Probabilidade e Estatística

Aula 5

Probabilidade: Distribuições de Discretas – Parte 1

Leitura obrigatória:

Devore, 3.1, 3.2 e 3.3

Objetivos

Nesta parte, vamos aprender:

- Como representar a distribuição de probabilidade de uma variável aleatória discreta
- Como calcular o valor esperado e a variância de uma variável aleatória discreta.
- Exemplos especiais de variáveis aletórias discretas: binomial, Poisson e hipergeométrica

Exemplo

- Exemplo: Suponha que estamos interessados na nota da população de uma turma de 10 alunos. Os valores das notas para esta população são:

  8, 8, 9, 9, 9, 9, 9, 10, 10, 10

- O experimento consiste em sortear um aluno aleatoriamente desta turma e registrar a nota dele.

- Qual é a variável de interesse? Esta variável é aleatória (pode mudar quando o experimento é repetido)?

Variáveis Aleatórias

- Uma **variável aleatória** representa um possível resultado ***numérico*** de um evento incerto.

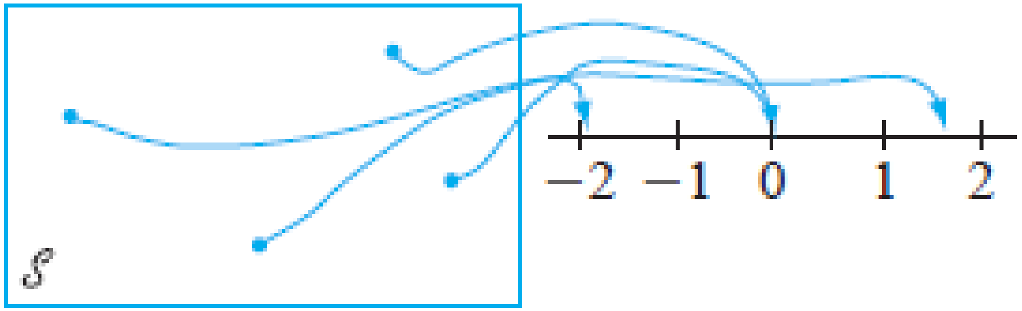
- **Exemplos:**
  - Número de lâmpadas em uma amostra de 10 lâmpadas que queimam antes de 10000 horas.
  - Peso total da bagagem de uma amostra de 25 passageiros de certo vôo.
  - Número de um acidentes em determinado trecho de rodovia em um mês.
  - Altura de um cidadão de Natal selecionado aleatoriamente.
  - Profundidade de um ponto selecionado aleatoriamente em um lago.

Variáveis Aleatórias

- Uma **variável aleatória** representa um possível resultado ***numérico*** de um evento incerto.

- Ou seja, dado um experimento aleatório com espaço amostral S, uma variável aleatória é <u>qualquer função que associe um valor (pertencente aos reais) a cada um dos resultados de S</u>.

Definição!



Variável de Bernoulli

Exemplo: Um estudante liga para a coordenação. Ele pode ter sorte e ser atendido imediatamente (S de sucesso) ou ele pode ficar na espera (F de fracasso).

- Espaço amostral: S = {S,F}
- Defina uma variável aleatória X tal que:
  - X(S)=1 e X(F)=0
  - Assim se X=1 o aluno é atendido imediatamente e se X=0 o aluno espera para ser atendido.

A variável que só pode assumir os valores 0 ou 1 é chamada de **variável aleatória de Bernoulli.**

Definição!

Variáveis Aleatórias

- Uma **variável aleatória** representa um possível resultado numérico de um evento incerto.
  - Ou seja, dado um experimento aleatório com espaço amostral S, uma variável aleatória é <u>qualquer regra que associe um valor (pertencente aos reais) a cada um dos resultados de S</u>.

- Uma variável aleatória **discreta** está associada a um espaço amostral enumerável (elementos são contáveis).

Definição!

