



UNIVERSIDADE  
**AbERTA**  
www.uab.pt

**Educational Resources / Recursos Educativos**

# **Inferência Causal**

**Luís Cavique**  
DCeT, Univ. Aberta  
Luis.Cavique@uab.pt

**Lisboa, outubro 2025**



Este trabalho está licenciado com uma Licença Creative Commons  
Attribution-NonCommercial-ShareAlike CC BY-NC-SA

## **Resumo**

Este documento pretende complementar a bibliografia da unidade curricular Extração de Conhecimento de Dados (Data Mining) oferecida nos cursos de mestrado MEIW (Mestrado em Engenharia Informática e Tecnologia Web) e MBB (Mestrado em Bioestatística e Biometria).

A inferência causal quantifica o efeito de uma intervenção. O problema central é que não é possível observar simultaneamente os resultados com e sem tratamento, exigindo métodos que estimem efeitos contrafactuais.

No conjunto de dados do Titanic, que contém informações sobre passageiros e sobrevivência, é possível analisar quais fatores influenciam a chance de sobrevivência.

São usadas duas abordagens. As abordagens contrafactuais, como NATE, T/S-Learner, Duplamente Robusto e MICE, estimam resultados que não podem ser observados diretamente, enquanto métodos de estratificação, como o Matching, comparam indivíduos semelhantes para reduzir viés.

A inferência causal permite, assim, compreender relações de causa e efeito mesmo em dados observacionais, fornecendo insights confiáveis sobre fatores determinantes de desfechos importantes.

### **Conteúdos**

1. Causalidade: descoberta e inferência
2. Problema fundamental da inferência causal
3. Conjunto de dados Titanic
4. Abordagens de inferência causal
  - 4.1. Contrafactuais: NATE, T/S-learner, DR, MICE
  - 4.2. Estratificação: Matching
  - 4.3. Resultados
5. Conclusões

## 1. Introdução à Causalidade

A controvérsia Rubin (1974) versus Pearl (2000) é um dos debates mais importantes na inferência causal, envolvendo duas abordagens diferentes para lidar com a causalidade.

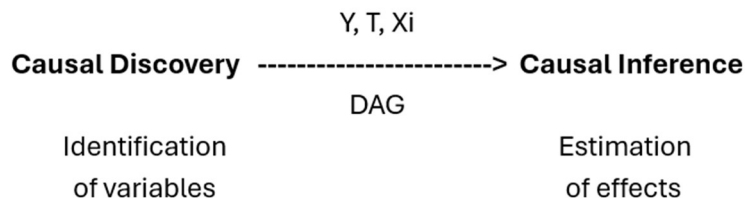
Donald Rubin desenvolveu o "Rubin Causal Model" (RCM), também conhecido como framework de resultados potenciais, utiliza dados contrafactuais onde para cada unidade existe um resultado potencial sob tratamento e outro sob controlo. Rubin utiliza uma linguagem estatística tradicional, mantendo ceticismo em relação aos modelos baseados em grafos, que considera úteis, mas não essenciais.

Por outro lado, Judea Pearl desenvolve uma abordagem baseada em grafos causais, introduzindo Directed Acyclic Graphs (DAGs) como representação visual das relações entre variáveis. Propôs uma hierarquia da causalidade que vai da associação, passando pela intervenção e finalizando nos contrafactuais. Definiu o do-calculus como formalismo para manipular distribuições sob intervenções, para determinar os efeitos causais. Pearl critica a dependência exclusiva do design experimental ao argumentar que a randomização sozinha é insuficiente sem uma compreensão causal adequada das relações subjacentes.

Aqui defendemos uma abordagem híbrida, usar DAGs para identificar as variáveis e depois estimar com métodos de Rubin. Vários investigadores defendem explicitamente essa abordagem híbrida, tais como Miguel Hernán (epidemiologia) e Guido Imbens (prémio Nobel da economia 2021).

O resultado  $Y$  com base no tratamento  $T$  é o objeto do estudo, onde resultado  $Y$  é uma variável dependente e  $T$  da variável independente. Para tanto, são utilizados conjuntos de dados de teste e controlo para tratamento realizado,  $T=1$ , e não realizado,  $T=0$ . A função explicativa utiliza três variáveis,  $Y=f(T, X)$  onde  $X$  são as variáveis de controlo ou covariáveis cujo efeito se procura neutralizar com vista a impedir que interfira entre as variáveis independentes e dependentes. Utiliza-se frequentemente a expressão  $Y = f(T|X)$  onde a ideia é isolar o efeito de  $T$  sobre  $Y$  mantendo  $X$  constante ('ceteris paribus').

A Figura 1.1 apresenta a sequência entre a Descoberta Causal e a Inferência Causal. Na descoberta causal identificamos as variáveis ( $Y, T, X$ ) e na inferência causal são estimados os efeitos do tratamento.



**Figura 1.1.** Relação entre a Descoberta Causal e a Inferência Causal

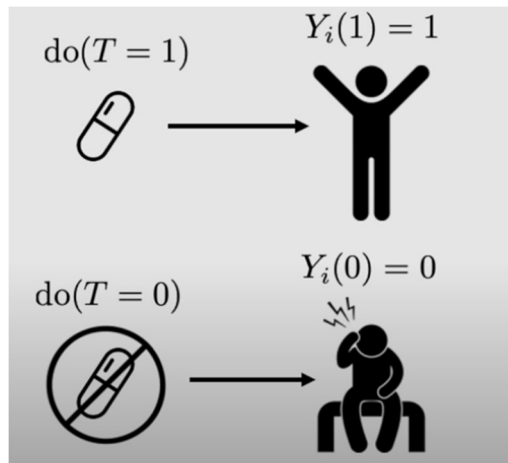
## 2. Problema Fundamental da Inferência Causal

O problema fundamental da inferência causal surge do facto de só podermos observar um resultado potencial para cada unidade individual sob um determinado tratamento, mas não ambos. Esta limitação torna difícil isolar o verdadeiro efeito causal de um tratamento.

