Análise de dados COVID-19 em Portugal

Analysis of Portuguese COVID-19 data

by

João F. Pereira

7 Outubro 2022

Índice

1. Introdução
   - Modelos Epidemiológicos
2. Análise Exploratória
   - Obtenção dos dados
   - Processamento
   - Transformação
   - Exploração
   - Resultados
3. Demonstração Prática
   - Some closing thoughts
4. Conclusões

Introdução

jn.pt, 23 Março 2021. Fonte: Lusa

COVID-19 inCTRL

Financiado pelo programa:
RESEARCH4COVID - FCT

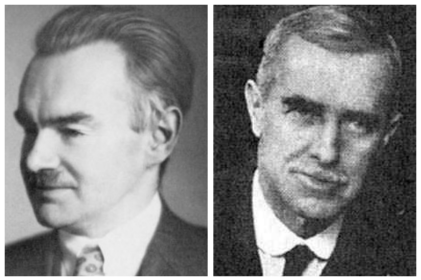
utad UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO

Instituto_Nacional de Saúde Doutor Ricardo Jorge

Modelos Epidemiológicos

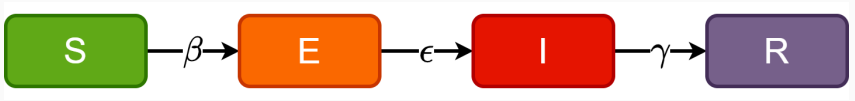
Modelo SIR

$$\begin{cases} \frac{dS}{dt} = -\beta \frac{SI}{N} \\ \frac{dI}{dt} = \frac{SI}{N} - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$



W.O. Kermack e A.G. McKendrick

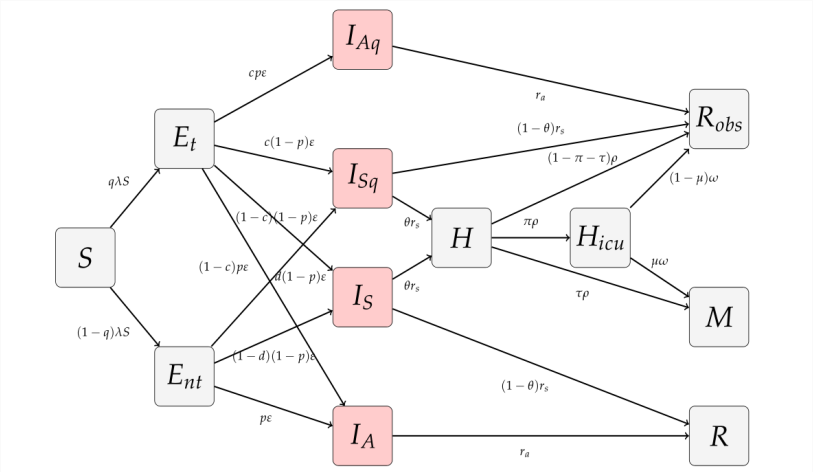
Modelo SEIR

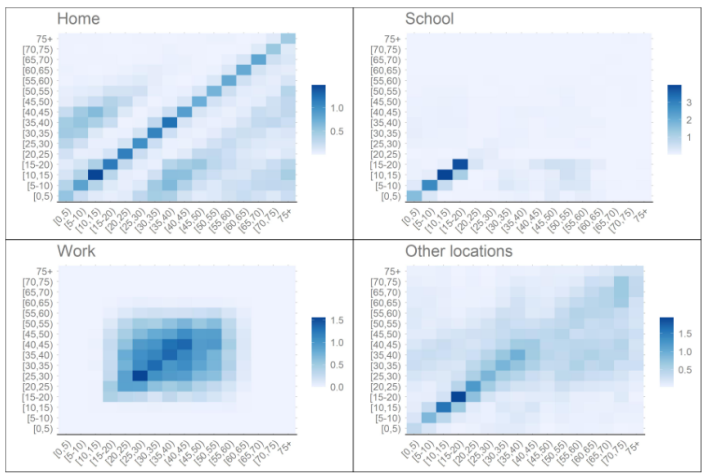


$$\begin{cases} \frac{dS}{dt} = -\beta \frac{SI}{N} \\ \frac{dE}{dt} = \beta \frac{SI}{N} - \epsilon E \\ \frac{dI}{dt} = \epsilon E - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$