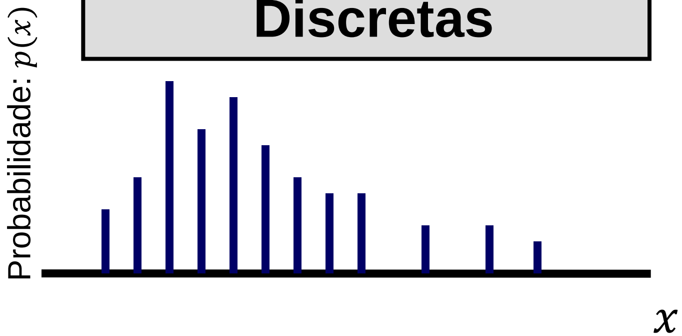
- Uma variável aleatória **contínua** está associada a um espaço amostral não-enumerável e convexo (« sem buracos »).

Definição!

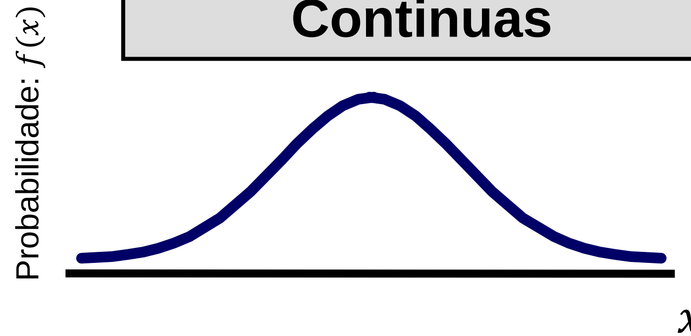
Variáveis Aleatórias

Variáveis Aleatórias

Variáveis Aleatórias Discretas



Variáveis Aleatórias Continuas

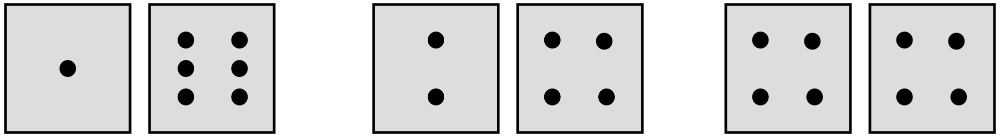


Observação:
- **representamos a variável aleatória com letras maiúsculas (ex: X)**
- **representamos os possíveis valores que esta variável pode assumir com letras minúsculas (ex: x ou $x_1, x_2, \dots x_n$).**
- **o conjunto de todos os valores que a variavél aleatória pode assumir é chamado de suporte.**

Exemplos de V.A. Discretas

Exemplos de variáveis aleatórias discretas:

- Lance um dado duas vezes. Seja X $=$ número de vezes que o resultado 4 aparece. O suporte de X é?



$Suporte\colon\ X = 0{,}1{,}2.$

- Lance uma moeda 5 vezes. Seja X $=$ número de caras que aparecem. Qual é o suporte de X?

$Suporte\colon\ X = 0{,}1{,}2{,}3{,}4{,}5.$

Variáveis Aleatórias

Exercício: Para cada um dos exemplos de variáveis aleatórias abaixo, indique se são discretas ou contínuas e forneça um valor possível para o suporte.

- Número de lâmpadas em uma amostra de 10 lâmpadas que queimam antes de 10000 horas.
- Peso total da bagagem de uma amostra de 25 passageiros de certo vôo.
- Número de um acidentes em determinado trecho de rodovia em um mês.
- Altura de um cidadão de Natal selecionado aleatoriamente.
- Profundidade de um ponto selecionado aleatoriamente em um lago.

Exemplo

Exemplo: Suponha que estamos interessados na nota da população de uma turma de 10 alunos. Os valores das notas para esta população são:

8, 8, 9, 9, 9, 9, 9, 10, 10, 10

O experimento consiste em sortear um aluno aleatoriamente desta turma e registrar a nota dele.

Qual o nosso modelo de probabilidade? Ou seja, como descrever as possibilidades de notas diferentes e a chance de cada nota ser sorteada?

