

Ein paar einfache Wahrscheinlichkeiten zum Coronavirus

Jean Roch Donsimoni, Damir Stijepic und Klaus Wälde¹
Johannes Gutenberg Universität Mainz, CESifo and IZA

9. März 2020

Die rasante Ausbreitung der bestätigten Erkrankungen mit dem Coronavirus (SARS-CoV-2) führt zu vielfältigen Befürchtungen. Diese kurze Notiz fasst die Fakten zusammen und leitet daraus einfache Wahrscheinlichkeiten ab. Wie hoch ist das Ansteckungsrisiko für eine einzelne Person pro Tag? Wie hoch ist die Wahrscheinlichkeit, innerhalb der nächsten Woche bzw. des nächsten Monats gesund zu bleiben? Auf diese Wahrscheinlichkeiten können Entscheidungen bezüglich Schulschließungen oder Absage von Massenveranstaltungen gegründet werden.

Zusammenfassung

Auf Grundlage der gemeldeten Neuerkrankungen des Robert Koch Instituts liegt die Wahrscheinlichkeit, sich innerhalb der letzten 7 Tage infiziert zu haben bei etwa einem halben Tausendstel von einem Promille pro Tag. Das heißt, pro Tag steckte sich im Schnitt eine von 500.000 Personen an. Die Wahrscheinlichkeit, eine Woche lang gesund zu bleiben liegt mit dieser durchschnittlichen Ansteckungsrate bei 99,999%. Die Wahrscheinlichkeit, ein Monat gesund zu bleiben liegt bei 99,995%.

Diese Zahlen erscheinen auf der individuellen Ebene klein. Bei unveränderten Ansteckungsraten bedeutet dies jedoch eine tägliche Wachstumsrate von circa 34% für die Anzahl der Neuerkrankungen. Damit steigt die Anzahl der Infizierten in Deutschland innerhalb der kommenden 7 Tage auf knapp 9000, das wären in einem Monat mehr als 7 Millionen. Die Vorhersage für die nächste Woche erscheint aufgrund der relativen Stabilität der Entwicklung von gemeldeten Neuinfektionen wahrscheinlich. Die Vorhersage für in einem Monat ist unsicherer und hängt vor allem auch von den ergriffenen Gegenmaßnahmen ab. Die Anzahl der berichteten Infektion in der Provinz Hubei mit der Hauptstadt Wuhan beläuft sich seit Datenerhebung auf knapp 68.000. Bei einer Bevölkerung von gut 59 Millionen ist etwas mehr als 1 von 1000 erkrankt. Somit erscheint eine Anzahl von berichteten Erkrankungen in Deutschland (mit einer Bevölkerung von 83 Millionen) jenseits von 83.000 Fällen als unplausibel.

Es erscheint jedoch allgemein einsichtig, dass aktuelle individuelle Infektionsraten zu hoch sind und auf individueller wie gesellschaftlicher Ebene Anstrengungen unternommen werden sollten, um Infektionsraten zu senken.

¹Prof. Dr. Klaus Wälde, Johannes Gutenberg Universität Mainz, Gutenberg School of Management and Economics, Jakob-Welder-Weg 4, D-55131 Mainz, Telefon + 49.6131.39-20143, waelde@uni-mainz.de

1 Die Datenlage

Das Robert-Koch Institut (RKI) veröffentlicht² täglich die Anzahl der gemeldeten Erkrankungen mit dem Coronavirus (SARS-CoV-2). Der starke Anstieg der Erkrankungen ist vermutlich einer der Ursache für die öffentliche Sorge um den weiteren Verlauf.

