Ciência dos dados

Variáveis aleatórias discretas

Aula de hoje

Ao final desta aula, o aluno deve ser capaz de:

- Descrever e aplicar propriedades de distribuições de probabilidades de variáveis aleatórias discretas
- Compreender e aplicar propriedades de esperança e variância.

Variáveis aleatórias

Quando tomamos decisões em face da incerteza, raramente elas se baseiam apenas na probabilidade.

Na maioria dos casos, devemos também saber algo sobre as consequências potenciais da tomada de decisão (perdas, lucros, penalidades ou recompensas).

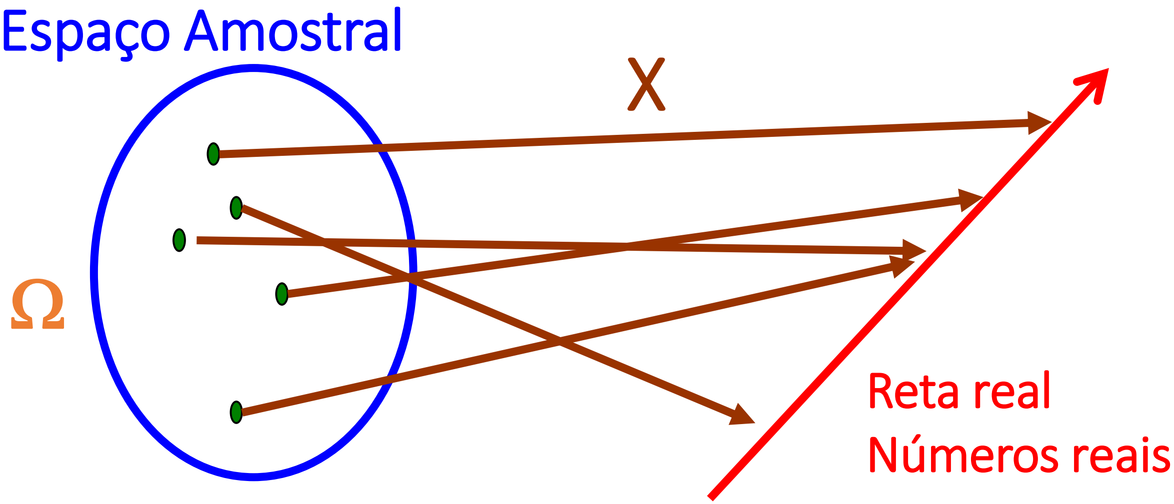
Variáveis aleatórias

Por exemplo, uma construtora precisa decidir se apresenta proposta para um projeto que lhe oferece a perspectiva de R$ 250.000,00 de lucro com probabilidade de 20% ou de um prejuízo de R$ 50.000,00 (em consequência de uma crise financeira no país) com 80% de probabilidade.

A probabilidade da construtora ter lucro não é muito grande, mas a quantia que ela pode ganhar é muito maior do que a que ela pode perder.

Este exemplo mostra a necessidade de um método que permita combinar probabilidades e consequências.

Variáveis aleatórias



Variável aleatória: função que associa um número real a cada ponto do espaço amostral (possível realização do experimento aleatório).

Exemplos de Variáveis Discretas

- Uma indústria de aviões recebe pedidos de um determinado tipo de jato comercial por ano (x: 0, 1, 2, 3, 4, ...)
- Um empresário com 20 escritórios comerciais quer saber o número de escritórios que estão alugados por mês (x : 0, 1, 2, 3, ..., 20)
- Número de vendas num dia de funcionamento de uma loja (x : 0, 1, 2, ...)
- Número de chamadas telefônicas recebidas numa central num dia (x: 0,1, 2, ...)
- Um vendedor de seguros aborda 5 clientes por dia, recebendo 50,00 de comissão a cada venda. A variável de interesse é o ganho diário do vendedor (x: 0,00; 50,00; 100,00; 150,00; 200,00; 250,00)
- Número de peças defeituosas num lote com 30 peças (x: 0, 1,..., 30)
- Número de papéis que fecharam em alta ao final de um pregão (x: 0, 1, ..., n)

Exemplo Comissão

Uma corretora de seguros paga uma comissão de R$50,00 a cada novo seguro que um corretor vende. A probabilidade de um cliente adquirir o seguro é de 0,20.

a) Descreva como pode se comportar a comissão se um corretor ao abordar 2 clientes de maneira independente um do outro.

b) Qual a probabilidade de um corretor ganhar apenas R$50,00?

c) Qual a comissão média diária do corretor se todos os dias ele aborda exatamente dois clientes?