Distribuição de Probabilidade Discreta

- A **função massa de probabilidade** (fmp) de uma variável aleatória discreta é definida para cada número x do suporte de uma variável aleatória X por:

Definição!

$$p(x) = P(X = x) = P(todos\ os\ eventos\ w \in S\ tais\ que\ X(w) = x)$$

Ou seja, é a probabilidade de todos os eventos, w, do espaço amostral S que estão associados ao valor x da v.a. X.

Lembrem-se que:

- X (maiúscula) representa a variável aleatória que é função de resultados do espaço amostral, w, por isso: $X(w)$
- x (minúscula) representa um dos valores possíveis que X pode assumir.
- Ao calcularmos $p(x)$ para todos os valores possíveis de, X, $(x_1, x_2, \ldots, x_n)$ obtemos o modelo de probabilidade

Distribuição de Probabilidade Discreta

Propriedades da função massa de probabilidade:

- (não negatividade) $p(x)$ é uma probabilidade, então $p(x) \geq 0$ para todo x; e
- (aditividade e normalização) $\sum_{i=1}^{n} p(x_i) = 1$, ou seja, a soma da probabilidade de todos os valores possíveis da variável aleatória é igual a 1.

Distribuição de Probabilidade Discreta

Em resumo, a função massa de probabilidade (fmp) para uma variável aleatória discreta é:

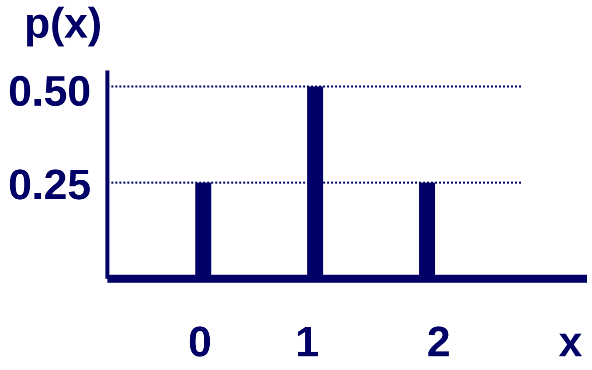
- uma **lista** de todos os **possíveis valores** mutuamente excludentes da variável aleatória: $x_1, x_2, \ldots, x_n$

- uma **probabilidade** para a ocorrência de cada valor da lista ($p(x_1), \ldots, p(x_n)$) e zero para os outros valores.

Distribuição de Probabilidade

Exercício: O experimento é o lançamento de duas moedas. Seja X = número de caras. Qual a função massa de probabilidade (ou fmp) de X?

Valor de x	Probabilidade
0	1/4 = 0.25
1	2/4 = 0.50
2	1/4 = 0.25



Definição:
Função Distribuição Acumulada

- A **função de distribuição acumulada** da variável aleatória X, representada por $F_X(.)$, ou simplesmente $F(.)$, é definida por:

Definição!

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

para todo valor x pertencente aos reais.

- Simplesmente "soma" ou "acumula" a probabilidade de a variável aleatória X assumir todos os valores abaixo ou iguais a x.

Distribuição Acumulada de v.a. Discreta

- A função de distribuição acumulada (FDA), $F(.)$, de uma variável aleatória discreta X com função massa de probabilidade $p(x)$ é definida para cada valor de x por:

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

Simplesmente soma a função massa de probabilidade para todos os valores possíveis da variável aleatória que estão abaixo de x.

Função Distribuição Acumulada: propriedades

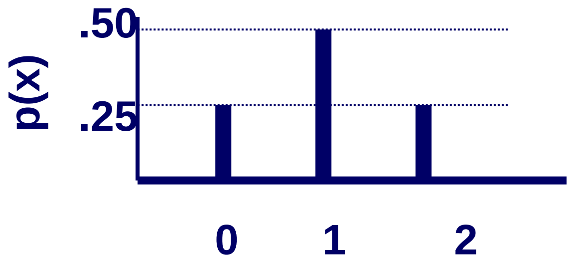
Uma função distribuição acumulada, $F(.)$, satisfaz as seguintes **propriedades:**

