

# **Curso de Vigilância em Saúde apoiado por plataforma BI**

**Aula 09- Modelagem epidêmica e cálculo do número efetivo de reprodução da infecção ( $R_t$ )**

## 1. Apresentação

Olá!

Seja bem-vindo à quarta aula do curso de Vigilância em Saúde apoiado por plataforma *Business Intelligence* (BI)! Nesse curso estamos aprendendo alguns conceitos de Vigilância Epidemiológica apoiada por ferramentas de BI para análise de dados temporais.

O objetivo do nosso curso é ajudar a análise epidemiológica em um contexto de vigilância em serviços de saúde, utilizando ferramentas que facilitem a vida dos profissionais de saúde.

Essa é nossa última aula! Parabenizamos você por ter chegado até aqui! Vamos falar sobre número efetivo de reprodução da infecção e como podemos calculá-lo no R.

## **2. Introdução**

Nessa última aula, vamos trabalhar com taxas de transmissão. Esse é um conceito muito importante em Epidemiologia e é fundamental para a Vigilância em Saúde, pois nos permite propor ações de contenção de epidemias e surtos (1, 2).

A Epidemiologia se preocupa em responder questões fundamentais do aspecto da saúde de desfechos que podem ser positivos ou negativos em saúde. Os fenômenos de saúde podem ser descritos por meio da descrição de local, tempo e pessoa (2).

Quando tratamos do tempo queremos entender a ocorrência das doenças e sua distribuição ao longo do tempo para propor ações que podem ser desenvolvidas pelos serviços de saúde (2, 3).

Compreender as taxas de transmissão de uma doença também é fundamental para prever os insumos necessários para o tratamento da doença e para prevenir sua transmissão.

## **3. Número efetivo de reprodução da infecção ( $R_t$ )**

A taxa de transmissão ou número efetivo de reprodução da infecção ( $R_t$ ) é a média de pessoas infectadas por cada indivíduo doente em um determinado período. Esse indicador nos permite estimar o número de pessoas infectadas a partir do primeiro caso.

Nas aulas anteriores, mostramos como estimar novos casos de uma determinada doença com base em um modelo de séries temporais. Esses modelos são muito úteis quando temos dados históricos de determinadas doenças e podem se aplicar para identificar surtos e epidemias de doenças já conhecidas e ocorrentes no território.

O  $R_t$  está estimando o número de novos casos com base na transmissão de um indivíduo infectado para outros em um período determinado e pode ser usado em casos de novas doenças ou em epidemias que já estão ocorrendo para estimar o alcance e o tempo em que mais pessoas podem ser infectadas.

#### 4. Modelagem Epidêmica no R

Você deve ter aberto o R pelo arquivo Rproject que está disponível em nossa plataforma. Caso não se lembre de como realizar essa operação, retorne à aula 4 com as instruções detalhadas.

Vamos iniciar carregando pacotes. Primeiro, vamos instalar o pacote *Pacman*, que permite a instalação e carregamento de quaisquer pacotes que estejam no CRAN.

```
#instalando e carregando pacotes
install.packages("pacman")

#carregando pacotes necessários
pacman::p_load(
  rio,          # importando dados
  here,         # locando dados
  tidyverse,    # manipulando dados e criando gráficos
  epicontacts,  # analisando redes de transmissão
  EpiEstim,     # estimando Rt
  projections,  # projetando a incidência
  incidence2,   # manipulando dados de incidência
  epitrix,      # usando funções de epidemiologia
  distcrete     # distribuições discretas de tempo
)
```

Nós vamos importar dados internacionais de Ebola. Esses dados foram importados do *The Epidemiologist R Handbook* (disponível em <https://epirhandbook.com>).

```
#importando dados
linelist <- import("dados/linelist_cleaned.rds")
#disponível em https://epirhandbook.com/en/epidemic-modeling.html
```

