Probabilidade e Estatística

Aula 6 - Distribuições Contínuas (Parte 01)

Leitura obrigatória: Devore, Capítulo 4

Objetivos

Nesta aula, vamos aprender:

- Representações de uma v. a. contínua: função densidade de probabilidade e função distribuição acumulada.
- Distribuição uniforme.
- Distribuição normal.
- Distribuição exponencial.

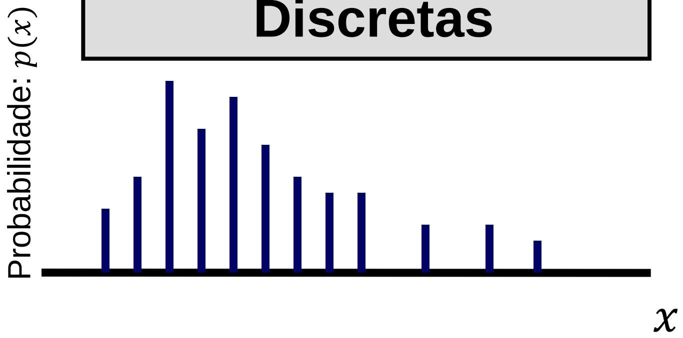
Variáveis Aleatórias

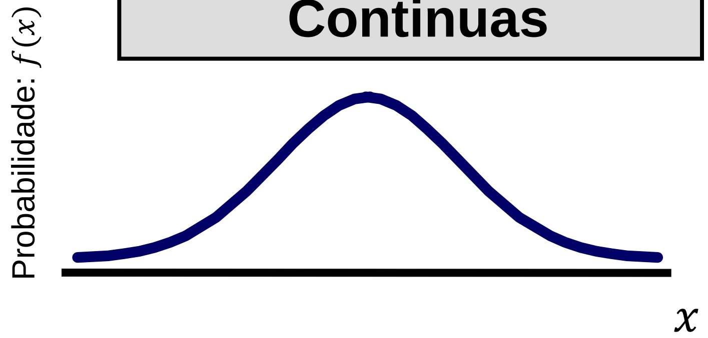
Variáveis Aleatórias

Variáveis Aleatórias Discretas

Variáveis Aleatórias Continuas

Variáveis Aleatórias Mistas





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 $x_1, x_2, \dots x_n$)**

Distribuições Continuas

Definição!

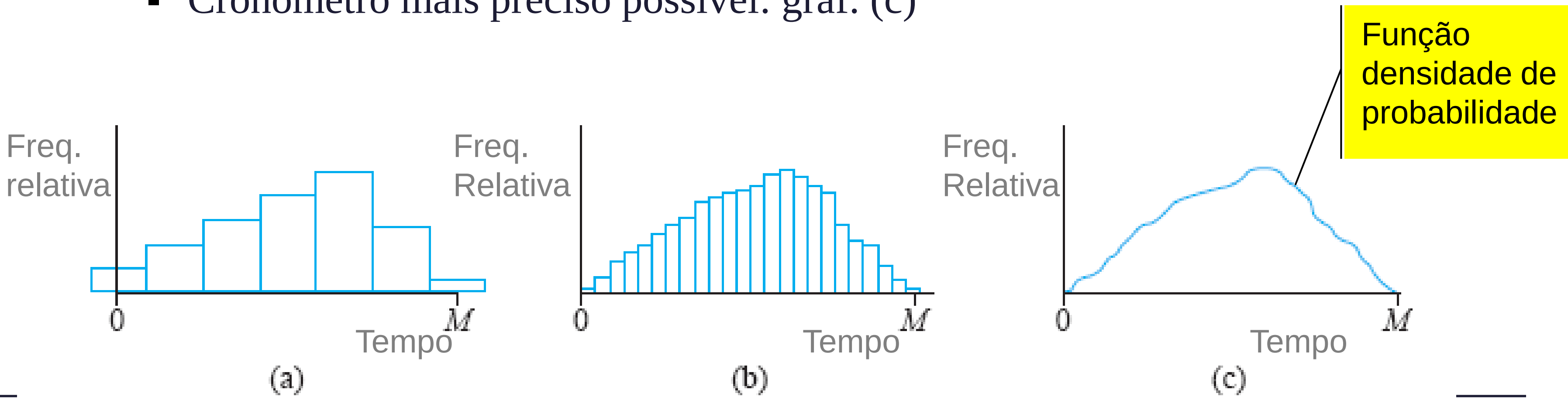
- Uma variável aleatória contínua é uma variável que assume valores em:
  - um conjunto não-enumerável (contínuo) e
  - convexo (combinações lineares de elementos do conjunto também pertencem ao conjunto), ou seja, sem buracos.