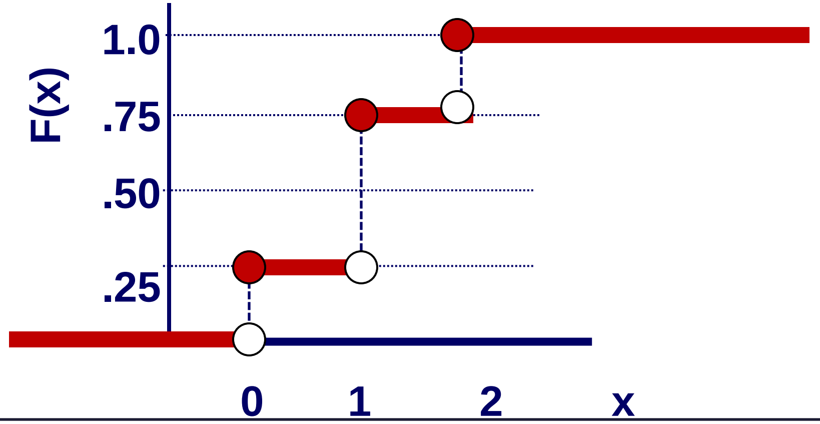
- P1: $F(.)$ é não-decrescente:
  - se $x_1 \leq x_2$, então $F(x_1) \leq F(x_2)$. Porque?
- P2: $F(.)$ é contínua à direita.
- P3: $\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to \infty} F(x) = 1$

Distribuição de Probabilidade Discreta

- A f.d. acumulada é definida para todos os valores da reta.
- Quando a v.a. é discreta, a lista de valores da v.a. representa os pontos de discontinuidade da FDA.

x	p(x)
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25





Distribuição de Probabilidade

- Exercício 1: O espaço amostral de um experimento aleatório é S={a, b, c, d, e, f} em que cada resultado é igualmente provável. Uma variável aleatória é definida como segue:

Resultado	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

- Determine as funções massa de probabilidade e distribuição de probabilidade acumulada.

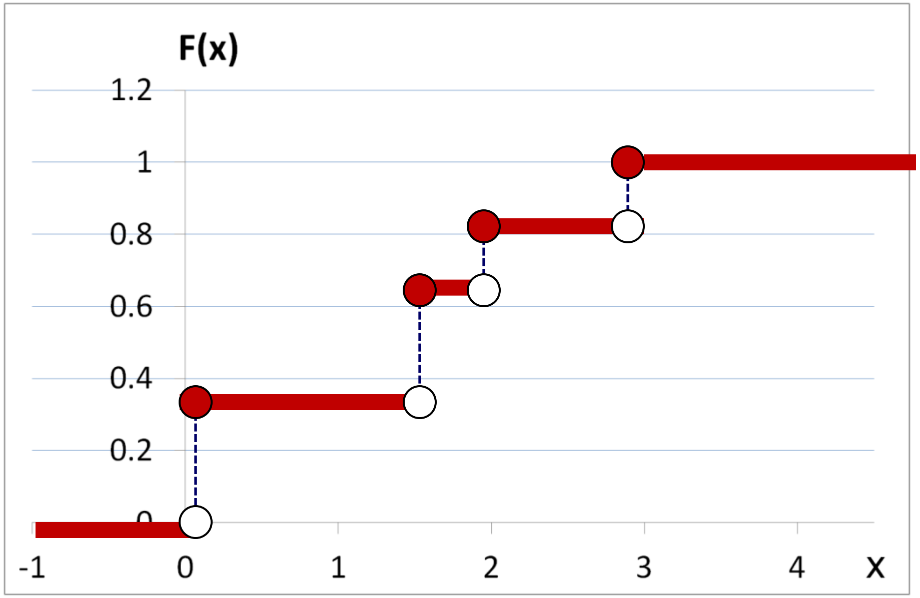
Distribuição de Probabilidade

Exercício 1: Solução

função massa de probabilidade (fmp):

x	p(x)
0	2/6
1.5	2/6
2	1/6
3	1/6

função distribuição acumulada (FDA):

