Berechnung von R Bru

Ausgangspunkt ist eine simple Exponentialfunktion:

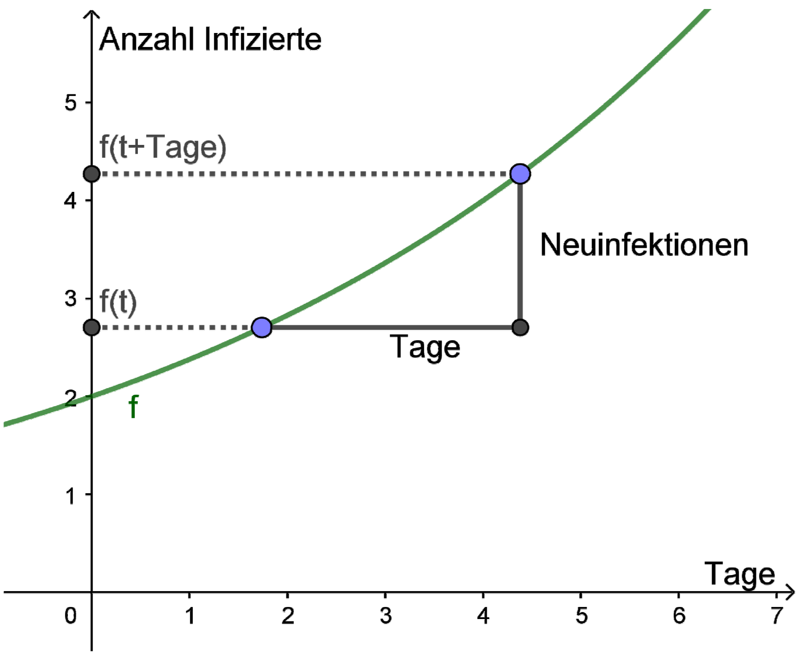
Sei B die Anzahl der Infizierten zu einem Startzeitpunkt 0. Unsere Funktion bekommen wir mit der Zusatzinformation, dass alle 4 Tage eine Generation zu Ende ist (eine Schätzung des RKI), d.h. nach 4 Tagen hat ein Infizierter seine Krankheit an R Personen weitergegeben (d.h. R = 2, wenn er es zwei weitergibt, R = 0,5, wenn zwei Personen es an eine weitere weitergeben.

Zeit t in Tagen	Anzahl Infizierter f(t)
0	B
4	$R \cdot B$
8	$R \cdot R \cdot B = R^2 \cdot B = R^{\frac{8}{4}} \cdot B$
12	$R \cdot R \cdot R \cdot B = R^3 \cdot B = R^{\frac{12}{4}} \cdot B$
t	$R^{\frac{t}{4}} \cdot B$

D.h. die Exponentialfunktion mit R lautet: $f(t) = R^{\frac{t}{4}} \cdot B$

Nun wollen wir jedoch R bestimmen, wissen jedoch B nicht, sondern lediglich die Anzahl der Neuinfizierten N pro Tag.

Die Neuinfizierten pro Tag entspricht einer Sekantensteigung bzw. ist die Einheit der Ableitung von f.
z.B. für R = 2 und B = 2:



Die Ableitung (wir werden es in Kürze lernen) von f ist:

f'(t) = $R^{\frac{t}{4}} \cdot B \cdot \frac{1}{4} \ln(R)$ ln(R) ist der natürliche Logarithmus mit Basis e

für uns erstmal einfach nur eine Zahl (TR!)

Wenn wir nun die Neuinfizierten für einen Tag kennen, haben wir annähernd f'(t) bzw. die Sekantensteigung zwischen zwei Tagen. Wenn ich nun einen Tag durch seinen vorhergehenden bzw. den folgenden f'(t+1) durch den aktuellen Tag f'(t) teile, erhalte ich eine wunderschöne Gleichung:

$$\frac{f'(t+1)}{f'(t)} = \frac{R^{\frac{t+1}{4}} B \frac{1}{4} \ln(R)}{R^{\frac{t}{4}} B \frac{1}{4} \ln(R)} = \frac{R^{\frac{t+1}{4}}}{R^{\frac{t}{4}}} = \frac{R^{\frac{t}{4}} R^{\frac{1}{4}}}{R^{\frac{t}{4}}} = R^{\frac{1}{4}}$$

Das ist schön, denn wenn wir nun die Neuinfektionen zweier aufeinander folgender Tage weiß, kann man R berechnen.
Hier die Grafik vom RKI. Probieren wir es aus.