- Exemplos:
  - espessura de um objeto
  - temperatura de uma solução
  - profundidade em lago
  - tempo para completar tarefa

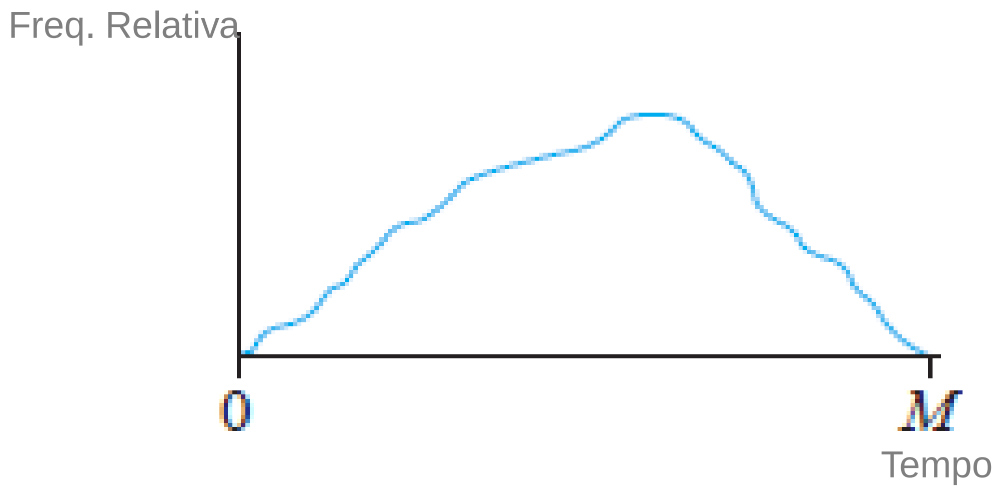
Função Densidade de Probabilidade

Exemplo: Variável aleatória é o tempo que um aluno leva para completar uma tarefa.

- Tempo máximo: M
- Temos 3 instrumentos de medida do tempo:
  - Cronômetro com precisão em minutos: hist (a)
  - Cronômetro com precisão em segundos: hist (b)
  - Cronômetro mais preciso possível: graf. (c)



Função Densidade de Probabilidade



(c)

Função densidade de probabilidade

=

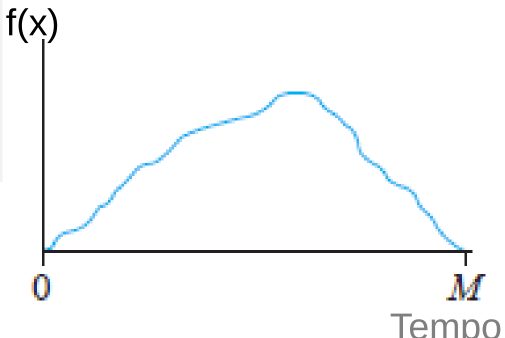
Frequência relativa de uma v.a. contínua

Função Densidade de Probabilidade

Definição!

A função $f(x)$ é dita **função de densidade de probabilidade** (fdp) se:

- $f(x) \geq 0$, para todo x real (não negatividade)
- $\int_{-\infty}^{\infty} f(x)dx = 1$ (normalização)



(c)

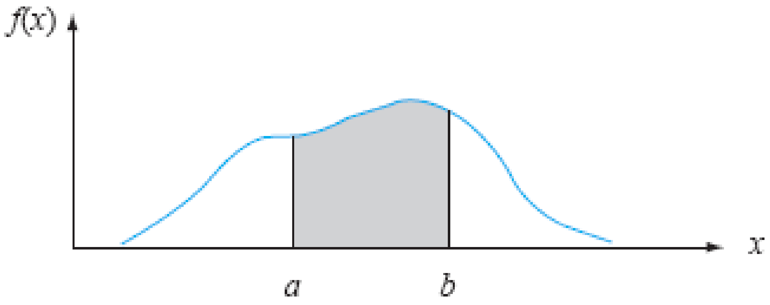
A probabilidade de levar entre 0 e M para executar uma tarefa é 100%.