Após a importação dos dados, que estão disponíveis no projeto em nossa plataforma, nós vamos calcular os casos por dia. Isso é importante para numerarmos os casos por dia e podermos calcular a infecção ao longo do tempo.

```
##calculando incidência desde a primeira data
cases <- incidence2::incidence(linelist, date_index =
date_onset) %>% # gnúmero de casos por dia
  tidyr::complete(date_index = seq.Date(
# incluindo todas as datas no período considerado
```

```

    from = min(date_index, na.rm = T),
    to = max(date_index, na.rm=T),
    by = "day"),
    fill = list(count = 0)) %>%
# convertendo NA em 0
    rename(I = count, #
renomeando as colunas para a coincidir com a função
estimate_R
    dates = date_index)

```

Nós vamos configurar alguns parâmetros para a nossa modelagem de acordo com o que encontramos na literatura para a média e desvio padrão de transmissão do Ebola.

```

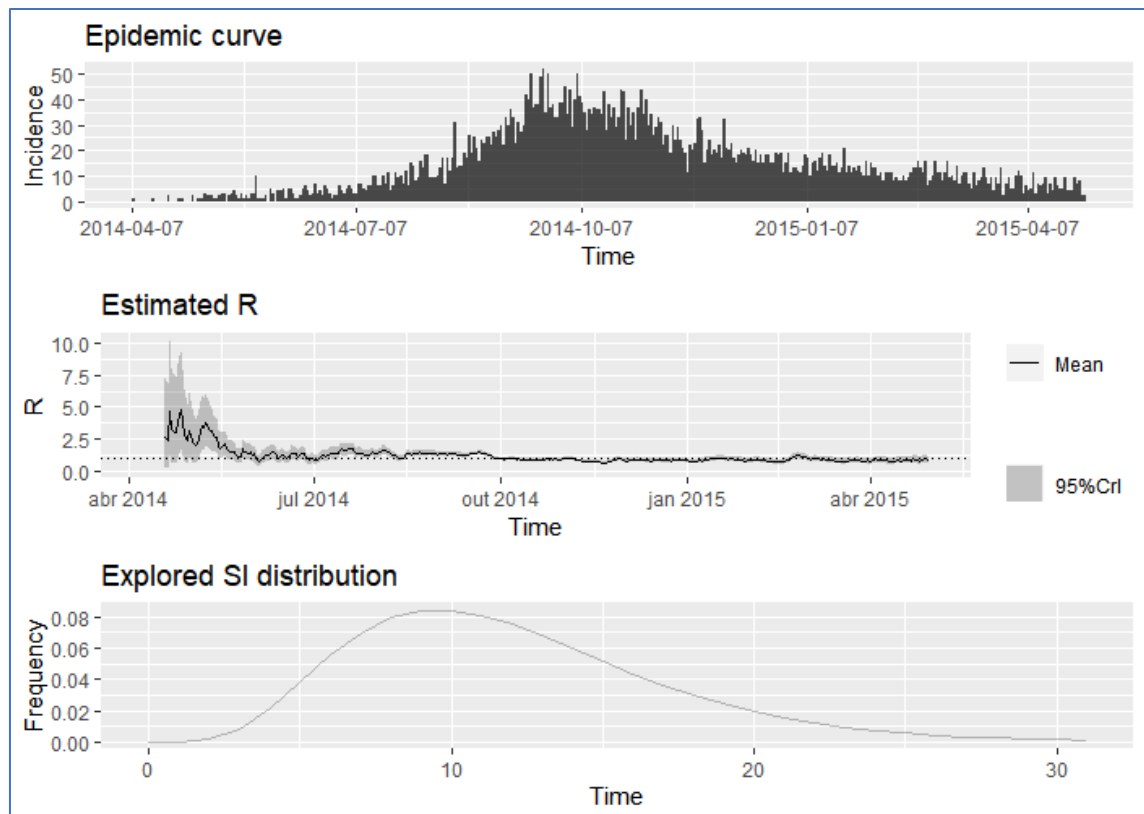
##gerando a configuração de tempo com os dados que estão na
literatura para a
##média e o desvio padrão do Ebola -> disponível em
https://www.nejm.org/doi/full/10.1056/nejmoa1411100
config_lit <- make_config(
  mean_si = 12.0,
  std_si = 5.2
)

epiestim_res_lit <- estimate_R(
  incid = cases,
  method = "parametric_si",
  config = config_lit
)

##plotando a curva epidêmica, o Rt estimado para o período
##e o intervalo serial (intervalo entre os sintomas dos casos
primários para os secundários)
plot(epiestim_res_lit)

```

Nós plotamos a curva epidêmica, o R estimado e a distribuição no tempo do intervalo serial, que se refere entre o intervalo entre os casos primários e os casos secundários.