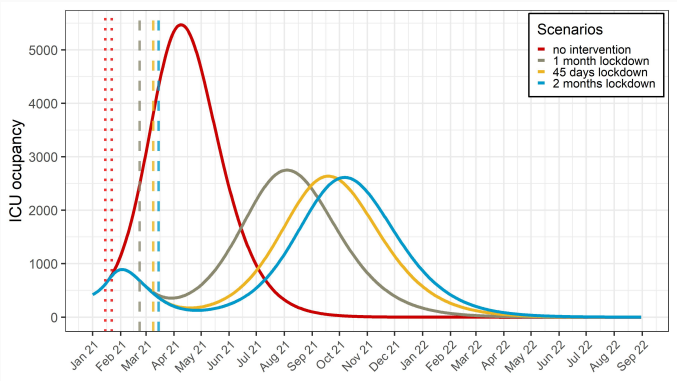
Modelos Epidemiológicos

Modelo COVID-19 inCTRL





$$
\begin{aligned}
S' &= -\lambda S, \\
E_t' &= q\lambda S - \epsilon E_t \\
I_A' &= p\epsilon E_{nt} + (1-c)p\epsilon E_t - r_a I_A \\
I_S' &= (1-d)(1-p)\epsilon E_{nt} + (1-c)(1-p)\epsilon E_t \\
&\quad - r_s I_S \\
E_{nt}' &= (1-q)\lambda S - \epsilon E_{nt} \\
I_{Aq}' &= cp\epsilon E_t - r_a I_{Aq} \\
I_{Sq}' &= c(1-p)\epsilon E_t + d(1-p)\epsilon E_{nt} - r_s I_{Sq} \\
H' &= \theta r_s (I_S + I_{Sq}) - \rho H \\
H_{icu}' &= \pi\rho H - \omega H_{icu} \\
M' &= \mu\omega H_{icu} + \tau\rho H \\
R_{obs}' &= (1-\theta) r_s I_{Sq} + r_a I_{Aq} + (1-\pi-\tau)\rho H \\
&\quad + (1-\mu)\omega H_{icu} \\
R' &= (1-\theta) r_s I_S + r_a I_A
\end{aligned}
$$



Análise exploratória

Como realizar uma Análise exploratória?

Análise exploratória

Como realizar uma Análise exploratória?

1. Obtenção dos dados

Análise exploratória

Como realizar uma Análise exploratória?

1. Obtenção dos dados
2. Processamento

Análise exploratória

Como realizar uma Análise exploratória?

1. Obtenção dos dados
2. Processamento
3. Transformação

Análise exploratória

Como realizar uma Análise exploratória?

1. Obtenção dos dados
2. Processamento
3. Transformação
4. Exploração

Análise exploratória

Como realizar uma Análise exploratória?

1. Obtenção dos dados
2. Processamento
3. Transformação
4. Exploração
5. Resultados

Análise exploratória

Como realizar uma Análise exploratória?

1. Obtenção dos dados
2. Processamento
3. Transformação
4. Exploração
5. Resultados

Obtenção dos dados

Processamento

Objetivo: Transformar os dados originais para outros mais utilizáveis e verificar possíveis incongruências.

1. Alteração de designações
2. Transformação de variáveis (ex: idade, grupos etários, LoS, etc.)
3. Problema de representação
4. Remoção de variáveis sem interesse

Dados Originais

Variable	Description
Health Institution	Hospital where the entry took place
Gender	Gender of the patient
Date of birth	Date of birth of the patient
Hospital process	Process identification number
Hospital episode	Episode identification number
Destination after discharge	Includes 21 types of discharges, including fatality
Date of entry	Date of entry of the patient in hospital
Admission to ICU	If patient received Intensive Care Unit (ICU) treatment during hospitalization
Type of episode	Designator of the length of stay in hospital
Diagnostic code	Code of the diagnosis, according with APR-DRG regulations
Diagnosis	The diagnosis corresponding to the diagnostic code
Order of diagnosis	The order the patient arrived in that day
Date of discharge	Date when the patient was discharged