Função Densidade de Probabilidade

- A *probabilidade* de X pertencer a um intervalo (a, b) é obtida pela área abaixo da fdp(.) de X:

Definição!

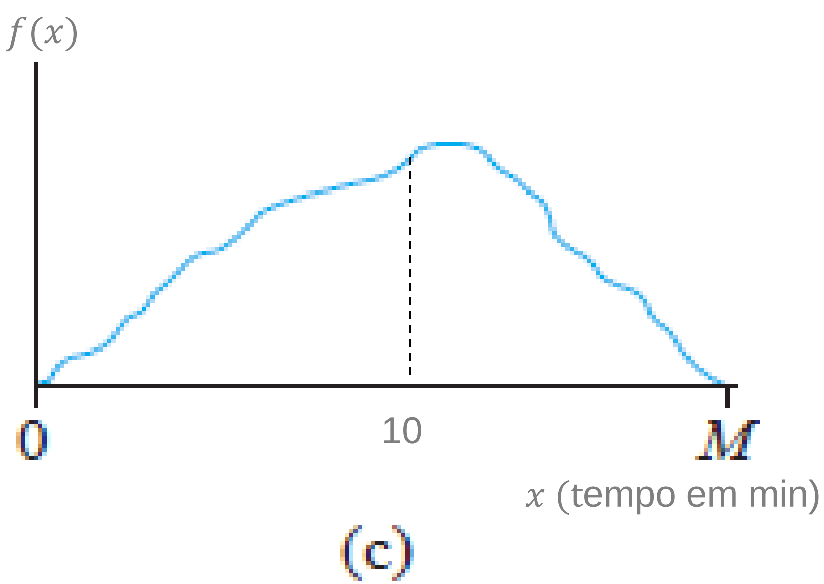
$$\boldsymbol{P(a \leq X \leq b) = \int_a^b f(x)dx}$$



A probabilidade de levar um tempo entre a e b para executar uma tarefa pode ser calculada como a área abaixo da densidade entre estes dois pontos. Equivalente a somar a probabilidade de todos os tempos possíveis com uma v.a. discreta.

Função Densidade de Probabilidade

- Qual é a probabilidade de o aluno levar exatamente 10 minutos para executar a tarefa?



$$P(X=10)=\int_{10}^{10} \boldsymbol{f(x)dx = 0}$$

o que é o mesmo que calcular a área de uma linha.

Pq isso acontece??
Estamos querendo calcular exatamente:
10.000000000.......... Min

que é diferente de:
10.01 min, 9.999min, ...

Função Densidade de Probabilidade

Analogia com a física:

- A probabilidade em um intervalo é o equivalente a massa de uma barra com densidade $f(x)$, ao longo do comprimento da seção da barra.

- A massa em um ponto preciso é igual a 0, pois queremos a massa de uma seção da barra com comprimento nulo!! Igual a cortar a barra em um certo ponto, ou seja, tem massa nula.

Função Densidade de Probabilidade

Propriedade: Se X é uma v.a. contínua, então

$$P(X = c) = 0, \text{ para qualquer valor } c.$$

Por essa razão, temos:

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X \leq b) \\ &= P(a \leq X < b) \\ &= P(a < X < b) \end{aligned}$$

ou seja, <u>os extremos do intervalo são irrelevantes</u>!

Função Densidade de Probabilidade

Exercício: Seja X o tempo entre passagem de dois carros em uma auto-estrada com fdp:

$$f(x) = \begin{cases} 0.15e^{-0.15(x-0.5)}, & se\ x \geq 0.5 \\ 0, & caso\ contrário \end{cases}$$

a) Verifique se $f(x)$ é uma fdp legítima.

b) Qual a probabilidade de $X \leq 5$?

c) Qual a probabilidade de $3 < X < 6$?