Nós também podemos estimar o intervalo serial com base nos contatos dos primeiros casos. Para isso, precisaremos de uma base em que tenhamos os contatos e os pacientes iniciais.

Na base que disponibilizamos temos as informações sobre contatos que foram infectados, por isso será possível estimar o intervalo serial. Primeiro vamos criar um objeto para os contatos e usar a função *fit\_disc\_gamma* para estimar o intervalo e colocá-lo em um novo modelo.

```
##gerando os pares de contatos
contacts <- linelist %>%
  transmute(
    from = infector,
    to = case_id
  ) %>%
  drop_na()

##criando um objeto com os pares de contatos
epic <- make_epicontacts(
  linelist = linelist,
  contacts = contacts,
  directed = TRUE
)
```

```
##estimando o intervalo serial gamma
serial_interval <- fit_disc_gamma(get_pairwise(epic,
"date_onset"))
```

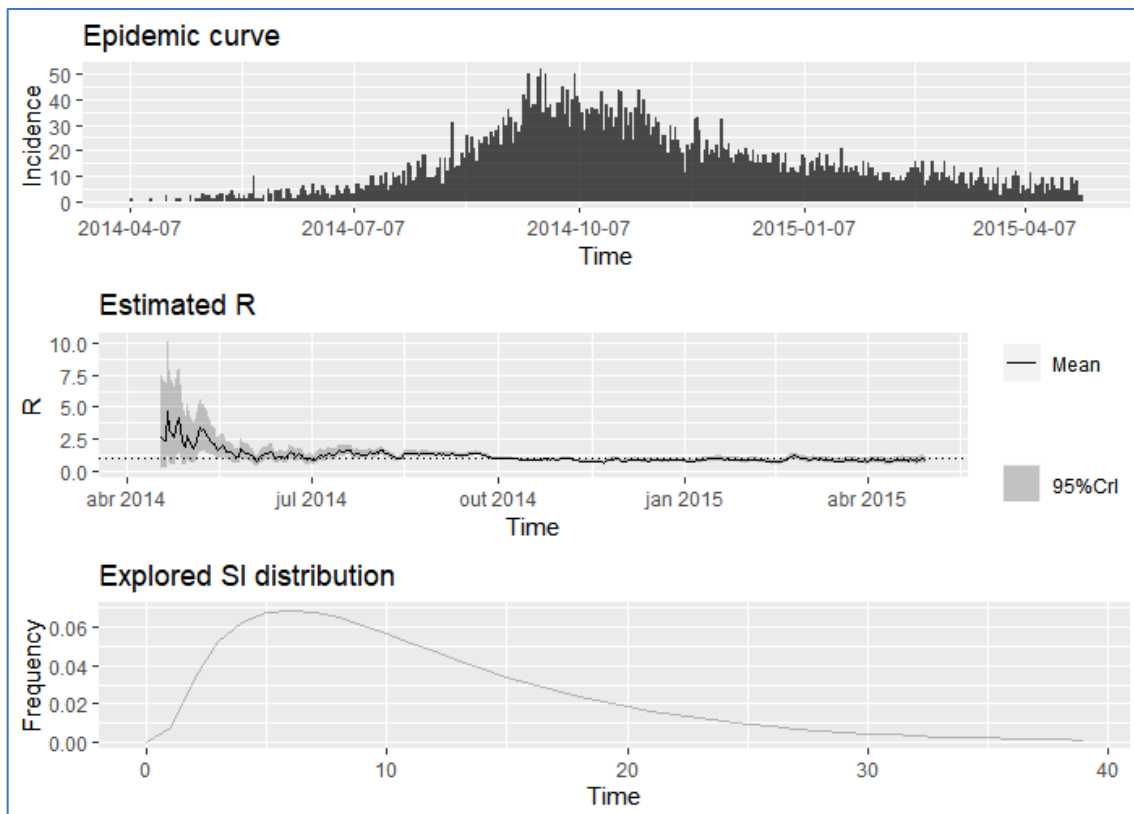
Após estimar o intervalo serial, vamos criar um modelo epidêmico com os dados estimados da nossa base de dados.

```
##usando o intervalo serial gamma para criar um modelo
config_emp <- make_config(
  mean_si = serial_interval$mu,
  std_si = serial_interval$sd
)

##rodando o novo modelo de estimativa do Rt
epiestim_res_emp <- estimate_R(
  incid = cases,
  method = "parametric_si",
  config = config_emp
)

##plotando o novo modelo de estimativa Rt
plot(epiestim_res_emp)
```