O resultado potencial do tratamento é expresso por  $Y(T=1)$  ou  $Y(1)$  ou ainda  $Y^1$ . Por outro lado, para o resultado potencial do controlo usa-se a notação  $Y(T=0)$  ou  $Y(0)$  ou ainda  $Y^0$ .

Na Figura 2.1 temos a possibilidade de tomar uma aspirina para aliviar as dores de cabeça. Se tomar o medicamento o paciente cura a dor de cabeça,  $Y(1)=1$ , e caso contrário mantém a dor de cabeça,  $Y(0)=0$ . Infelizmente, não se podem observar os dois resultados potenciais.

Se ele tomar a aspirina,  $Y(1)$ , corresponde ao resultado factual, sendo  $Y(0)$  o resultado contrafactual. Inversamente, se o paciente não tomar a aspirina,  $Y(0)$ , corresponde ao resultado factual, sendo  $Y(1)$  o resultado contrafactual.



**Figura 2.1.** Valores de  $Y(1)=1$  e  $Y(0)=0$

A Tabela 2.1 mostra os valores dos resultados potenciais do tratamento  $Y(T=1)$  e do controlo  $Y(T=0)$ . Nem todos os valores podem ser preenchidos, dado que só podemos observar para cada unidade individual o tratamento ou o não-tratamento (controlo). O problema dos dados em falta ('missing values') é um desafio na inferência causal.

**Tabela 2.1.** Valores de  $Y(1)$  e  $Y(0)$  com valores em falta

$i$	$T$	$Y$	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	1	?	?

### 3. O dataset Titanic

O naufrágio do paquete Titanic é um dos naufrágios mais conhecidos da história, pelas piores razões. A 15 de abril de 1912, durante a sua viagem inaugural, o RM Titanic, considerado ‘inafundável’, afundou após colidir com um icebergue. Infelizmente, não havia botes salva-vidas suficientes para todos a bordo, resultando na morte de cerca de 70% dos passageiros e tripulantes.

O dataset Titanic (Yasser 2021) original foi limpo e reduzido, contando com os seguintes atributos (Y,T,X). Os resultados potenciais Y corresponde ao atributo binário Survived. O tratamento distingue os passageiros da 1ª classe e os restantes passageiros. Se o passageiro viajava em 1ª classe (PClass), T=1, caso contrário, T=0. Os restantes atributos são Sex, Age, Sibsp (Siblings/Spouses, i.e., Irmãos/Cônjuges) e Parch (Parents/Children, i.e., Pais/Filhos). A Tabela 3.1 apresenta o Dataset Titanic reduzido a seis atributos, que irá ser utilizado nas secções seguintes.

A pergunta de trabalho que se coloca é: será que o facto de viajar em 1ª classe influenciou na sobrevivência do paquete Titanic?

**Tabela 3.1.** Dataset Titanic

	pclass	sex	age	sibsp	parch	survived
1	1	0	29	0	0	1
2	1	1	1	1	2	1
3	1	0	2	1	2	0
4	1	1	30	1	2	0
5	1	0	25	1	2	0
6	1	1	48	0	0	1
7	1	0	63	1	0	1
8	1	1	39	0	0	0
9	1	0	53	2	0	1
10	1	1	71	0	0	0
11	1	1	47	1	0	0
12	1	0	18	1	0	1
13	1	0	24	0	0	1
14	1	0	26	0	0	1
15	1	1	80	0	0	1

#### 4. Abordagens na Inferência Causal

Um Ensaio Clínico Randomizado (Randomized Controlled Trial, RCT) é amplamente considerado o melhor standard para o estabelecimento de relações causais; no entanto, também apresenta desvantagens e limitações significativas. Os RCT são frequentemente dispendiosos, demorados e requerem muitos recursos, tornando-os impraticáveis em muitos contextos do mundo real. Podem também surgir desafios éticos, especialmente quando a recusa de tratamentos potencialmente benéficos a um grupo de controlo não é aceitável. Uma abordagem alternativa é basear-se em dados observacionais, que são informações já recolhidas e propriedade das organizações, fornecendo insights valiosos sem as restrições do desenho experimental.