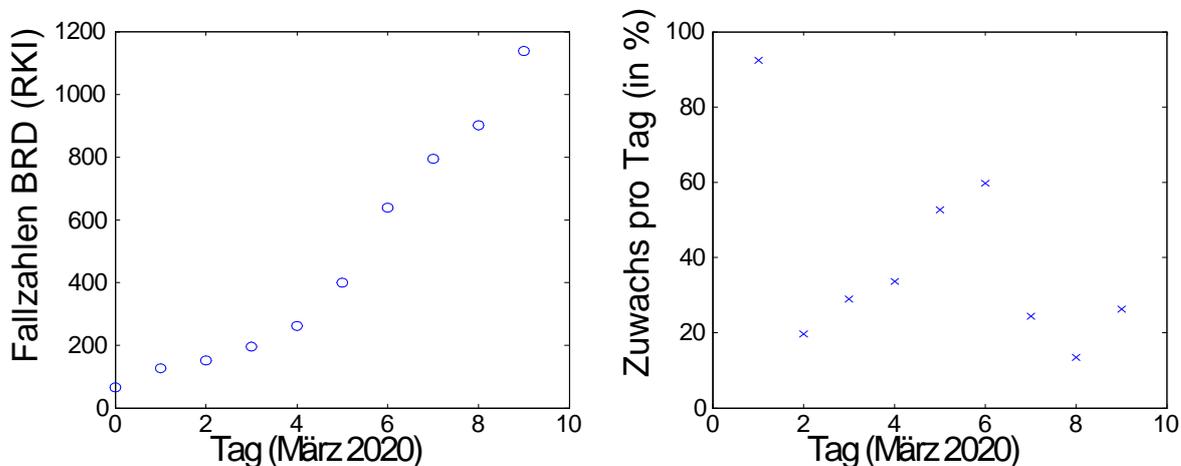


Figure 1 Die Anzahl der Erkrankungen und die tägliche Wachstumsrate

Die linke Abbildung zeigt augenscheinlich ein exponentielles Wachstum. Beeindruckend sind dabei die in der rechten Abbildung sichtbaren täglichen Wachstumsraten. Auch wenn diese u.a. auch erfassungs- und meldebedingt stark schwanken, liegen sie im Schnitt bei circa 39% pro Tag. Die Werte am 7. und 8. März fielen auf Samstag und Sonntag. Würde man diese Wachstumsrate von 39% pro Tag fortschreiben, wären in gut einem Monat die gesamte Bevölkerung der Bundesrepublik erkrankt.

2 Erkrankungswahrscheinlichkeiten

2.1 Die tägliche Wahrscheinlichkeit

- Einfache Betrachtung

Was bedeuten diese Zahlen für jede einzelne Person? Wenn man die täglichen Infektionen durch die Anzahl der gesunden Bevölkerung teilt, bekommt man die Infektionswahrscheinlichkeit: Wenn an einem Tag 1000 Menschen gesund sind, am nächsten Tag 1 krank, dann war die Infektionswahrscheinlichkeit $1/1000$. Nimmt man also die Neuerkrankungen pro Tag vom RKI und teilt durch die Bevölkerung in der Bundesrepublik abzüglich der Krankheitsfälle, bekommt man die Erkrankungswahrscheinlichkeit p_t am Tag t .³

²https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html

³Siehe Gleichung (A.1) im Anhang zu Details.

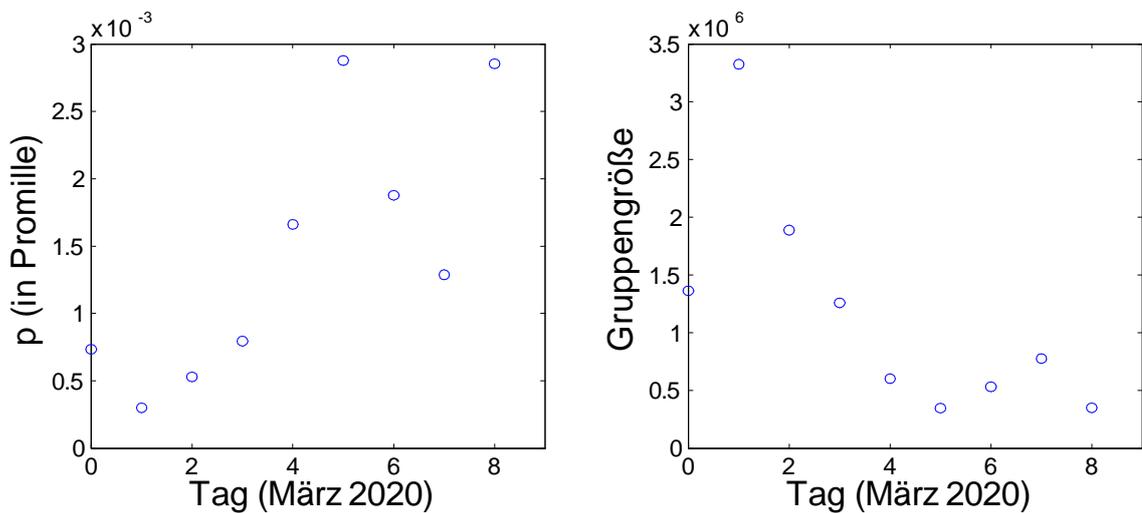


Figure 2 Individuelle Erkrankungs wahrscheinlichkeiten (links) und Gruppengröße für eine Neuerkrankung (rechts)