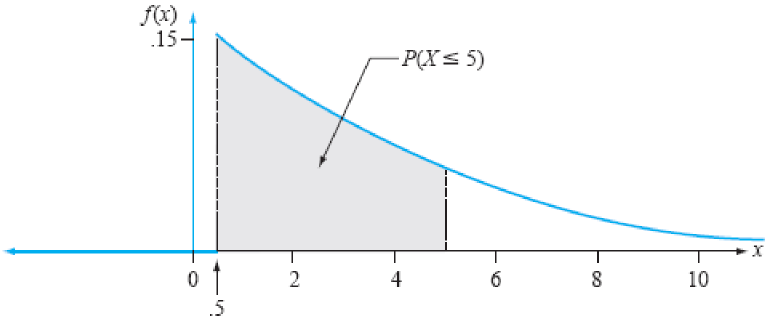
Função Densidade de Probabilidade

Exercício: Solução

$$f(x) = \begin{cases} 0.15e^{-0.15(x-0.5)}, & se\ x \geq 0.5 \\ 0, & caso\ contrário \end{cases}$$

a) $f(x) \geq 0$ e $\int_{-\infty}^{\infty} f(x)dx = 1?$

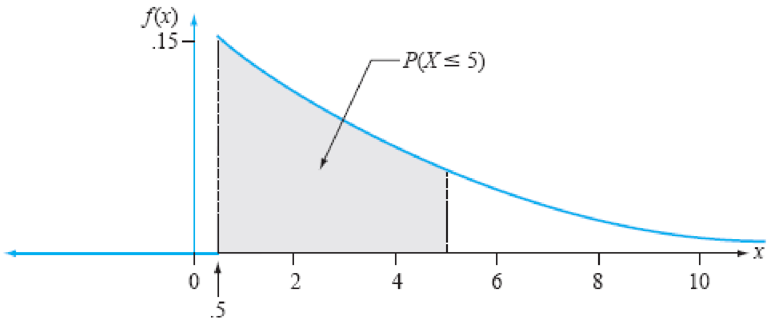
b) $P(X \leq 5) = ?$



Função Densidade de Probabilidade

Exercício: Solução

b) $P(X \leq 5) = ?$



$$P(X \leq 5) =$$

$$= \int_{-\infty}^{5} f(x)dx = \int_{-\infty}^{0.5} f(x)dx + \int_{0.5}^{5} f(x)dx$$

$$= 0 + \int_{0.5}^{5} e^{-0.15(x-0.5)} dx$$

$$= 0.15e^{0.75}\left(-\frac{1}{0.15}e^{-0.15x}\Big|_{x=0.5}^{x=5}\right)$$

$$= 0.491$$

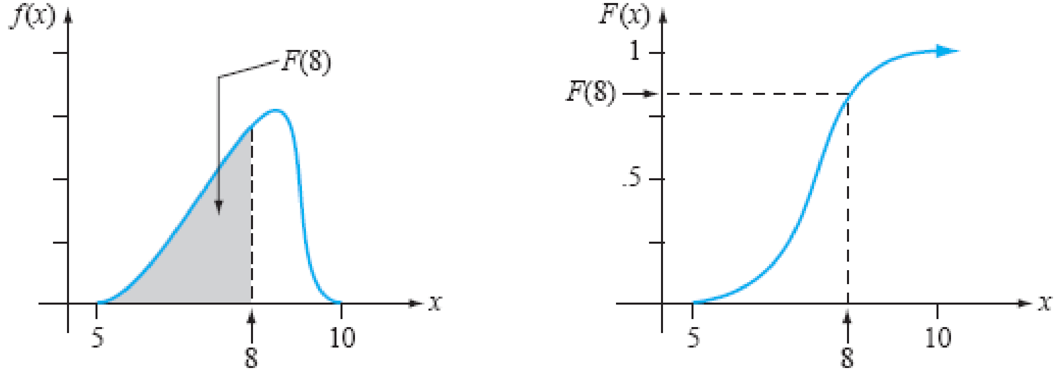
Função Distribuição Acumulada

Def: A **função de distribuição acumulada** $F(x)$ de uma v.a. contínua X é definida para todo x real:

Definição!

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(v)dv$$

A FDA em x, $F(x)$, é a "soma" da probabilidade de todos os valores menores do que x! Para $x = 8$ na figura abaixo:



Função Distribuição Acumulada

!