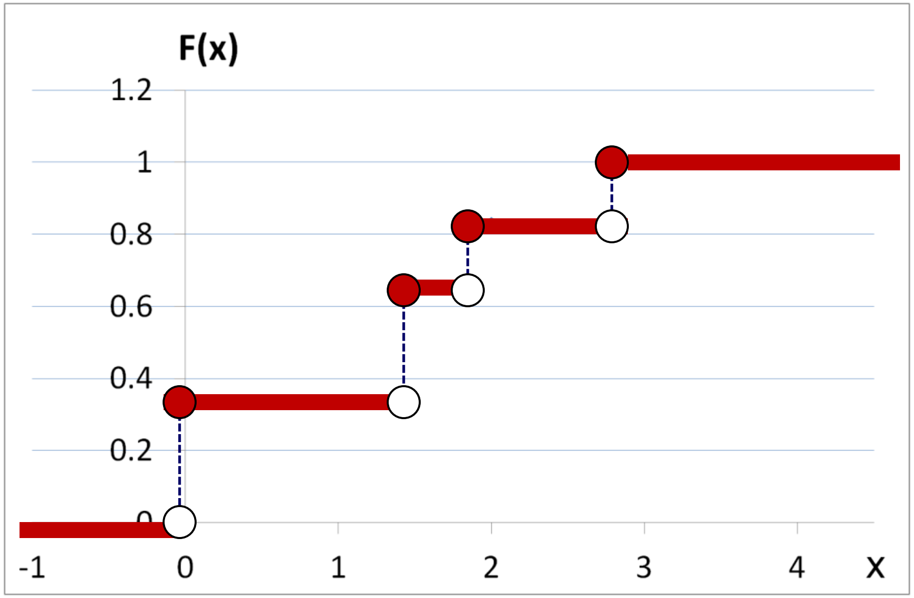


Distribuição de Probabilidade

fmp:

x	p(x)
0	2/6
1.5	2/6
2	1/6
3	1/6

função distribuição acumulada (FDA):



$$F_X(x) = \begin{cases} 0, & \text{se } x < 0 \\ 2/6, & \text{se } 0 \le x < 1.5 \\ 4/6, & \text{se } 1.5 \le x < 2 \\ 5/6, & \text{se } 2 \le x < 3 \\ 1, & \text{se } x \ge 3 \end{cases}$$

Distribuição de Probabilidade

- Exercício 2: Use a função de massa de probabilidade do exercício anterior para determinar as seguintes probabilidades:

  a) $P(X = 1.5)$
  b) $P(0.5 < X < 2.7)$
  c) $P(X > 3)$
  d) $P(0 \leq X < 2)$

Distribuição de Probabilidade

- Exercício 2: Solução

a) $P(X=1.5) = p(1.5) = 2/6 = 0.333$

b) $P(0.5 < X < 2.7) = p(1.5) + p(2) = 3/6 = 0.5$

c) $P(X > 3) = 0$

d) $P(0 \leq X < 2) = p(0) + p(1.5) = 4/6 = 0.667$

x	**p(x)**
0	2/6
1.5	2/6
2	1/6
3	1/6

Distribuição de Probabilidade

Proposições:

P1: Para quaisquer dois números a e b com $a \leq b$,

$$P(a \leq X \leq b) = F(b) - F(a^-)$$

em que $F(a^-) = \lim_{x \to a^-} F(x)$, ou seja, é o limite de $F(.)$ em a pela esquerda. Portanto, a^- é o maior valor de x que é menor do que a.