Na plotagem percebemos que houve uma mudança entre os valores estimados com os parâmetros da literatura e com os dados que estamos usando. O intervalo serial, que no modelo anterior caminhava entre o primeiro e o 30º dia, com a concentração no 10º dia passou a caminhar até o 40º dia e se concentrar no 5º dia.



Você também poderá alterar os dados, incluindo um intervalo conhecido para rodar um novo modelo.

```
#Especificando intervalos de estimação de tempo

##definindo um objeto com as datas iniciando a partir de 1
de junho
start_dates <- seq.Date(
  as.Date("2014-06-01"),
  max(cases$dates) - 7,
  by = 1
) %>%
  `--`(min(cases$dates)) %>%
  as.integer()

##adicionando um intervalo de 6 dias a partir da data de
início
end_dates <- start_dates + 6

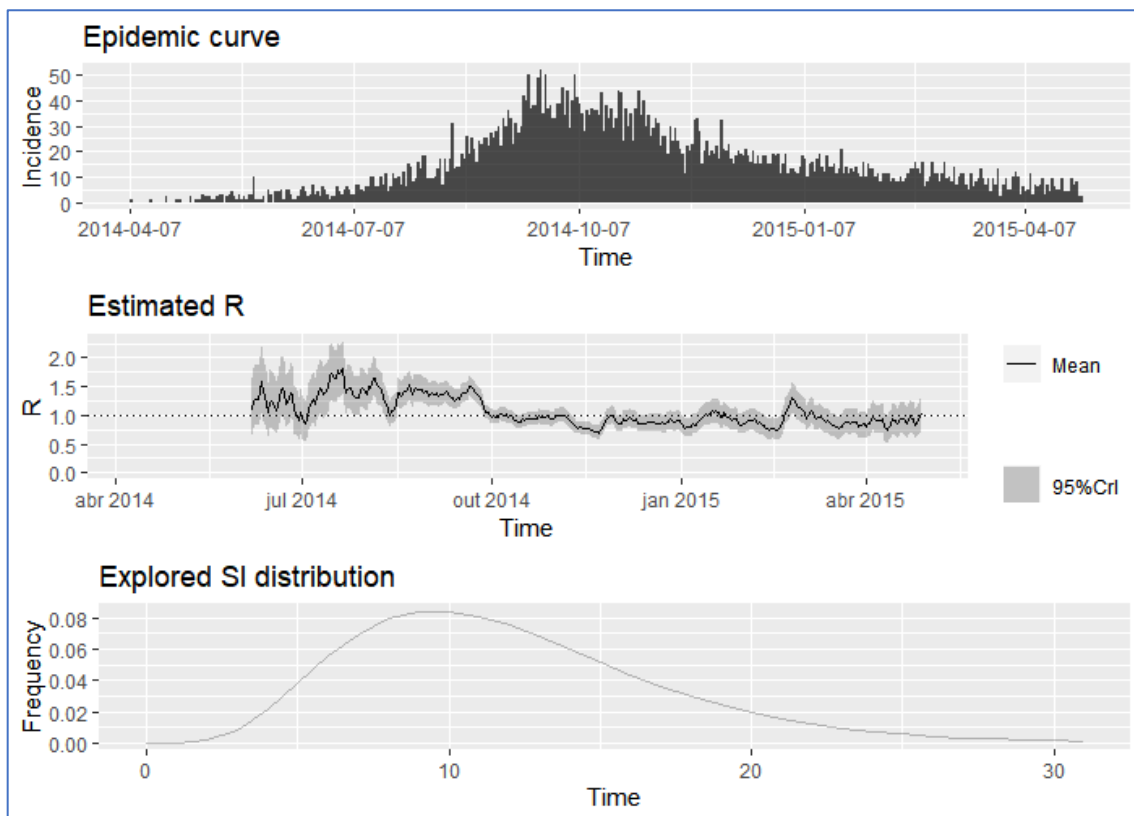
##configurando a estimativa com base nas datas de início e
fim do intervalo proposto
config_partial <- make_config(
  mean_si = 12.0,
  std_si = 5.2,
  t_start = start_dates,
  t_end = end_dates
)
```



```
##rodando o novo modelo
epiestim_res_partial <- estimate_R(
  incid = cases,
  method = "parametric_si",
  config = config_partial
)

##plotando os outputs do intervalo
plot(epiestim_res_partial)
```