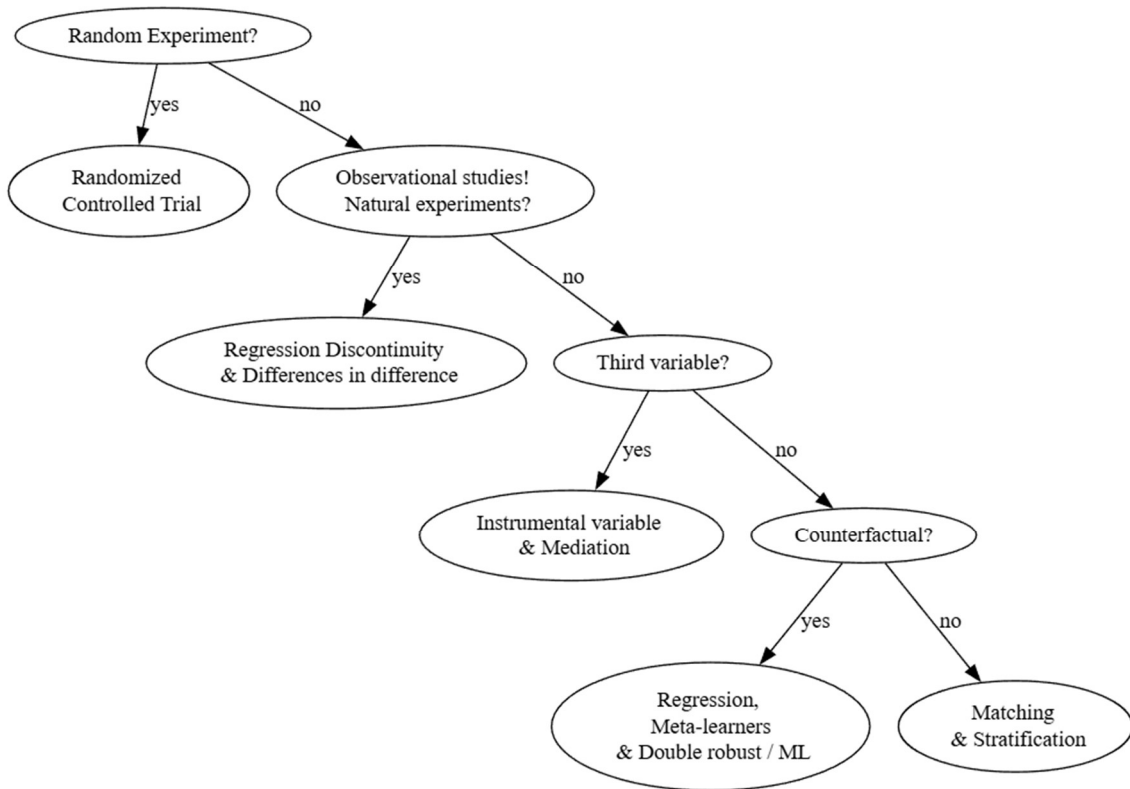
Para extrair alegações causais de dados observacionais, os economistas empregam experiências naturais, utilizando abordagens como a descontinuidade de regressão e a diferença em diferenças (Cunningham, 2021). Estas abordagens quase experimentais, como as experiências naturais e os métodos econométricos relacionados, fornecem alternativas fiáveis para a identificação de efeitos causais.

Para além do par de variáveis (T-tratamento, Y-resultado) existem abordagens com uma terceira variável: a análise de mediação e as variáveis instrumentais. A análise de mediação (Baron, Kenny, 1986) é uma técnica estatística utilizada para compreender como uma variável independente influencia uma variável dependente através de uma variável mediadora (M). As variáveis instrumentais são igualmente uma técnica estatística utilizada para estimar relações causais. Um instrumento (Z) é uma variável que influencia a variável independente do tratamento, mas não tem um efeito direto sobre a variável dependente (Angrist, Krueger, 2001).

Outras abordagens são mais comuns em áreas como a aprendizagem automática e a bioestatística. Na aprendizagem automática, a inferência causal baseia-se frequentemente no raciocínio contrafactual e emprega algoritmos de meta-aprendizagem, como o T-learner, o S-learner ou o X-learner (Künzel et al., 2019). Estes algoritmos permitem a estimativa de efeitos de tratamento heterogêneos através do raciocínio contrafactual, complementando as abordagens econométricas tradicionais.

Por outro lado, a bioestatística emprega métodos de ‘pontuação de propensão’, particularmente o emparelhamento (Zubizarreta et al., 2023), bem como a estratificação. A técnica de emparelhamento ajuda a criar um par de unidades de comparação que se aproxima do contrafactual, combinando unidades tratadas e não tratadas com características semelhantes. Outra estratégia, como o emparelhamento, passa por estratificar ou subclassificar a população, criando estratos com características semelhantes. O emparelhamento e a estratificação aderem ao conceito de *Ceteris Paribus* (sendo todas as outras coisas iguais), tal como referido na inferência causal (Imbens & Angrist, 1994; Angrist & Pischke, 2015). Austin (2011), citando Rosenbaum e Rubin (1984), demonstrou que a divisão dos indivíduos em quintis com base no escore de propensão pode eliminar aproximadamente 90% do viés dos fatores de confusão medidos quando se estima um efeito linear do tratamento.

A Figura 4.1. apresenta uma taxonomia das técnicas de Inferência Causal que identifica os RCT, os experimentos naturais, as abordagens com três variáveis, abordagens com dados contrafactuais e as abordagens de matching/estratificação.



**Figura 4.1.** Taxonomia das abordagens de Inferência Causal

De uma forma mais reduzida teremos:

- caso Experimento Aleatório então Teste de Controlo Aleatório
- caso Experimento Natural então Regressão Descontinuidade ou DiD
- caso Terceira Variável então Variável Instrumental ou Mediação
- caso Contrafactual então Regressão ou Meta-learning
- caso contrário Matching ou Estratificação

Neste trabalho iremos exemplificar as abordagens com contrafactuais e as abordagens de matching.

#### 4.1. Estimaco com Modelos Contrafactuais

Para exemplificar a estimaco com modelos contrafactuais vamos utilizar um subconjunto de dados do dataset Titanic. Na Tabela 4.1 a varivel do tratamento T corresponde a Pclass e a varivel do resultado potencia Y  a varivel Survive.