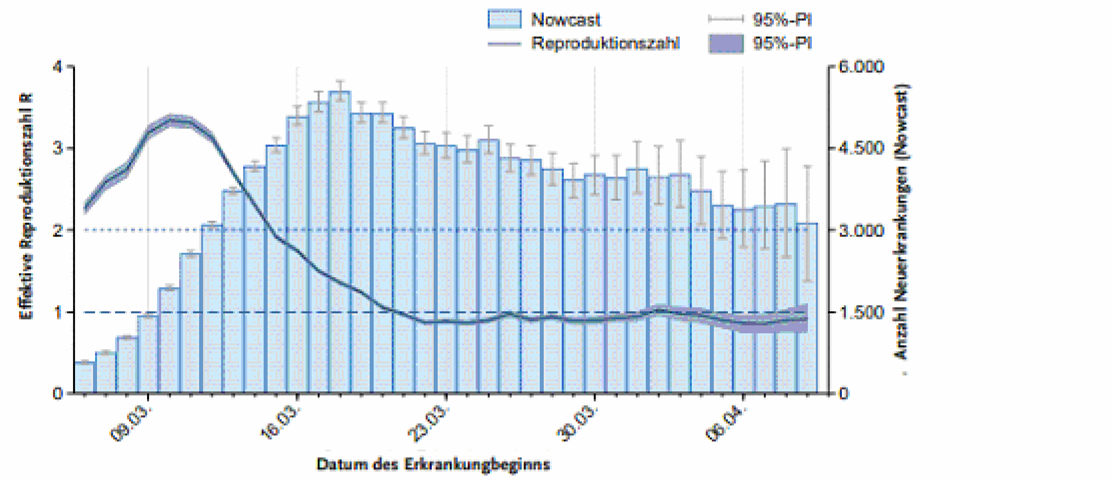


Abb. 4 | Schätzung der effektiven Reproduktionszahl R für eine angenommene Generationszeit von 4 Tagen und die durch das Nowcasting geschätzten Anzahlen von Neuerkrankungen, auf denen die R-Schätzung beruht.

Am 9.3. waren es etwa 1350 Neuerkrankte und am Folgetag etwa 1950:

$$R^{\frac{1}{4}} = \frac{1950}{1350} \Rightarrow R = \left(\frac{1950}{1430}\right)^4 \approx 3{,}45$$

Das ist zwar etwas hoch, kommt aber in etwa hin. Das RKI rechnet so jedoch nicht.

Das RKI weiß natürlich, dass die einzelnen Werte statistisch recht ungenau sind (außerdem wissen die Mediziner nicht genau was hoch ein Viertel überhaupt sein soll ;-)), deswegen betrachten sie einen Zeitraum von acht Tagen. Sie summieren die Werte vom 6.3.-9.3. und die Werte vom 10.3.-13.3. und betrachten diese als zwei Funktionswerte, die im Schnitt vier Tage auseinander liegen. Das vereinfacht schön die Rechnung:

$$\frac{f'(t+4)}{f'(t)} = \frac{R^{\frac{t+4}{4}} B \frac{1}{4} \ln(R)}{R^{\frac{t}{4}} B \frac{1}{4} \ln(R)} = \frac{R^{\frac{t+4}{4}}}{R^{\frac{t}{4}}} = \frac{R^{\frac{t}{4}} R^{\frac{4}{4}}}{R^{\frac{t}{4}}} = R^{\frac{4}{4}} = R$$

Wenn man versucht die Zahlen nachzurechnen (Schwer abzulesen, wer eine Tabelle mit den Zahlenwerten findet – bitte her damit):
Kommt man für den 6.3.-9.3. auf etwa: 3780 pro 4 Tage (Änderungsrate!)
Und für 10.3. – 13.3. auf etwa 11200 pro 4 Tage

Das ergibt ungefähr ein R = 11200 : 3780 = 2,96

Den Wert trägt das RKI nun für den 13.3. als Wert ein („Der so ermittelte R-Wert wird dem letzten dieser 8 Tage zugeordnet, weil erst dann die gesamte Information vorhanden ist"). Da mir die genauen Zahlenwerte fehlen, stimmt das natürlich nur sehr ungefähr. (Ich habe die starke Vermutung, dass das trotzdem nicht so ganz stimmt und die Werte irgendwie doch früher eingetragen werden)
Nun zu ihren Aufgaben:

1. Versuchen Sie das gleiche für den Zeitraum 7.3.-10.3. und 11.3.-14.3. und prüfen sie ihren Wert.
2. Warum ist es ein grober Fehler, wenn das RKI die Anzahl der Neuinfektionen für die Berechnung von R verwendet und nicht den Anteil der Neuinfektionen an der Stichprobe?
3. Sie wissen, dass von 1.3.-8.3. etwa 7000 Tests pro Tag, von 9.3.-15.3. etwa 31000 Tests pro Tag und von 16.3.-22.3. etwa 64700 Tests pro Tag gemacht wurden. Welche Werte für R würden Sie in diesen Zeiträumen grob geschätzt erwarten, wenn sich der Anteil der Infizierten unter der Bevölkerung NICHT ändert?

Bildquelle: https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2020/17/Art_02.html