Propriedade: uma função de distribuição acumulada $F(x)$, definida por:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(v)dv$$

Tem as seguintes propriedades:

- $\lim_{x\to-\infty} F(x) = 0$ e $\lim_{x\to\infty} F(x) = 1$
- $F(x)$ é não-decrescente

Funções Distribuição

Exercício: Seja X a espessura de uma determinada Capa de metal, com fdp:

$$f(x) = \frac{1}{B-A}, \qquad para\ A \leq x \leq B$$

a) Determine a função distribuição acumulada de X.

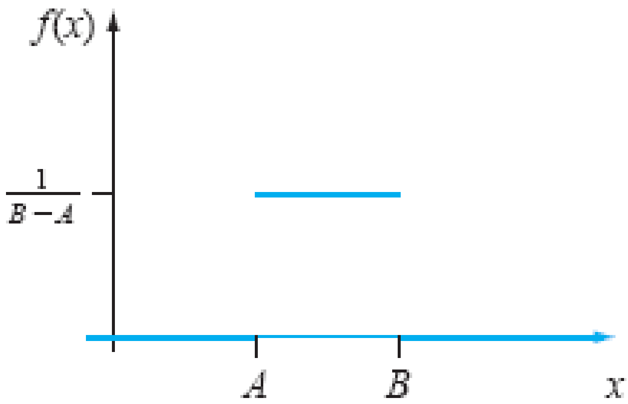
b) Faça o grafico da fdp de $X, f(x)$, e da FDA de $X, F(x)$.

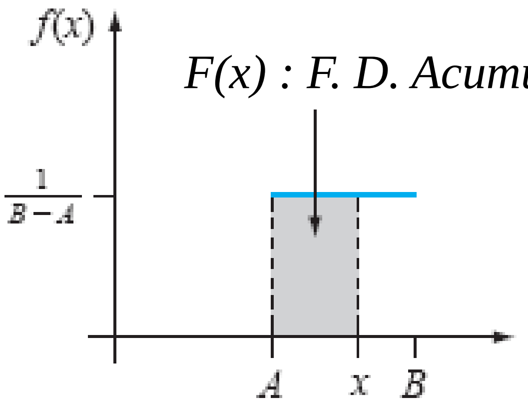
Veremos na aula seguinte que esta distribuição de probabilidade, de tão usada, tem um nome específico: **distribuição uniforme**. Ela é o equivalente a probabilidade clássica (eventos equiprováveis) para o caso contínuo.

Funções Distribuição

Exercício: Solução

$$f(x) = \frac{1}{B-A}, \qquad para\ A \le x \le B$$





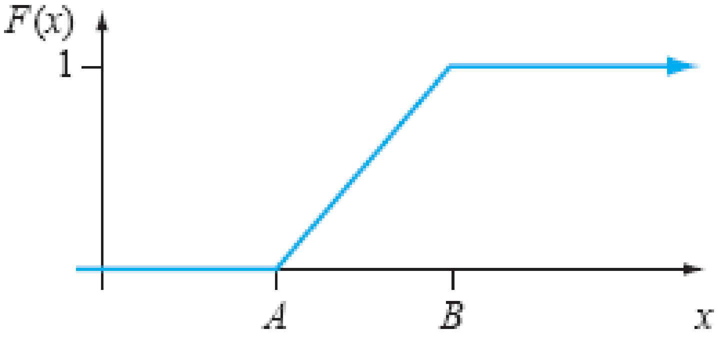
Funções Distribuição

Exercício: Solução

$$F(x) = \int_{-\infty}^{x} f(v)dv = \int_{-\infty}^{x} \frac{1}{B-A} dv.$$

Daí:

$$F(x) = \begin{cases} \int_{-\infty}^{x} 0dv = 0, & x < A \\ \int_{-\infty}^{x} \frac{1}{B-A} dv = \frac{x-A}{B-A}, & A \leq x \leq B \\ \int_{-\infty}^{x} \frac{1}{B-A} dv = 1, & x \geq B \end{cases}$$