**Tabela 4.1.** Subconjunto de dados do dataset Titanic

sex	age	sibsp	parch	fare	embarked	pclass =T	survived =Y
0	29	0	0	211	3	1	1
1	1	1	2	152	3	1	1
0	2	1	2	152	3	1	0
1	30	1	2	152	3	1	1
0	25	1	2	152	3	1	1
1	42	0	0	8	3	0	0
1	13	0	2	20	3	0	0
1	16	1	1	20	3	0	0
0	35	1	1	20	3	0	1
0	16	0	0	8	3	0	1

Contrafactuais so cenrios alternativos que respondem a perguntas do tipo ‘e se algo tivesse sido diferente?’. Existe uma relao forte entre os contrafactuais com dados em falta (‘missing data’), pois ambos lidam com situaes hipotticas ou desconhecidas. Assim os cenrios alternativos correspondem aos dados em falta, associando o que poderia ter acontecido com a informao ausente se estivesse disponvel.

Na Tabela 4.2 representa o padro dos dados em falta do dataset Titanic relativos aos sobreviventes, onde Y=1 corresponde aos sobreviventes e Y=0 caso contrrio. A varvel  $Y(T=1)$ ,  $Y(1)$  ou  $Y^1$  corresponde aos resultados para os passageiros de 1ª classe e  $Y(0)$  os resultados para passageiros que viajavam noutras classes.

Dado o princpio fundamental da inferncia causal, para  $Y(1)$  e  $Y(0)$  existem um conjunto de dados conhecidos e um conjunto de dados em falta. Para  $Y(1)|T=1$  e  $Y(0)|T=0$  existem dados disponveis. Por outro lado, para  $Y(1)|T=0$  e  $Y(0)|T=1$  os dados no esto disponveis, pois correspondem a situaes contrafactuais.

**Tabela 4.2.** Padro dos dados em falta do dataset Titanic

T	Y	$Y(T=1) = Y(1)$	$Y(T=0) = Y(0)$	$Y(1)$	$Y(0)$
1	1	1	?	$Y(1) T=1$	missing data $Y(0) T=1$
1	1	1	?		
1	0	0	?		
1	1	1	?		
1	1	1	?		
0	0	?	0	missing data $Y(1) T=0$	$Y(0) T=0$
0	0	?	0		
0	0	?	0		
0	1	?	1		
0	1	?	1		

Nesta secção abordaremos uma técnica naïve e de seguida os modelos T-learner, S-learner, e Double robust, terminando com a utilização do algoritmo MICE desenvolvido para tratar dados em falta.

#### 4.1.a. Naïve ATE

O valor de ATE ('average treatment effect') é dado por  $ATE = E(Y^1 - Y^0) = E(Y^1) - E(Y^0)$ , contudo os valores que temos são os valores de  $E(Y^1|T=1)$  e  $E(Y^0|T=0)$ . Assim, só nos é possível calcular um valor aproximado, o NATE ('naïve average treatment effect').

A Tabela 4.3. apresenta o cálculo do NATE. O Naïve ATE são utilizados só os dados disponíveis, usando a expressão:  $ATE = E[Y^1|T=1] - E[Y^0|T=0]$ .

**Tabela 4.3.** Cálculo do NATE usando os dados disponíveis  $E(Y^1|T=1)$  e  $E(Y^0|T=0)$

Y(1) T=1	Y(0) T=0	NATE
1		
1		
0		
1		
1		
	0	
	0	
	0	
	1	
	1	
4/5	2/5	2/5
$E[Y(1) T=1]$	$E[Y(0) T=0]$	$E[Y(1) T=1] - E[Y(0) T=0]$

O código R para o NATE apresenta-se da seguinte forma, usando o símbolo # para comentário, em que Y=survived:

```
# Naive ATE
EY1_T1= mean (df$survived [df$T==1]) # E[Y(1) | T=1]
EY0_T0= mean (df$survived [df$T==0]) # E[Y(0) | T=0]
print (EY1_T1 - EY0_T0)
# Naive ATE do dataset Titanic com 1309 linhas = 0.314
```

Uma outra forma de calcular o NATE, Lift ou Uplift passa pela utilização de uma matriz de contingência, utilizando as variáveis Y e T:

	Y=0	Y=1	total
T=0	3	2	5
T=1	1	4	5

As probabilidades  $Pr(Y=1|T=0)=2/5$  e  $Pr(Y=1|T=1)=4/5$ , pelo que o 'tratamento é bom' visto que  $Pr(Y=1|T=1)-Pr(Y=1|T=0) > 0$ . O valor de NATE, Lift ou Uplift é igual a 2/5.

#### 4.1.b. T-learner

No T-learner ou Two-learner a ideia base é preencher os dados em falta utilizando modelos preditivos. A Figura 4.2 compara o modelo preditivo com a T-learner. No modelo preditivo, utilizamos os dados já classificados para treinar o modelo e, em seguida, prever a classe de novos dados. O T-learner vai realizar duas previsões para a população tratada  $Y(1)$  e para a população não tratada  $Y(0)$ , conforme ilustrado na Figura 4.3.