A distribuição do intervalo serial permaneceu igual à primeira estimativa, mas o  $R_t$  mudou com essa análise, se tornando mais distribuído ao longo do tempo.



Agora vamos criar *data.frames* para armazenar os dados e vamos plotar o potencial de transmissão no tempo e o  $R_t$ .

```
#analizando os outputs

##criando um data.frame no formato wide para a mediana
df_wide <- epiestim_res_lit$R %>%
  rename_all(clean_labels) %>%
  rename(
    lower_95_r = quantile_0_025_r,
```

```

    lower_90_r = quantile_0_05_r,
    lower_50_r = quantile_0_25_r,
    upper_50_r = quantile_0_75_r,
    upper_90_r = quantile_0_95_r,
    upper_95_r = quantile_0_975_r,
  ) %>%
  mutate(
    dates = epiestim_res_emp$dates[round(map2_dbl(t_start,
t_end, median))],
    var = "R[t]"
  ) %>%
  left_join(cases, "dates") %>%
  mutate(
    across(
      lower_95_r:upper_95_r,
      ~ .x*I,
      .names = "{str_replace(.col, '_r', '_risk')}"
    )
  ) %>%
  pivot_longer(
    contains("median"),
    names_to = c(".value", "variable"),
    names_pattern = "(.+)_(.+)"
  ) %>%
  mutate(variable = factor(variable, c("risk", "r")))

##criando um data.frame no formato long para os quantis
df_long <- df_wide %>%
  select(-variable, -median) %>%