P2: Para quaisquer dois números a e b com $a \leq b$,

$$P(a < X \leq b) = F(b) - F(a)$$

P3: Para quaisquer dois números a e b com $a \leq b$,

$$P(a \leq X < b) = F(b^-) - F(a^-)$$

Distribuição de Probabilidade

Exercício: Use a FDA (função distribuição acumulada) do exercício anterior para determinar:

$$F(x) = \begin{cases} 0, & se\ x < 0 \\ \frac{2}{6}, & se\ 0 \leq x \leq 1.5 \\ \frac{4}{6}, & se\ 1.5 \leq x < 2 \\ \frac{5}{6}, & se\ 2 \leq x < 3 \\ 1, & se\ x \geq 3 \end{cases}$$

a) $P(0 \leq X \leq 2)$

b) $P(0 \leq X < 2)$

c) $P(0 < X < 2)$

Distribuição de Probabilidade

- Exercício: Solução

$$F(x) = \begin{cases} 0, & se\ x < 0 \\ \frac{2}{6}, & se\ 0 \le x \le 1.5 \\ \frac{4}{6}, & se\ 1.5 \le x < 2 \\ \frac{5}{6}, & se\ 2 \le x < 3 \\ 1, & se\ x \ge 3 \end{cases}$$

a) O evento de interesse inclui o 2 e o 0. Vou usar a acumulada até o 2 e tirar a acumulada até « antes do 0 »: $P(0 \le X \le 2) = F(2) - F(0^-) = 5/6 - 0$

b) O evento de interesse não inclui o 2 e inclui o 0. Vou usar a aumulada até « antes » do 2 e excluir a acumulada até « antes » do 0. $P(0 \le X < 2) = F(2^-) - F(0^-) = 4/6 - 0 = 4/6$

Distribuição de Probabilidade

Exercício 3: Verifique que a seguinte função é uma função massa de probabilidade e determine as probabilidades abaixo:

$$p(x) = \frac{8}{7}\left(\frac{1}{2}\right)^{x}, \text{ para } x = 1, 2, 3.$$

a) $P(X \leq 1.5)$

b) $P(X > 1)$

c) $P(2 < X < 6)$

d) $P(X \leq 1 \text{ ou } X > 1)$

Distribuição de Probabilidade

Exercício 3: Solução

x	p(x)
1	8/(7*2) = 4/7
2	8/(7*4) = 2/7
3	8/(7*8) = 1/7
Total	1

- Para verificarmos que p(x) é uma fmp, note que: p(x) é maior igual a zero para todo x e que a soma de p(x) para todos os valores possíveis de x (1,2,3) é 1.

Distribuição de Probabilidade

Exercício 3: Solução

x	**p(x)**
1	8/(7*2) = 4/7
2	8/(7*4) = 2/7
3	8/(7*8) = 1/7
Total	1

a) $P(X \leq 1.5) = p(1) = 4/7$

b) $P(X > 1) = p(2) + p(3) = 6/7$

c) $P(2 < X < 6) = p(3) = 1/7$

d) $P(X \leq 1 \text{ ou } X > 1) = p(1) + p(2) + p(3) = 1$

Resumo

- Aprendemos a caracterizar um modelo de probabilidade de uma variável aleatória associada a um experimento aleatória através de duas funções:
  - A função massa de probabilidade (fmp), $p(x)$; OU
  - A Função acumulada de proabilidade (FDA), $F(x)$.
  - Precisamos de apenas uma das funções, pois através de uma delas podemos obter a outra. COMO??

- Agora, vamos analisar as características da distribuição de probabilidade, através de medidas numéricas:
  - A média (valor esperado ou esperança) de uma variável aleatória;
  - A média (valor esperado ou esperança) de uma função de uma variável aleatória;
  - A variância e o desvio-padrão de uma v.a. ou de uma função da v.a..

Valor Esperado

- O valor esperado é a média da população!

- Interpretação do valor esperado:

  Se o experimento for repetido um número infinito de vezes, qual é o valor médio obtido?

- O valor esperado **<u>não</u>** representa a moda da população, ou seja, não é o valor que esperamos sair com maior frequência!!