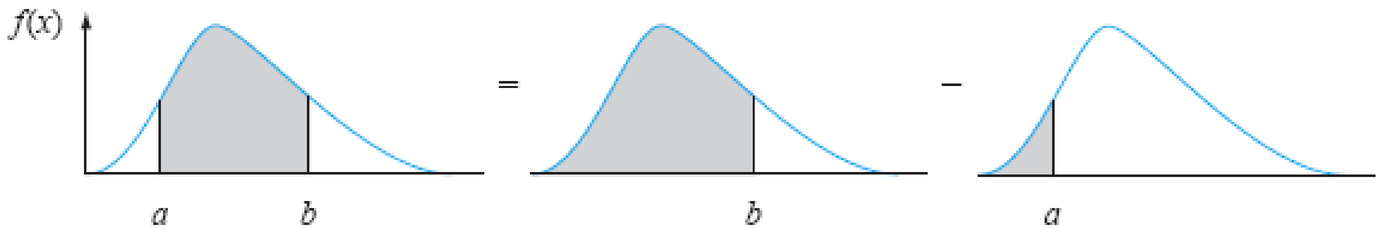


F(X): Probabilidades

Propriedade: Seja X uma variável aleatória contínua com fdp $f(x)$ e $F(x)$. Então, para a e b reais:

- $P(X > a) = 1 - F(a)$

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

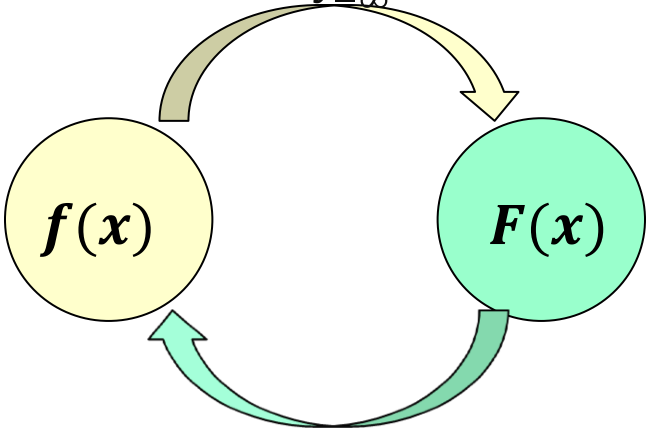


$F(b)$ $F(a)$

Relação entre fdp e FDA

- Se quisermos saber a probabilidade de um evento, por exemplo, $P(a < X < b)$, tanto podemos usar a f. densidade quanto a f. Acumulada.
- A fdp e FDA tem a seguinte relação:

$$F(x) = \int_{-\infty}^{x} f(v)dv$$



$$f(x) = \frac{dF(x)}{dx}$$

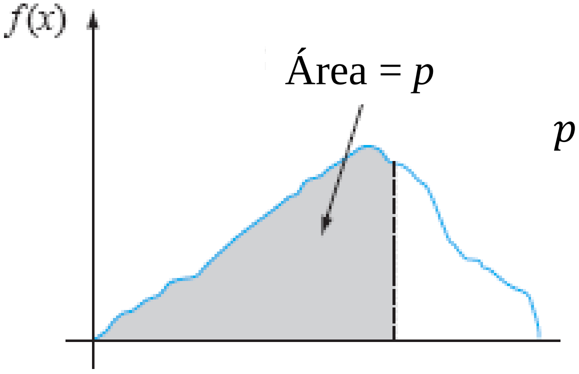
Percentil

Seja p um numero entre 0 e 1 (uma probabilidade)

O **percentil 100*p %** da distribuição de uma v.a. contínua X, representado por x_p, é definido por:

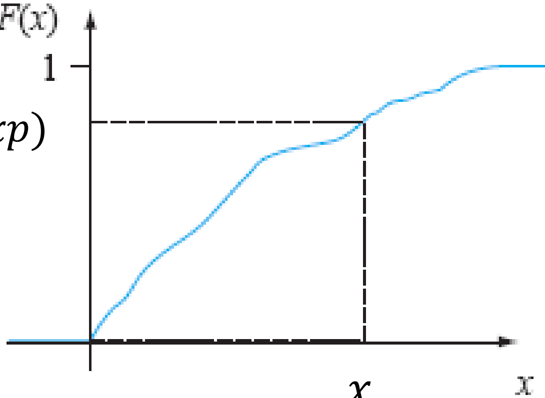
Definição!

$$p = P(X \leq x_p) = F(x_p) = \int_{-\infty}^{x_p} f(v)dv$$



x_p

$p = F(xp)$



x_p

Percentil

Seja p um numero entre 0 e 1 (uma probabilidade)

O **percentil 100% p,** x_p, é definido por:

$$p = P(X \leq x_p) = F(x_p)$$

Assim, o percentil-p, é o valor da variável aleatória que:

- deixa $100 * p\%$ dos valores da v.a. X abaixo dele;
- é maior do que $100 * p$ % dos valores da v.a. X;
- É ultrapassado (ou excedido) por $(1 - p) * 100\%$ dos valores de X.