Wie kann man eine Infektionswahrscheinlichkeit verstehen? Nach der linken Abbildung lag sie am 2. März 2020 bei $0,5 \times 10^{-3} = 0,0005$ Promille, also $0,00005\%$.⁴ Wenn man also die Bevölkerung in Gruppen von $100/0,00005 = 2$ Millionen Personen aufteilt, dann hat im Schnitt jede Gruppe am 3. März einen Erkrankten. Die Wahrscheinlichkeit, am 2. März erkrankt zu sein war also 1 zu 2 Millionen (rechte Abbildung). Nimmt man dann alle 83 Millionen in der Bundesrepublik zusammen, dann ergibt sich eine Zunahme von $83/2 = 41,5$ Personen, was (abgesehen von Rundungsfehlern) dem tatsächlichen Anstieg vom 2. auf den 3. März entspricht.

Die Infektionswahrscheinlichkeit am 5. März lag dann bei knapp $3 \times 10^{-3} = 0,003$ Promille (linke Abbildung). Es reichte also schon eine Gruppengröße von 330.000 Individuen aus, um im Schnitt einen Neuerkrankten am 6. März zu bekommen. Das Risiko, am 5. März zu erkranken lag also bei 1 zu 330 Tausend (rechte Abbildung).

- Inkubationszeit

Die Daten erfassen die gemeldeten Infektionsfälle. Es ist damit zu rechnen, dass zwischen der Infektion, dem Auftreten von Symptomen und der Meldung beim RKI mindestens eine Woche verstreicht. Somit erfassen die gemeldeten Infektionsfälle an einem Tag t vermutlich eher die Infektionswahrscheinlichkeiten ein bis zwei Wochen vorher. Man könnte also aus obiger Abbildung 2 (links) mit dem Wert am 9. März die Infektionswahrscheinlichkeit (mit folgender Erkrankung) eine Woche vorher, also am 2. März (oder noch früher) beschreiben.

Dieser Zusammenhang lässt Raum für Optimisten und Pessimisten. Optimisten würden sagen, dass die Meldezahlen in 1 Woche sicher niedriger sein werden. Also ist die Wahrscheinlichkeit am 9. März zu erkranken niedriger als in der Abbildung 2 angegeben. Pessimisten würden sagen, die Meldezahlen werden weiter steigen, so dass die Erkrankungswahrscheinlichkeit am 9. März höher ist als die Abbildung zeigt.

2.2 Die Wahrscheinlichkeit, gesund zu bleiben

Wenn wir eine optimistischere Frage stellen, können wir berechnen, wie hoch die Wahrscheinlichkeit ist, bei unverändertem Verhalten soundsoviel weitere Tage gesund zu bleiben. Die

⁴Es ist selbstverständlich, dass es nicht eine Erkrankungswahrscheinlichkeit für alle Einwohner, die sich mindestens in Geschlecht, Alter, allgemeinen Gesundheitszustand, Exposition und Hygieneverhalten unterscheiden, gibt. Ziel hier ist es, Infektionswahrscheinlichkeiten für einen durchschnittlichen Einwohner zu veranschaulichen.

Wahrscheinlichkeit, morgen gesund zu sein ist gleich der Wahrscheinlichkeit, heute nicht krank zu werden, also $1 - p_t$. Nehmen wir an, die Erkrankungswahrscheinlichkeit würde sich ab heute nicht mehr ändern sondern konstant bleiben (also $p_\tau = p_t$ für alle $\tau \geq t$). Dann ist die Wahrscheinlichkeit, nach n Tagen immer noch gesund zu sein gegeben durch $(1 - p_t)^n$.

Berechnet man diese mit $p_{5. \text{März } 2020} = 0,00029\%$, dann ist die Wahrscheinlichkeit, nach einer Woche noch gesund zu sein bei 99,998%. Die Wahrscheinlichkeit, nach einem Monat noch gesund zu sein liegt immerhin noch bei 99,9914%. Dies klingt sehr hoch und suggeriert, es gäbe keine Notwendigkeit, individuelles Verhalten zu ändern.