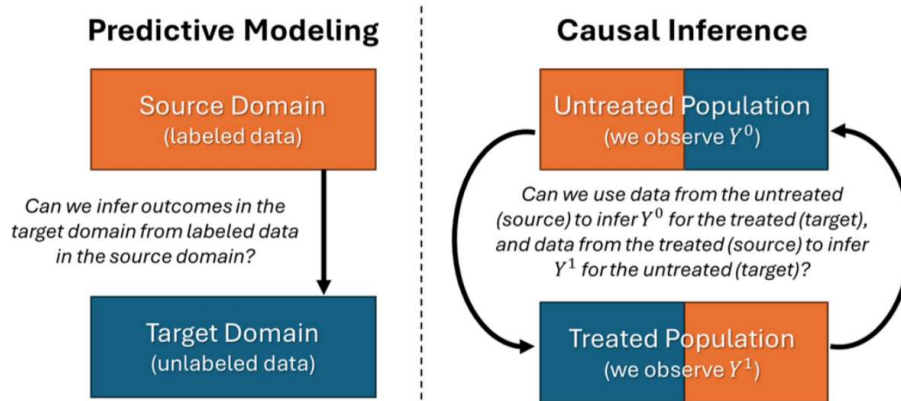


Figura 4.2. Comparação de um modelo preditivo com o T-learner

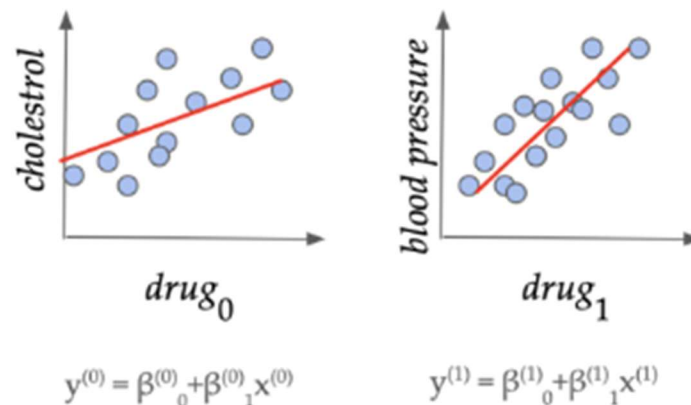


Figura 4.3. Modelos para população não tratada  $Y(0)$  e para a população tratada  $Y(1)$

O código R para o T-learner apresenta-se da seguinte forma, usando o símbolo # para comentário, em que  $Y = \text{survived}$  e ITE corresponde “individual treatment effect”.

```
lm1.out <- lm(survived ~ sex+age+sibsp+parch, data =df[df$treat==1,])
df$Y1_hat=predict(lm1.out, newdata = df)
lm0.out <- lm(survived ~ sex+age+sibsp+parch, data =df[df$treat==0,])
df$Y0_hat=predict(lm0.out, newdata = df)
df$ITE <- df$Y1_hat - df$Y0_hat
print(mean(df$ITE))
# ATE T-learner = 0.293
```

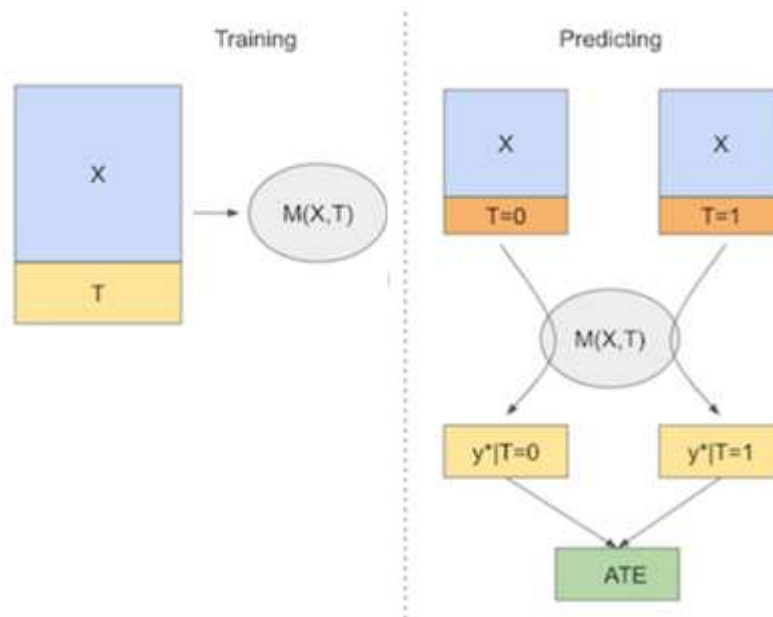
#### 4.1.c. S-learner

O S-learner ou Single-learner, ao contrário do anterior, utiliza um único modelo. O S-learner aprende o resultado esperado tanto com quanto sem tratamento, a partir de dados observacionais.

O modelo aprende a função  $Y^*=f(X, T)$  para estimar o efeito causal individual (ITE) de uma pessoa com características  $X_i$ :

$$\text{ITE}(X_i)=f(X_i, T=1) - f(X_i, T=0)$$

A Figura 4.4 apresenta as fases de aprendizagem e previsão do S-learner. Num primeiro passo é criado um único modelo. No segundo passo são preenchidos os valores em falta (ou contrafactuais) para  $T=0$  e  $T=1$ .



**Figura 4.4.** Fases de aprendizagem e previsão do S-learner

O código R para o T-learner apresenta-se da seguinte forma, usando o símbolo # para comentário, em que  $Y=$ survived:

```
lm.out <- lm(survived ~ sex+age+sibsp+parch+treat, data = df)
data_treated <- df; data_treated$treat <- 1
data_control <- df; data_control$treat <- 0
df$Y1_hat=predict(lm.out, newdata = data_treated)
df$Y0_hat=predict(lm.out, newdata = data_control)
df$SITE <- df$Y1_hat - df$Y0_hat
print(mean(df$SITE))
# ATE S.learner = 0.304
```

#### 4.1.d. Double robust

Sim, o double robust (DR) combina um termo de regressão (modelo de resultado), e um termo de ponderação dos resíduos (IPW, inverse probability weighting).