Valor Esperado

- **Exemplo:** Suponha que estamos interessados na nota da população de uma turma de 10 alunos. Os valores das notas para esta população são:

  8, 8, 9, 9, 9, 9, 9, 10, 10, 10

- Qual é a nota média da população?

$$\mu = \frac{8+8+9+9+9+9+9+10+10+10}{10}$$

$$\mu = \frac{8*2+9*5+10*3}{10}$$

$$\mu = 8*0.2+9*0.5+10*0.3$$

O valor esperado é
a média da população.
Podemos calcular a média listando os valores possíveis da variável (8,9,10) vezes a probabilidade de cada valor sair (0.2, 0.5, 0.3)

Valor Esperado

- O **valor esperado** (ou **média**) de uma distribuição discreta é dado pela média ponderada dos valores da variável aleatória, em que os pesos são a probabilidade de cada valor.

Definição!

- Representações: $E(X)\ ou\ \mu\ ou\ \mu_X$.

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

O valor esperado é a média da variável aleatória na população. Em vez de somarmos os valores e dividirmos pelo número total, ponderamos cada valor possível da variável aleatória pela sua chance de ocorrência.

Valor Esperado

- Valor esperado:

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

- **Exemplo:** Lance duas moedas. Seja X = n° de caras. Qual é o valor esperado de X?

x	$p(x)$
0	$1/4$
1	$2/4$
2	$1/4$

$$\mathrm{E}(X) = (0)(0.25) + (1)(0.50) + (2)(0.25) = 1$$

Valor Esperado

- Exercício: Seja X uma variável aleatória com a seguinte função massa de probabilidade:

$$p(x) = \frac{8}{7}\left(\frac{p}{2}\right)^{x}, para\ x = 1,2,3.$$

Determine p tal que $p(x)$ seja uma fmp válida.
Calcule $E(X)$.

Valor Esperado de Função

Qual o valor esperado de uma função de X, $Y = h(X)$??

- Y também é variável aleatória, então usamos a definição de valor esperado de Y:

$$\mu_Y = E(Y) = \sum_{i=1}^{n} y_i p(y_i)$$

$$p(y_i) = p(h(x_i)) = p(x_i)$$

$$\mu_Y = E(h(X)) = \sum_{i=1}^{n} h(x_i) p(x_i)$$

Pois, função $h(.)$ é função determinística (dado x, sabemos com certeza qual o valor de $h(x)$), então a probabilidade de cada valor de $y_i = h(x_i)$é a mesma de x_i.

Valor Esperado de Função

- Exercício: Lance duas moedas. Seja X = o número de caras. Para cada cara, você recebe 30 reais. Calcule o valor esperado da quantia recebida.

$\boldsymbol{x}$	$\boldsymbol{y}$	$\boldsymbol{p(x) = p(y)}$
0	0	$1/4$
1	30	$2/4$
2	60	$1/4$

$$\mu_Y = E\big(h(X)\big) = \sum_{i=1}^{n} h(x_i)p(x_i)$$
$$= 0 * \frac{1}{4} + 30 * \frac{2}{4} + 60 * \frac{1}{4} = 30$$

Valor Esperado: propriedades

- Propriedade: Seja X uma variável aleatória e a e b números reais.

  Então: $\boldsymbol{E(aX + b) \ = \ a * E(X) \ + \ b}$

  Prova:

Valor Esperado de Função

Exercício: Suponha que uma livraria compre 5 cópias de um livro a 6 reais cada. Cada livro será posto a venda por 12 reais. Após 3 meses, os livros não vendidos podem ser devolvidos para a editora por 2 reais cada. Seja *X* a demanda aleatória por livros em 3 meses com a seguinte função massa de probabilidade:

x	0	1	2	3	4	5
p(x)	0.3	0.2	0.2	0.15	0.1	0.05

Qual o valor esperado de ganho da livraria (receita esperada) com este livro?

Valor Esperado de Função

Exercício: Solução

A receita depende da número de livros vendidos que, neste caso, é igual a demanda por livros, X. Vamos chamá-la de $R(X)$.

Se a demanda é igual a X, e 5 cópias foram compradas, sobram $(5 - X)$ cópias não vendidas.

Para cada cópia vendida, o livreiro ganha (12-6) reais.

Para cada cópia não vendida, o livreiro perde (2-6) reais.