Dados Processados

Variable	Designation	Levels	Description
Region	Region	Norte, Centro, LVT, Alentejo and Algarve	General Health Administration (ARS) of the health institution.
Type of hospitalization	Type	nonICU and ICU	Non-ICU patients did not receive intensive care unit (ICU) treatment, ICU patients did.
Age	Age	Whole numbers from 0 to 105	Age of the patient, at the time of entry.
Age Group	Group	[0,5), [5,10), (...), [75,80), [80,150)	Designated age group of the patient, based on age at entry.
Gender	Gender	Feminine and Masculine	Gender of the patient.
Entry date	Entry.date	date in format "%d/%m/%Y"	Date of patient's entry into the institution.
Discharge date	Discharge.date	date in format "%d/%m/%Y"	Patient discharge date from the health institution.
Outcome	Outcome	Deceased and Discharged	If the patient died in hospital care or was discharged.
Length of stay	Time.stay	Whole numbers	Number of days between the discharge and entry date.
Month of entry	Entry.month	date in format "%m/%Y"	Month of entry in hospital care.
Month of discharge	Discharge.month	date in format "%m/%Y"	Month of discharge from hospital care.

Transformação

Objetivo: Transformar os dados noutros com nova informação.

LoS	mean	median	25%	75%	90%	95%	99%	max	IQR
Non-ICU	12.54	8	4	15	27	38	67	348	11
ICU	24.31	18	11.0	30.3	51.0	66.0	112.0	216	19.3

Dados Processados

Variable	Designation	Levels	Description
Region	Region	Norte, Centro, LVT, Alentejo and Algarve	General Health Administration (ARS) of the health institution.
Type of hospitalization	Type	nonICU and ICU	Non-ICU patients did not receive intensive care unit (ICU) treatment, ICU patients did.
Age	Age	Whole numbers from 0 to 105	Age of the patient, at the time of entry.
Age Group	Group	[0,5), [5,10), (...), [75,80), [80,150)	Designated age group of the patient, based on age at entry.
Gender	Gender	Feminine and Masculine	Gender of the patient.
Entry date	Entry.date	date in format "%d/%m/%Y"	Date of patient's entry into the institution.
Discharge date	Discharge.date	date in format "%d/%m/%Y"	Patient discharge date from the health institution.
Outcome	Outcome	Deceased and Discharged	If the patient died in hospital care or was discharged.
Length of stay	Time.stay	Whole numbers	Number of days between the discharge and entry date.
Month of entry	Entry.month	date in format "%m/%Y"	Month of entry in hospital care.
Month of discharge	Discharge.month	date in format "%m/%Y"	Month of discharge from hospital care.

→

Designations	Description
date	Row day, from 1 March 2020 to 31 March 2021.
entries	Entries by day.
discharges	Discharges by day.
entries_nonICU	Entries of non-ICU patients, by day.
discharges_nonICU	Discharges of non-ICU patients, by day.
in_nonICU	Non_ICU patients in hospital care, on the respective day.
entries_ICU	Entries of ICU patients, by day.
discharges_ICU	Discharges of ICU patients, by day.
in_ICU	CU patients in hospital care, on the respective day.
deaths_nonICU	Deceased non-ICU patients, by day.
deaths_ICU	Deceased ICU patients, by day.
sum_deaths_nonICU	Cumulative sum of deceased non-ICU patients.
sum_deaths_ICU	Cumulative sum of deceased ICU patients.

Variable	Designations	Description
Group	Group	One of the following: Total, Region, Gender, Group and Entry.month.
Level	Level	Row level, according to group.
Entries	Entries_(Type)	Number of entries.
Percentage to ICU	$Prop_ICU = \frac{Entries_ICU}{Entries_Total}$	ICU patients percentage in Total population.
Fatalities	Fatalities_(Type)	Number of fatalities.
Fatality rate	$Fatal_rate_(Type) = \frac{Fatalities_(Type)}{Entries_(Type)}$	Proportion of entries with the *Deceased* outcome.
Proporcion of all entries	$Prop_Entries_(Type) = \frac{Entries_(Type)}{sum(Entries_(Type))}$	Proportion of the total *Type* entries.
Population	Pop	Population numbers.
Percentage of the population	$Prop_Entries_Pop_(Type) = \frac{Entries_(Type)}{Pop}$	Proportion of the population of each *Level*.

Exploração

Objetivo: Relacionar as diferentes variáveis de forma a encontrar relações de interesse.

Resultados

Objetivo: Identificar que variáveis tem interesse para a questão em estudo.