Ex: Para $p = 0.75$, então $x_{0.75} = Q_3$, ou seja, $x_{0.75}$ é o 3º quartil.
Mas, $x_{0.75}$ é:

- maior do que 75% dos dados (deixa 75% dos valores abaixo dele)
- é excedido por apenas 25% dos dados (deixa 25% dos valores acima dele).

FDA e Percentil

Exercício: A distribuição da quantidade de cascalho (em toneladas) vendidas por uma empresa de materiais de construção em uma semana é uma v.a. contínua X com fdp:

$$f(x) = \frac{3}{2}(1 - x^2), \qquad 0 \leq x \leq 1.$$

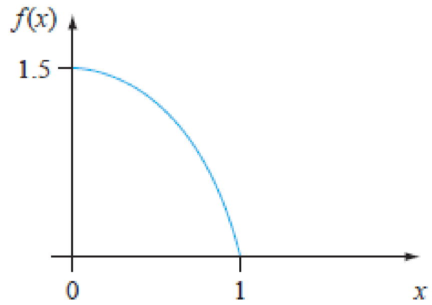
a) Determinar a função distribuição acumulada, $F(X)$.

b) Encontre a mediana da quantidade de cascalho vendida em uma semana (percentil 50%).

FDA e Percentil

Exercício: Solução.

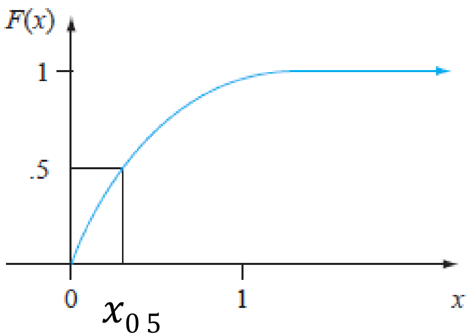
a) Para x entre 0 e 1:



$$F(x) = \int_0^x \frac{3}{2}\left(1 - y^2\right) dy$$

$$= \frac{3}{2}\left(y - \frac{y^3}{3}\right)\Bigg|_{y=0}^{y=x} = \frac{3}{2}\left(x - \frac{x^3}{3}\right)$$

b) Percentil 50% (mediana): $x_{0.5}$



$$p = F(x_p) = \frac{3}{2}\left(x_p - \frac{{x_p}^3}{3}\right)$$

$$0.5 = \frac{3}{2}\left(x_{0.5} - \frac{{x_{0.5}}^3}{3}\right) \Rightarrow {x_{0.5}}^3 - 3x_{0.5} + 1 = 0 \Rightarrow x_{0.5} = 0.347$$

Valor Esperado

- O **valor médio ou esperado** de uma v.a. contínua X com função densidade de probabilidade, $f(x)$, é:

Definição!

$$\boldsymbol{\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx}$$

Novamente o valor esperado de uma v.a. X é a média ponderada de cada valor possível x pela "probabilidade" de o valor sair.

- Para uma função qualquer $h(.)$:

$$\boldsymbol{\mu_{h(X)} = E\big(h(X)\big) = \int_{-\infty}^{\infty} h(x) f(x) dx}$$

- e $E(h(X)) \neq h(E(X))$ para toda função $h(.)$ não linear

Valor Esperado

Exercício: A distribuição da quantidade de cascalho (em toneladas) vendidas por uma empresa de materiais de construção em uma semana é uma v.a. contínua X com fdp:

$$f(x) = \frac{3}{2}(1 - x^2), \qquad 0 \leq x \leq 1$$

Calcule o valor esperado da quantidade de cascalho (ou a quantidade esperada de cascalho) a ser vendida pela empresa na próxima semana.

Valor Esperado

Exercício: Solução.

