

task specified in SP2

This practical focus on developing a simulation-based method to infer Covid-19 fatal incidence rates from death data in English hospitals. The goal is to estimate when new fatal infections occurred, which is crucial for assessing the effectiveness of control measures during the pandemic.

Objective

Write R code to implement a simulation method that infers fatal incidence rates from Covid-19 death data. The method will use a simple iterative approach to match simulated death times with actual recorded deaths.

Data and Parameters

- Data source: File 'engcov.txt' (first 150 rows)
- Columns: julian (day of the year), nhs (number of Covid deaths in English hospitals)
- Infection-to-death distribution: Log-normal (meanlog=3.152, sdlog=0.451)

Algorithm Overview

1. Calculate probabilities for disease durations (1-80 days) using log-normal distribution
2. Initialize infection times for each fatality
3. Iterate the following steps:
 - Generate simulated death days
 - Compute goodness of fit (P statistic)
 - Propose changes to infection times and accept if fit improves

Function Specification

Implement a function named `deconv` with the following signature:

```
deconv(t, deaths, n.rep = 100, bs = FALSE, t0 = NULL)
```

The function should return a list containing:

- P: Vector of P statistic history
- inft: Matrix of estimated infections per day
- t0: Final state of infection times

Additional Tasks

- Implement bootstrapping for uncertainty quantification
- Create a final plot showing estimated incidence trajectory, uncertainty, death data, and lockdown date

Requirements

- Use only base R (no additional packages)
- Include well-structured comments and documentation
- Ensure efficiency (runtime under 3 minutes)
- Collaborate using git and GitHub in a group of 3
- Submit one R code file per group by 12:00, 18th October 2024

Remember to include a statement of team member contributions at the beginning of your code.

本次实践重点是开发一种基于模拟的方法，从英格兰医院的死亡数据中推断新冠肺炎的致命发病率。目标是估计新的致命感染发生的时间，这对于评估大流行期间控制措施的有效性至关重要。

目标

编写R代码来实现一种模拟方法，从新冠肺炎死亡数据中推断致命发病率。该方法将使用简单的迭代方法来匹配模拟的死亡时间与实际记录的死亡。

数据和参数

- 数据来源：文件'engcov.txt'（前150行）
- 列：julian（一年中的天数），nhs（英格兰医院的新冠死亡数）
- 感染至死亡分布：对数正态分布（meanlog=3.152, sdlog=0.451）

算法概述

1. 使用对数正态分布计算疾病持续时间（1-80天）的概率
2. 初始化每个死亡病例的感染时间
3. 迭代以下步骤：
 - 生成模拟的死亡日期
 - 计算拟合优度（P统计量）
 - 提出感染时间的变更，如果拟合改善则接受

函数规范

实现一个名为 `deconv` 的函数，其签名如下：

```
deconv(t, deaths, n.rep = 100, bs = FALSE, t0 = NULL)
```

该函数应返回一个包含以下内容的列表：

- P：P统计量历史的向量
- inft：每日估计感染数的矩阵
- t0：感染时间的最终状态

附加任务

- 实现引导法进行不确定性量化
- 创建最终图表，显示估计的发病轨迹、不确定性、死亡数据和封锁日期

要求

- 仅使用基础R（不使用额外的包）

- 包含结构良好的注释和文档
- 确保效率（运行时间在3分钟以内）
- 3人一组使用git和GitHub进行协作
- 每组在2024年10月18日12:00前提交一个R代码文件

请记得在代码开头包含团队成员贡献声明。

在本实践中，您将编写代码，实现一个非常简单的基于模拟的方法，用于推断英国医院Covid死亡的致命发病率。从概念上讲，这个方法是这样工作的。我们首先给每个受害者分配一个猜测的感染时间。然后，我们从感染到死亡的分布中随机抽取每个感染时间，以得到隐含的死亡时间。最初，由此得到的死亡时间分布与真实分布很不匹配。因此，对于每一次感染，我们依次随机提议将其移动几天，只有在改进了模拟与真实死亡时间分布的匹配后，我们才接受提议的移动。这个过程只是简单地迭代，直到我们很好地匹配死亡时间分布。

您将在文件engcov.txt中处理Covid死亡数据，该文件将使用read.table进行读取。专栏julian给出了一年的日子(2020年1月1日为第1天)，专栏nhs给出了当天英国医院Covid死亡的数量。为此，只使用前150行数据。您将使用对数正态分布(见?dlnorm)作为感染至死亡时间的分布。根据ISARIC研究，适当的参数是meanlog=3.152和sdlog=0.451(其他研究给出了非常类似的结果)。让 d_i 表示一年中每一天的实际死亡人数， d_i^s 表示模拟的死亡人数。模拟拟合的优度可以通过修改后的Pearson统计量来判断

$$P = \sum_i (d_i - d_i^s)^2 / \max(1, d_i^s)$$