  pivot_longer(
    contains(c("lower", "upper")),
    names_to = c(".value", "quantile", "variable"),
    names_pattern = "(.+)_(.+)_(.+)"
  ) %>%
  mutate(variable = factor(variable, c("risk", "r")))

##plotando o gráfico
ggplot() +
  geom_ribbon(
    data = df_long,
    aes(x = dates, ymin = lower, ymax = upper, alpha =
quantile),
    color = NA
  ) +
  geom_line(
    data = df_wide,
    aes(x = dates, y = median),
    alpha = 0.2
  ) +
  facet_wrap(
    ~ variable,

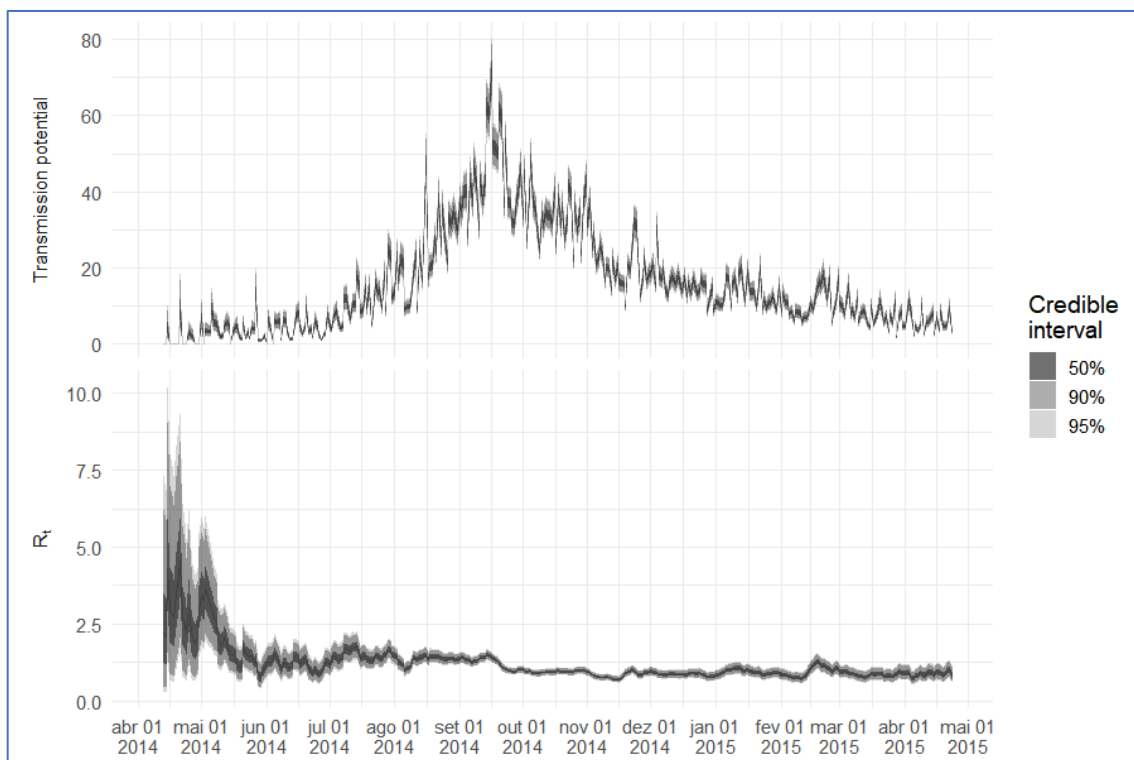
```

```

ncol = 1,
scales = "free_y",
labeller = as_labeller(c(r = "R[t]", risk =
"Transmission~potential"), label_parsed),
strip.position = 'left'
) +
scale_alpha_manual(
values = c(`50` = 0.7, `90` = 0.4, `95` = 0.2),
labels = function(x) paste0(x, "%")
) +
labs(
x = NULL,
y = NULL,
alpha = "Credible\\ninterval"
) +
scale_x_date(
date_breaks = "1 month",
date_labels = "%b %d\\n%Y"
) +
theme_minimal(base_size = 14) +
theme(
strip.background = element_blank(),
strip.placement = 'outside'
)

```

Plotando os gráficos de  $R_t$  e potencial de transmissão, temos uma ideia de como será a transmissão nos dias considerados.



Com esses dados também podemos projetar a incidência esperada para os próximos dias.