O modelo é dado por  $E[m_1(X) - m_0(X)]$  e o resíduo, ou erro =  $Y - m(X)$ . O IPW entra como um ajuste (ou correção) ao modelo de regressão. Vimos como combinar regressão linear com o escore de propensão para criar um estimador duplamente robusto. Ele é chamado de Duplamente Robusto porque basta que um dos dois modelos — o de escore de propensão ou o resultado — esteja correto para que possamos identificar o efeito causal. Mais formalmente teremos:

$$\hat{\tau}_{DR} = \underbrace{\mathbb{E}[\hat{m}_1(X) - \hat{m}_0(X)]}_{\text{Outcome Model (regression)}} + \underbrace{\mathbb{E}\left[\frac{T - e(X)}{e(X)(1 - e(X))}(Y - \hat{m}_T(X))\right]}_{\text{Correction (IPW residual)}}$$

onde:

$T_i$  é o tratamento

$m_{T_i}(X_i)$  o modelo preditivo sob o tratamento  $T_i$

$Y_i$  o resultado observado

$e(X_i)$  é o escore de propensão

A Figura 4.5 apresenta o DR. O código R para o Doule Robust apresenta-se da seguinte forma, usando o símbolo # para comentário, em que  $Y = \text{survived}$ :

```
# 1. Propensity score: Pr(T = 1 | X)
modelo_g <- glm(treat ~ sex+age+sibsp+parch, data = df, family = binomial())
df$pscore <- predict(modelo_g, type = "response")
# 2. Modelo do resultado: E[Y | T, X]
modelo_Q <- lm(survived ~ treat+sex+age+sibsp+parch, data = df)
df$y_hat <- predict(modelo_Q)
# 3. DR Estimator (ATE, Average Treatment Effect)
df$dr_2term <- ((df$treat - df$pscore) / (df$pscore * (1 - df$pscore))) * (df$survived - df$y_hat)
ate_dr <- mean(df$y_hat[df$treat == 1]) - df$y_hat[df$treat == 0] + mean(df$dr_2term)
# Resultado final
cat("Estimativa Double Robust (manual com lm/glm):", round(ate_dr, 4), "\n")
# Estimativa Double Robust (manual com lm/glm): 0.311
```

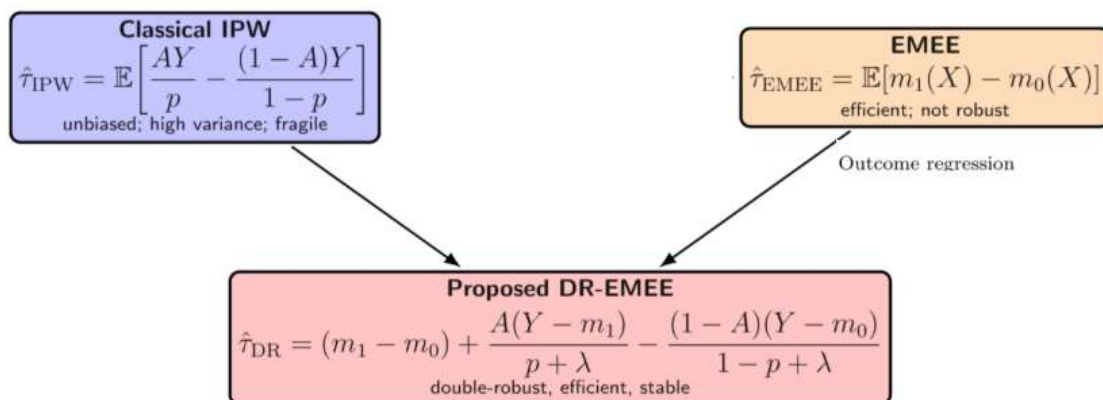


Figura 4.5. Fases do Duplamente Robusto

#### 4.1.e. MICE

Em inferência causal não é costume utilizar algoritmos usados em ‘missing data’. Contudo, neste trabalho utilizando o algoritmo MICE. O algoritmo MICE (Multivariate Imputation by Chained Equations, Imputação Multivariada por Equações Encadeadas) é um método amplamente utilizado para imputar dados em falta.

O MICE funciona criando várias imputações (ou substituições) para os valores em falta, com base em modelos preditivos construídos a partir das outras variáveis disponíveis. O processo ocorre de forma iterativa: cada variável com valores em falta é modelada em função das demais, e os valores imputados são atualizados sucessivamente até que o algoritmo convirja para uma solução estável. O resultado são múltiplos conjuntos de dados completos. O pseudocódigo de MICE é apresentado no Algoritmo 1.