S_i: o cliente *i* compra o seguro

N_i: o cliente *i* não compra o seguro

Aula passada: Probabilidade condicional

A **probabilidade condicional** de A dado B é dada por

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

$$\xRightarrow{\text{isto é}}$$

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

Aula passada: Eventos Independentes

Dois eventos A e B quaisquer contidos ao mesmo espaço amostral são **independentes** quando

$$P(A|B) = P(A) \quad \text{ou} \quad P(B|A) = P(B) \qquad (1)$$

Ainda, a **probabilidade condicional** de A dado B é dada por

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \xRightarrow{\text{isto é}} P(A \cap B) = P(B)P(A|B)$$

Logo, se A e B forem independentes (substituímos (1) na equação acima), temos:

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

Função de probabilidade

Seja uma variável aleatória (v.a.) discreta X, que assume valores x´s.

A função que associa a probabilidade de ocorrência em cada valor x, isto é,

$$f(x) = P(\mathrm{X} = x)\ ,$$

chama-se função de probabilidade.

Distribuição de probabilidade de uma v.a. discreta

A distribuição da v.a. X (ou distribuição de probabilidades da v.a. X) é o conjunto de todos os pares formados por

$$\{x, P(\mathrm{X} = x)\},$$

isto é, pelos valores de X e as respectivas probabilidades da variável assumir tais valores.

Propriedades da função de probabilidade de uma v.a. discreta

Seja X uma v.a. discreta assumindo valores x´s:

$$0 \leq P(\mathrm{X} = x) \leq 1$$

$$\sum_{x} P(\mathrm{X} = x) = 1$$

Média e Variância
de uma variável aleatória
discreta

Exemplo Comissão

Qual a comissão média diária do corretor se todos os dias ele aborda exatamente dois clientes?

E o desvio-padrão?

Distribuição de probabilidades de X

x	P(X=x)
0	0,64
50	0,32
100	0,04
Soma	**1,00**

Esperança (média ou valor esperado) de uma variável discreta

O **valor esperado** (ou esperança ou média) de uma variável aleatória **discreta** é uma medida de tendência central dada por

$$E(\mathrm{X}) = \sum_{x} x\, P(\mathrm{X} = x)$$

Variância

A **variância** de uma variável aleatória **discreta** X é dada por

$$Var(\mathrm{X}) = \sum_{x} [x - E(\mathrm{X})]^2 \, P(\mathrm{X} = x)$$

Variância - alternativamente

Como a **variância** é o valor esperado de $[x - E(X)]^2$, prova-se que outra maneira de calcular é:

$$Var(\mathrm{X}) = E(\mathrm{X}^2) - [E(\mathrm{X})]^2,$$

em que

$$E(\mathrm{X}^2) = \sum_x x^2 P(\mathrm{X} = x)$$

Exemplo Comissão – Nova proposta

Distribuição de probabilidades de X

x	P(X=x)
0	0,64
50	0,32
100	0,04
Total	1,00

A corretora de seguros irá fornecer um aumento na comissão diária dos corretores. Entretanto, cada corretor poderá escolher uma das seguintes opções:

1. Nova comissão será a comissão diária atual mais um fixo de R$ 20,00.
2. Nova comissão será o dobro da atual comissão diária.

Determine as distribuições de probabilidades das novas comissões. Calcule esperança e variância. Diga qual você escolheria?

Exemplo Comissão – Nova proposta

Construa a **distribuição de probabilidades** da comissão diária considerando as duas propostas e calcule o **valor esperado** e a **variância** de cada opção.

Escolha qual delas é melhor para aumentar o ganho diário de um corretor. Justifique sua resposta.