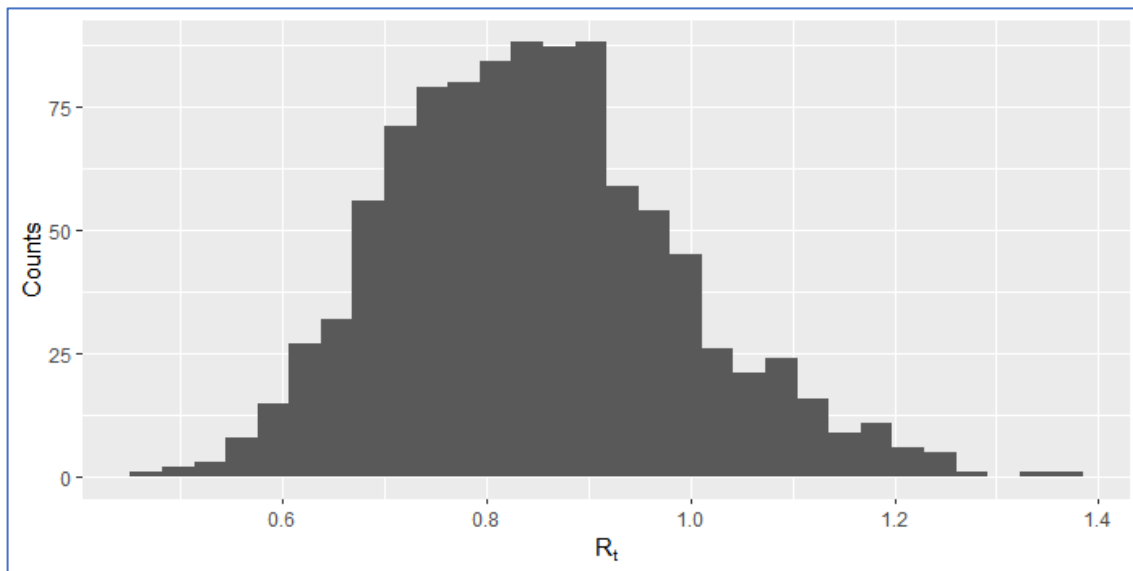
```
#projetando a incidência

##criando um objeto para as datas de incidência de casos
inc <- incidence::incidence(linelist$date_onset)

##extraído valores de r possíveis para as estimativas mais recentes
mean_r <- tail(epiestim_res_emp$R$`Mean(R)` , 1)
sd_r <- tail(epiestim_res_emp$R$`Std(R)` , 1)
shapyscale <- gamma_mucv2shapyscale(mu = mean_r, cv = sd_r/mean_r)
plausible_r <- rgamma(1000, shape = shapyscale$shape, scale = shapyscale$scale)

##verificando a distribuição
qplot(x = plausible_r, geom = "histogram", xlab = expression(R[t]), ylab = "Counts")
```

Para isso, vamos criar um objeto com a incidência de casos e verificar a distribuição dos dados para essa projeção com um histograma de distribuição dos dados.



Após a análise, vamos criar a projeção, com um gráfico para os próximos dias e com um data.frame com as incidências esperadas por data.

```
##criando a projeção
proj <- project(
  x = inc,
```

```

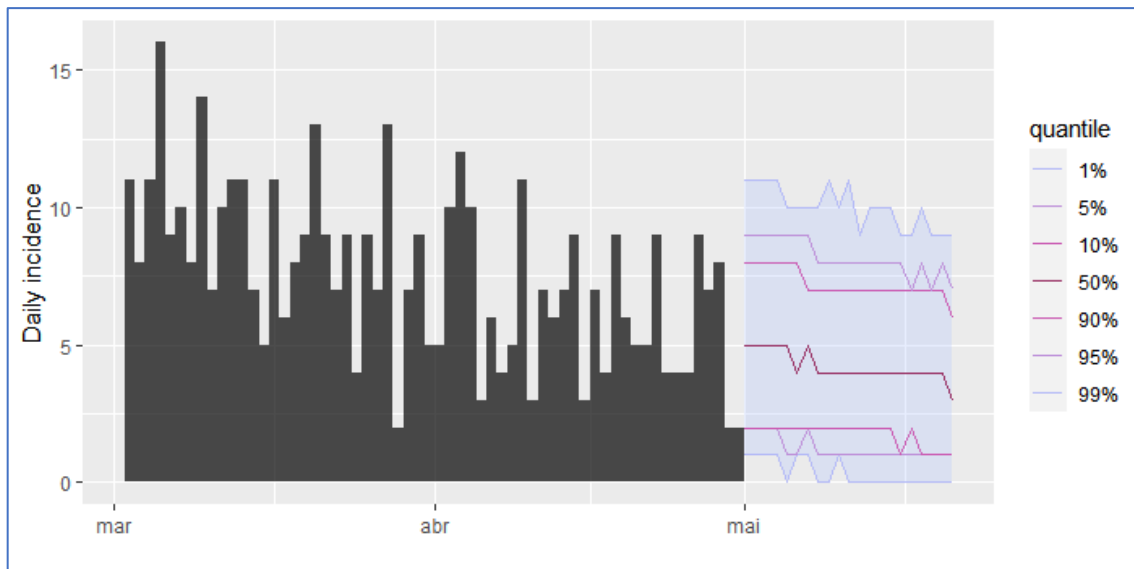
R = plausible_r,
si = serial_interval$distribution,
n_days = 21,
n_sim = 1000
)

##plotando a projeção
plot(inc[inc$dates > as.Date("2015-03-01")]) %>%
  add_projections(proj)

##convertendo os dados em um data.frame
proj_df <- as.data.frame(proj)
View(proj_df)