Portanto a receita é: $R(X) = 6 * X - 4 * (5 - X) = 10X - 20$

Como a receita é uma função linear da demanda X, temos:

$$E\left(R(X)\right) = E(10X - 20) = 10E(X) - 20$$

Basta calcular a demanda esperada por livros, $E(X)$:

$E(X) = 0 * 0.3 + 1 * 0.2 + 2 * 0.2 + 3 * 0.15 + 4 * 0.1 + 5 * 0.05 =$1.7

Portanto: $E\left(R(X)\right) = 10 * 1.7 - 20 = -3.$ A expectativa é que o livreiro perca dinheiro se comprar 5 livros.

Desafio: quantos livros o livreiro deveria comprar??

Dispersão

- A Variância de uma variável aleatória discreta é definida por:

Definição!

$$\sigma^2 = \sum_{i=1}^{n} [x_i - \mu_X]^2 p(x_i)$$

- O desvio-padrão de uma variável aleatória discreta é dado por:

Definição!

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{n} [x_i - \mu_X]^2 p(x_i)}$$

em que:

$\mu_X = E(X)$ = valor esperado da v. a. X

x_i = i^{esimo} valor da variável aleatória X

$p(x_i)$ = probabilidade de ocorrência do i^{esimo} valor da v. a. X

Dispersão

Exercício: Lance 2 moedas. Seja X = n° de caras. Calcule o desvio-padrão (lembre-se que $E(X) = 1$).

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{n} [x_i - \mu_X]^2 p(x_i)}$$

$$\sigma = \sqrt{(0-1)^2 * 0.25 + (1-1)^2 * 0.5 + (2-1)^2 * 0.25} = 0.707$$

Valores possíves para o n° caras= 0, 1 ou 2

Dispersão

- Exercício: Seja X uma variável aleatória com a seguinte função massa de probabilidade:

$$p(x) = \frac{8}{7}\left(\frac{1}{2}\right)^{x}, para\ x = 1,2,3.$$

- Calcule a variância de X, $V(X)$.

Dispersão: propriedades

Propriedade: Se X é uma variável aleatória, então:

$$\boldsymbol{V}(\boldsymbol{X}) = \boldsymbol{E}(\boldsymbol{X}^2) - \mu^2.$$

Prova:

OBS: esta forma alternativa simplifica bastante os cálculos da variância em alguns casos!

Dispersão: propriedades

Propriedade: Seja X uma variável aleatória e a,b numeros reais.

Então:

$$\boldsymbol{V(aX+b) = a^2V(X)}$$

Prova:

Exercício

Uma loja de eletrodomésticos vende três modelos diferentes de freezer verticais com 13.5, 15.9, 19.1 pés^3 de espaço. Seja X = volume de armazenagem comprado pelo próximo clientes. Suponha que a fmp de X seja

x	13.5	15.9	19.1
p(x)	0.2	.5	0.3

a) Calcule $E(X)$, $E(X^2)$ e $V(X)$.

b) Se o preço de um freezer com X pés^3 de capacidade for $25X - 8.5$, qual será o preço esperado pago pelo proximo cliente?

c) Qual é a variância do preço $(25X - 8.5)$ pago pelo próximo cliente?

d) Suponha que, apesar da capacidade nominal de um freezer ser X, a capacidade real seja $h(X) = X - 0.01X^2$. Qual é a capacidade real esperada do freezer comprado pelo próximo cliente?

Resumo

Aprendemos:

- variável aleatória: discreta, contínua ou mista.
- Modelo de probabilidade para variáveis aleatórias discretas
  - função massa de probabilidade (fmp), $p(x)$;
  - Função acumulada de proabilidade (FDA), $F(x)$.

- Medidas numéricas para uma variável aleatória:
  - O valor esperado (ou média) de uma variável aleatória
  - O valor esperado (ou média) de uma **função** de variável aleatória
  - Propriedades do valor esperado
  - A variância (valor esperado de uma função particular) e o desvio-padrão
  - Propriedades da variância e do desvio-padrão.