Propriedades da Esperança

Seja X uma variável aleatória qualquer, então

(i) $E(\mathrm{X}+d)=E(\mathrm{X})+d$, onde d é uma constante.

(ii) $E(c\,\mathrm{X})=c\,E(\mathrm{X})$, onde c é uma constante.

(iii) Combinando (i) e (ii):

$$E(c\,\mathrm{X}+d)=c\,E(\mathrm{X})+d,$$

onde c e d são constantes.

Propriedades da Variância

Seja X uma variável aleatória qualquer, então

(i) $Var(\mathrm{X} + d) = Var(\mathrm{X})$, onde d é uma constante.

(ii) $Var(c\,\mathrm{X}) = c^2\,Var(\mathrm{X})$, onde c é uma constante.

(iii) Combinando (i) e (ii):

$$Var(c\,\mathrm{X} + d) = c^2\,Var(\mathrm{X}),$$

onde c e d são constantes.

Exemplo 2

Uma empresa de segurança visita, em um dia de trabalho, dois potenciais clientes para oferecer seus serviços. A probabilidade de fechar contrato com o primeiro cliente visitado no dia é da ordem de 10%. Quando a primeira visita resulta em contrato fechado, a probabilidade de se fechar contrato na segunda visita quadruplica, caso contrário, ela se mantém em 10%.

Admitindo que o **custo do dia de trabalho** seja da ordem de **R$30,00** e que a receita obtida com **cada contrato fechado** seja da ordem de **R$500,00**:

a) Encontre a distribuição de probabilidades da variável Lucro.

b) Qual a probabilidade de se ter prejuízo num dia?

c) Quanto se espera lucrar por dia de trabalho?

d) Qual a variância do lucro?

e) Qual valor esperado e variância do Lucro, se esse cair em 10%?

Exercícios

Exercício 1

Um rapaz está pensando em convidar sua namorada para sair. O problema é que as despesas correm por sua conta. Eles podem ir ao cinema ou ao teatro. 70% das vezes ela prefere ir ao cinema, nesse caso, ele gasta $70,00 com os ingressos. Quando eles vão ao teatro, o gasto fica em $190,00. Se eles forem ao cinema, ele sabe que em 80% das vezes ela pede para ir jantar, a despesa adicional do jantar fica em $130,00; 20% das vezes, eles vão direto para casa. Levando a namorada ao teatro, em 40% das vezes ela pede para ir jantar e 60% das vezes eles vão direto para casa.

a) Qual a distribuição de probabilidades do gasto que o rapaz tem com a namorada?

b) Qual o gasto médio? E o seu desvio-padrão?

c) Com a inflação deste ano, o gasto total aumentou até agora $9, mas com a crise geral, o casal resolveu reduzir esse novo gasto total em 15%. Calcule o novo gasto médio e respectivo desvio padrão.

Exercício 2 - Montgomery e Runger

Exercício 3-33

Um arranjo consiste em três componentes mecânicos.

Suponha que as probabilidades de o primeiro, o segundo e o terceiro componentes satisfazerem as especificações sejam iguais a 0,95; 0,98 e 0,99.

Considere que os componentes sejam independentes.

Construa a distribuição de probabilidades do número de componentes no arranjo que satisfazem as especificações.

Exercício 3 - Montgomery e Runger

Exercício 3-67

O sistema de controle aéreo, chamado PASS (Primary Avionics Software Set), do ônibus espacial uso quatro computadores independentes trabalhando em paralelo.

Em cada etapa crítica, os computadores "votam" para determinar a etapa apropriada. A probabilidade de o computador pedir para manobrar para a esquerda quando uma manobra para a direita seria apropriada é de 0,0001.

Seja X o número de computadores que votam em uma manobra para a esquerda quando uma manobra para a direita seria apropriada.

Construa a distribuição de probabilidades de X. Calcule média e desvio padrão.

Atividade

- Download do notebook pelo Github
- Fazer individual e discutir em grupo
- Usar arquivo:

`Aula10_Atividade_....ipynb`

Exercício

- Download do notebook pelo Github
- Fazer individual e discutir em grupo
- Usar arquivo:

`Aula10_Exercicio_....ipynb`