```

O gráfico final mostra os a incidência esperada por dia considerando os percentis esperados. Os valores estarão no *data.frame* “proj\_df”.



### Referências Bibliográficas

1. WALDMAN, Eliseu Alves et al. **Vigilância em saúde pública**. NAMH/FSP-USP; Peirópolis, 1998.
2. Medronho, R. D. A., Bloch, K. V., Luiz, R. R., & Werneck, G. L. (2009). **Epidemiologia**. Editora Atheneu.
3. ROTHMAN, Kenneth; GREENLAND, Sander; LASH, Timothy. **Epidemiologia Moderna-3ª Edição**. Artmed Editora, 2016.
4. CARRAPATO, Pedro; CORREIA, Pedro; GARCIA, Bruno. **Determinante da saúde no Brasil: a procura da equidade na saúde**. Saúde e Sociedade, v. 26, p. 676-689, 2017.
5. Brasil. Ministério da Saúde. Secretaria de Vigilância à Saúde. Secretaria de Atenção à Saúde. **Diretrizes Nacionais da Vigilância em Saúde/** Ministério da Saúde, Secretaria de Vigilância em Saúde, Secretaria de Atenção à Saúde. – Brasília: Ministério da Saúde, 2010.

## Atividades

1	Taxa de transmissão não é um conceito do campo da Epidemiologia.	V	
		F	X
2	Taxa de transmissão é um conceito muito importante em Epidemiologia e é fundamental para a Vigilância em Saúde	V	X
		F	
3	Analisar a taxa de transmissão permite propor ações de contenção de epidemias e surtos	V	X
		F	
4	A Epidemiologia preocupa-se em responder apenas questões sobre o tempo de ocorrência das doenças	V	
		F	X
5	A Epidemiologia se preocupa em responder questões fundamentais do aspecto da saúde de desfechos que podem ser positivos ou negativos em saúde.	V	X
		F	
6	Os fenômenos de saúde podem ser descritos por meio da descrição apenas do tempo e pessoa	V	
		F	X

7	Os fenômenos de saúde podem ser descritos por meio da descrição de apenas local e pessoa	V	
		F	X
8	Os fenômenos de saúde podem ser descritos por meio da descrição de local e tempo	V	
		F	X
9	Os fenômenos de saúde podem ser descritos por meio da descrição de local, tempo e pessoa.	V	X
		F	
10	Quando tratamos do tempo queremos entender a ocorrência das doenças e sua distribuição ao longo do tempo para propor ações que podem ser desenvolvidas pelos serviços de saúde.	V	X
		F	
11	A análise temporal não interfere nas ações de serviços de saúde que devem funcionar de forma contínua e promover a aquisição periódica de insumos.	V	
		F	X
12		V	X



	Compreender as taxas de transmissão de uma doença também é fundamental para prever os insumos necessários para o tratamento da doença e para prevenir sua transmissão.	F	
13	A taxa de transmissão ou número efetivo de reprodução da infecção ( $R_t$ ) é a média de pessoas infectadas por cada indivíduo doente em um determinado período.	V	X
		F	
14	O $R_t$ é um indicador nos permite estimar o número de pessoas infectadas a partir do primeiro caso.	V	X
		F	
15	O $R_t$ estima o número de novos casos com base na transmissão de um indivíduo infectado para outros em um	V	X
		F	

	<p>período determinado e pode ser usado em casos de novas doenças ou em epidemias que já estão ocorrendo para estimar o alcance e o tempo em que mais pessoas podem ser infectadas.</p>		
--	---	--	--