulados e independentes se sobrepõem.

Função massa de probabilidade da Binomial

de média $m = np$ e $\sigma^2 = np(1-p)$ com m não muito pequeno e n elevado pode ser aproximado por:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(k-m)^2}{2\sigma^2}}$$

, ou seja a distribuição normal desde que $m = np$ e $Var(X) = np(1-p)$

• Distribuição exponencial

→ é não negativa;

→ surge em exercícios envolvendo filas de espera e fiabilidade
- tempo até um PC avariar
- tempo entre chegada de utentes às urgências

→ está relacionada c/ a distribuição (discreta) de Poisson
(se o nº de acontecimentos que ocorre num intervalo seguem distribuição de Poisson, o tempo entre eles segue distribuição exponencial)

- $$f_X(x) = \begin{cases} 0 & , x < 0 \\ \lambda e^{-\lambda x} & , x \geq 0 \end{cases}$$

- $$F_X(x) = \begin{cases} 0 & , x < 0 \\ 1 - e^{-\lambda x} & , x \geq 0 \end{cases}$$

- $E(X) = \dfrac{1}{\lambda}$

- $var(X) = \dfrac{1}{\lambda^2}$

exemplo de aplicação:

A vida útil, em milhares de horas, de um componente de um robô é uma variável aleatória com distribuição exponencial de valor médio 10 (milhares de horas).
Qual a probabilidade de um desses componentes selecionado ao acaso durar menos de 4000 horas?

$$\lambda = \frac{1}{10} = 0,1 \qquad P(X<4) = \int_0^4 0,1\, e^{-0,1x}\, dx$$
$$= 0,33$$

·Distribuição dos primeiros dígitos
(Distribuição de Benford)

função de probabilidade

$$P(d) = \log_{10}(d+1) - \log_{10}(d)$$
$$= \log_{10}\left(\frac{d+1}{d}\right)$$
$$= \log_{10}\left(1 + \frac{1}{d}\right)$$

• Distribuição de Zipf

A lei de Zipf prevê que num dado texto, a probabilidade de ocorrência p(n) de uma palavra esteja ligada à sua ordem n na ordem das frequências por uma lei da forma:

$$p(n) = \frac{K}{n}$$

K é uma constante dependente da língua

p(n) é estimada c/ base na contagem de ocorrências de palavras num texto ou conjunto de textos