---

**Algorithm 1: MICE Algorithm**

---

```
Input: Features to be imputed; Prediction Models  
Output: Imputed Features, MICE Model  $\Phi$   
/* The imputation can be based upon the percentage of missing  
   values or at random. */  
select Imputation Order;  
 $\Phi \leftarrow$  list;  
/* initialise imputation chain */  
for each feature in input do  
  | fill missing values using a starting guess;  
end  
while change in predictions > tolerance do  
  /* The input can be sorted randomly each loop or from the  
   least missing to most missing feature. */  
  sort input;  
  for each feature in input do  
    | train predictor;  
    | impute missing values in feature;  
    | append trained predictor to  $\Phi$ ;  
    | update input;  
  end  
end
```

---

O código R para o MICE apresenta-se da seguinte forma, usando o símbolo # para comentário, em que Y=survived:

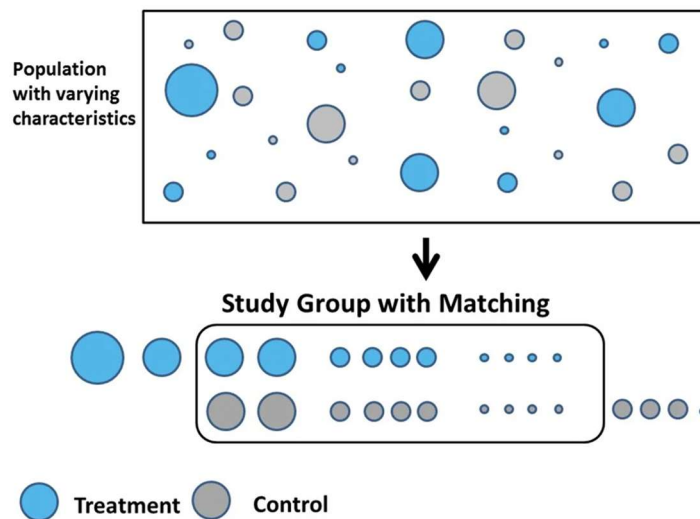
```
# MICE #install.packages("mice")  
library(mice); md.pattern(df)  
tempData <- mice(df,m=5, maxit=50, meth='pmm', seed=500)  
summary(tempData)  
completedData <- complete(tempData,1)  
md.pattern(completedData)  
completedData$ATE=completedData$Y1_hat - completedData$Y0_hat  
print(mean(completedData$ATE))  
# ATE MICE= 0.296
```

## 4.2. Matching

O *matching* e a estratificação são técnicas padrão para reduzir o viés em estudos observacionais.

O *matching* de pares de unidades tratadas e de controlo com covariáveis semelhantes, visando simular a randomização. A estratificação divide a amostra em subgrupos com base nas covariáveis, garantindo que as comparações são feitas dentro de grupos mais homogéneos. Austin (2011), citando Rosenbaum e Rubin (1984), demonstrou que a divisão dos indivíduos em quintis com base no escore de propensão (propensity score) pode eliminar aproximadamente 90% do viés dos fatores de confusão medidos quando se estima um efeito linear do tratamento.

A Figura 4.6 ilustra o procedimento de Matching. São escolhidos os itens semelhante no tratamento e no controlo, sendo as restantes removidas da população.

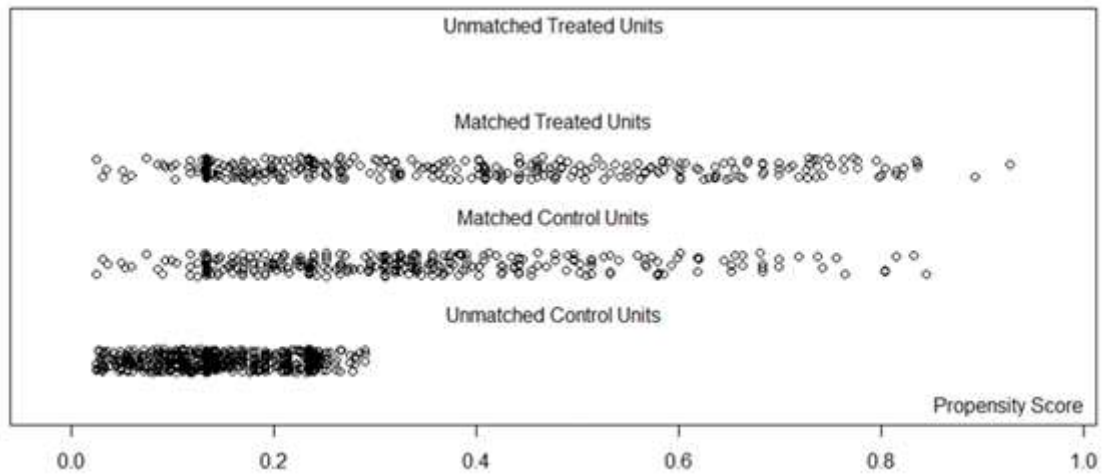


**Figura 4.6.** Procedimento de Matching

O código R para o Matching apresenta-se da seguinte forma, usando o símbolo # para comentário, em que Y=survived:

```
# match 2 conjuntos, com Y=survived, T=treat=pclass
library(MatchIt) # matchit (treatment ~ covariates)
m.out = matchit (treat ~ sex+age+sibsp+parch, data=df, method="nearest", ratio=1)
summary(m.out)
plot(m.out, type = "jitter", interactive=FALSE)
# usando os dados emparelhados
m.data = match.data(m.out)
ATE <- with(m.data, mean(survived[treat == 1]) - mean(survived[treat == 0]))
print(ATE)
# ATE MatchIt= 0.294
```

A Figura 4.7 apresenta a distribuição dos Propensity Scores, com itens emparelhados e não emparelhados, como resultado do comando Plot.



**Figura 4.7.** Distribuição dos PS com itens emparelhados e não emparelhados

A Tabela 4.4. apresenta os pares de dados emparelhados (subclass) com Propensity Score semelhante. O algoritmo de Matching minimiza a diferença (distance) entre os Propensity Scores dos indivíduos tratados e do controle, garantindo comparabilidade entre os grupos e reduzindo o viés da seleção.