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{1} x * \frac{3}{2}(1 - x^2)dx$$

$$= \frac{3}{2}\int_{0}^{1}(x - x^3)dx$$

$$= \frac{3}{2}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Bigg|_{x=1} - \frac{3}{2}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Bigg|_{x=0} = \frac{3}{2}\left(\frac{1}{2} - \frac{1}{4}\right)$$

$$= 0.375$$

Valor Esperado de função

Exercício: Duas espécies estão competindo em uma região pelo controle de uma quantidade limitada de um determinado recurso.

Seja X = a proporção de recurso controlado pela espécie 1 e suponha que X tem fdp:

$$f(x) = 1, para\, 0 \leq X \leq 1.$$

A espécie que controla a maioria dos recursos controla:

$$h(X) = \max(X, 1 - X).$$

Determine $E(h(X))$.

Valor Esperado de função

Exercício: Solução

$$h(X) = \begin{cases} 1 - X, \;\; se\ 0 \leq X < 0.5 \\ \quad X, \;\; se\ 0.5 \leq X \leq 1 \end{cases}$$

O valor esperado dos recursos controlados pela espécie "dominante" é:

$$E\big(h(X)\big) = \int_0^1 h(x)f(x)dx = \int_0^{0.5} (1-x)dx + \int_{0.5}^1 x = \frac{3}{4}$$

Variância

- A **variância** de uma v.a. contínua X com fdp $f(x)$ é:

Definição!

$$V(X) = \sigma^2 = E\left((X-\mu)^2\right) = \int_{-\infty}^{\infty} \boldsymbol{(x-\mu)^2 f(x) dx}$$

- E também vale a propriedade:

$$V(X) = \sigma^2 = E(X^2) - \mu^2$$

Propriedades

Propriedade: Continuam valendo as seguintes propriedades para a média (valor esperado) ou variância de uma função linear:

$$E(aX + b) = aE(X) + b$$

$$V(aX + b) = a^2V(X)$$

Variância

Exercício: A distribuição da quantidade de cascalho (em toneladas) vendidas por uma empresa de materiais de construção em uma semana é uma v.a. contínua X com fdp:

$$f(x) = \frac{3}{2}(1 - x^2), \qquad 0 \leq x \leq 1$$

Calcule a variância da quantidade de cascalho vendida pela empresa na próxima semana.

Variância

Exercício: Solução.

No exercício anterior obtemos: $E(X) = \mu = 0.375$

Vamos usar a propriedade: $V(X) = E(X^2) - \mu^2$

Com:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2 * \frac{3}{2}(1 - x^2)dx$$

$$= \frac{3}{2}\int_0^1 (x^2 - x^4)dx = \frac{3}{2}\left(\frac{x^3}{3} - \frac{x^5}{5}\right)\Bigg|_{x=0}^{x=1} = \frac{3}{2}\left(\frac{1}{3} - \frac{1}{5}\right)$$

$$= 0.2$$

Assim: $V(X) = 0.2 - 0.375^2 = 0.059$

Resumo

Definição	**Discretas**	**Contínuas**
Função probabilidade	fmp: $p(x)$	fdp: $f(x)$
F. Probabilidade Acumulada	$F(x) = \sum_{x_i \leq x} p(x_i)$	$F(x) = \int_{-\infty}^{x} f(v)dv$
Probabilidade de X pertencer a um intervalo $[a, b]$	$P(a \leq X \leq b) = \sum_{a \leq x_i \leq b} p(x_i)$ $= F(b) - F(a^-)$	$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$ $= F(b) - F(a)$
Valor esperado	$E(X) = \mu = \sum_{i=1}^{\infty} x_i p(x_i)$	$E(X) = \mu = \int_{-\infty}^{\infty} xf(x)dx$
Valor esperado de uma função, $h(X)$	$E(h(X)) = \sum_{i=1}^{\infty} h(x_i)p(x_i)$	$E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$
Variância	$V(X) = \sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 p(x_i)$	$V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx$

Resumo

Nesta parte vimos:

- Como representar modelos de probabilidade para variáveis aleatórias contínuas:
  - Função densidade de probabilidade, $f(x)$
  - Função distribuição acumulada, $F(x)$
- As propriedades da fdp ($f(x)$) e FDA ($F(x)$)
- Como obter fdp a partir de uma FDA e vice-versa
- Como calcular o p-percentil de uma v.a. contínua
- A média e variância de uma v.a. contínua