**Tabela 4.4.** Dados emparelhados (subclass)

survived	treat	distance	weights	subclass
1	1	0.30495164	1	1
0	0	0.29481640	1	1
1	1	0.02549531	1	2
1	0	0.02549531	1	2
0	1	0.05310200	1	3
1	0	0.05340342	1	3
0	1	0.17188723	1	4
0	0	0.17030115	1	4
0	1	0.22472136	1	5
1	0	0.22576467	1	5
1	1	0.46060067	1	6
1	0	0.36722632	1	6
1	1	0.83621327	1	7
0	0	0.81528005	1	7
0	1	0.30987851	1	8
0	0	0.29481640	1	8

### 4.3. Resultados

Existem dois tipos de abordagens a abordagem contrafactual com seis algoritmos e a abordagem de estratificação com um algoritmo. A Tabela 4.5 apresenta os resultados ATE nos vários algoritmos. A média geral do ATE ronda os 30%, o desvio padrão é de 0,009 e o coeficiente de variação é de 3%, o que indica que os algoritmos um têm desempenho semelhante, sugerindo que os resultados são consistentes e estáveis, dada a baixa variabilidade.

**Tabela 4.5.** Resultados ATE nos vários algoritmos

Type	Algorithm	ATE
Counterfactual	Naive ATE	0,315
	T-learner	0,293
	S-learner	0,304
	Double robust	0,311
	MICE	0,296
Stratification	Matching	0,294

### 5. Conclusões

Este trabalho explorou diferentes abordagens em Inferência Causal para estimar efeitos causais utilizando o dataset Titanic, focando-se na questão: será que viajar em primeira classe influenciou a sobrevivência no naufrágio do Titanic?

O trabalho apresenta uma nova taxonomia das técnicas de inferência causal que permite uma visão holística da área e facilita a visão do analista.

O problema fundamental da inferência causal impõe que nunca podemos observar simultaneamente os resultados potenciais  $Y(1)$  e  $Y(0)$  para o mesmo indivíduo, criando um desafio metodológico que foi abordado através de duas estratégias principais: modelos contrafactuais e técnicas de matching.

Os resultados obtidos demonstram notável consistência entre os diferentes métodos. Todos os algoritmos testados, desde o NATE (0,314) aos métodos mais sofisticados como T-learner (0,293), S-learner (0,304), Double Robust (0,311), MICE (0,296) e MatchIt (0,294), convergem para um efeito médio de tratamento (ATE) em torno de 30%. Esta convergência é confirmada pelo baixo coeficiente de variação de apenas 3%, indicando robustez e estabilidade dos resultados.

A resposta à pergunta de trabalho é que viajar em primeira classe teve um impacto causal positivo e substancial na probabilidade de sobrevivência, aumentando-a em aproximadamente 30 pontos percentuais.

Do ponto de vista metodológico, este trabalho demonstra que tanto as abordagens baseadas em raciocínio contrafactual quanto as técnicas de emparelhamento podem ser eficazes para extrair inferências causais de dados observacionais. O estimador Double Robust destaca-se pela sua propriedade de proteção dupla, enquanto o Matching oferece maior interpretabilidade.

Este estudo reforça ainda a importância de utilizar múltiplas abordagens para validar conclusões causais, especialmente quando se trabalha com dados observacionais onde a randomização não foi possível.

## **Bibliografia**

Angrist J., A. Krueger (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, vol.15 (4): 69–85. doi:10.1257/jep.15.4.69. hdl:1721.1/63775.

Angrist J.D., J.-S. Pischke (2015). *Mastering Metrics: The Path from Cause to Effect*, Princeton University Press, ISBN-13: 978-0691152844.

Austin, P.C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies, *Multivariate Behavioral Research*, vol. 46 (3): 399-424, doi: 10.1080/00273171.2011.568786.

Baron R.M., D.A. Kenny (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, *Journal of Personality and Social Psychology*, vol. 51 (6), pp. 1173–1182, <https://doi.org/10.1037/0022-3514.51.6.117>.

Cunningham S. (2021). *Causal Inference: The Mixtape*, Yale University Press, ISBN-13: 978–0300251685.

Imbens, G.W., Angrist, J.D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, vol. 62(2), pp. 467–475. <https://doi.org/10.2307/2951620>.

Künzel S.R., J.S. Sekhon, P.J. Bickel, & B. Yu (2019), Metalearners for estimating heterogeneous treatment effects using machine learning, *Proc. Natl. Acad. Sci. U.S.A.* 116 (10) 4156-4165, <https://doi.org/10.1073/pnas.1804597116>.

Pearl J. (2000), *Causality: models, reasoning, and inference*, Cambridge University Press.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, vol. 79(387), 516–524. <https://doi.org/10.2307/2288398>.

Rubin D.B. (1974), Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, vol. 66(5), pp. 688–701, <https://doi.org/10.1037/h0037350>.

Yasser M. (2021), *Titanic Dataset: Titanic Survival Prediction Dataset*, Kaggle, <https://www.kaggle.com/datasets/yasserh/titanic-dataset>

Zubizarreta J.R., E.A. Stuart, D.S. Small, P.R. Rosenbaum, Eds (2023). *Handbook of Matching and Weighting Adjustments for Causal Inference*, *Handbooks of Modern Statistical Methods*, Chapman & Hall/CRC, ISBN-13 9780367609528.