Fragt man sich jedoch, wie hoch die Wahrscheinlichkeiten sind, dass alle Individuen in einer Gruppe von 1000 Einwohnern gesund bleiben, dann sinkt die Wahrscheinlichkeit für eine Woche auf 98,0%. Fragt man nach der Wahrscheinlichkeit, daß eine Gruppe von 1 Million Einwohnern gesund bleibt, dann liegt diese bei 0%.

3 Notwendigkeit behördlicher Anordnungen?

Denkt man aus volkswirtschaftlicher Sicht über die Notwendigkeit staatlichen Eingreifens nach (sollen Fußballspiele vor leeren Rängen stattfinden oder sollen Schulen geschlossen werden?), dann stellt sich routinemäßig die Frage, ob individuelle Entscheidungen zu einem gesellschaftlich wünschenswerten Ergebnis führen. Bei individuellen Entscheidungen wird üblicherweise davon ausgegangen, dass ein Individuum alle Einflüsse seines Handelns auf sein Wohlergehen berücksichtigt (die Effekte auf andere Individuen jedoch vernachlässigt).⁵ Aus Sicht dieses sehr hilfreichen Analyserahmens folgt unmittelbar, dass ein Individuum zwar vorsichtiger ist in seinen Handlungen, wenn daraus das Risiko einer *individuellen* Infektion erwächst. Das Individuum wird aber nicht ausreichend die Effekte auf *andere* Individuen in seinem Umfeld berücksichtigen. Diese sind jedoch durch die Ansteckungsgefahr für andere Menschen gegeben.⁶

Aufgrund dieser Ansteckungsgefahr ist es Aufgabe des Staates bzw. der zuständigen Behörden, Maßnahmen zu ergreifen, die über die selbstgewählte Vorsicht einzelner hinausgeht. Auch wenn individuell nicht gewünscht kann es gesellschaftlich wünschenswert sein, Großveranstaltungen zu untersagen oder Schulen zu schließen.

Idealerweise würden Entscheidungen dieser Art auf einem Modell aufbauen, das obige Infektionsraten in die Zukunft projiziert. In einem solchen Rahmen könnten dann, ebenfalls idealerweise auf regionaler Ebene, die Auswirkungen verschiedener Maßnahmen detailliert untersucht werden.

Betrachtet man die wachsenden Wachstumsraten in Abbildung 1 rechts (wobei 7 Beobachtungen nicht viele sind, um von einem Wachstumsprozess reden zu können), dann erscheint es logisch, dass öffentliche Maßnahmen zu einem früheren Zeitpunkt besser sind als zu einem späteren Zeitpunkt. Jede nicht stattgefundene Infektion heute reduziert die zukünftigen Infektionen um ein Vielfaches.⁷

⁵Das ist die in Teilen realistische aber zum Glück nicht in allen Lebens- und auch nicht Berufsbereichen relevante Annahme des egozentrierten homo oeconomicus.

⁶Dies ist ein klassisches Beispiel einer negativen Externalität.

⁷Die Version vom 6. März 2020 ist die erste Version dieser Notiz. Kommentare, gerne auch aus epidemiologischer Sicht, sind sehr erwünscht. Erweiterungen dieser Notiz, gerne auch mit weiteren Autoren, sind geplant.

A Anhang

A.1 Die Berechnungen

Die Bevölkerung in der Bundesrepublik wird mit $N = 83,1$ Millionen Einwohnern angenommen.⁸ Die Anzahl der Erkrankten E_t am Tag t kommt von der oben erwähnten Seite des RKI. Somit ergibt sich

$$p_t = \frac{E_{t+1} - E_t}{N - E_t} \quad (\text{A.1})$$

Da Ansteckung und Symptome nicht zusammenfallen sondern Symptome erst nach einer Inkubationszeit von n Tagen ausbrechen, wäre die Erkrankungswahrscheinlichkeit in t besser berechnet durch

$$p_t^{Inkub} = \frac{E_{t+1+n} - E_{t+n}}{N - E_{t+n}}.$$

Man betrachtet für die Erkrankungswahrscheinlichkeit heute also den Anstieg der Erkrankungen in 1 Woche (für $n = 7$) relativ zur gesunden Bevölkerung in 1 Woche. Mit dieser Sichtweise wäre die Erkrankungswahrscheinlichkeit am 27. Februar 2020 also so hoch gewesen, wie in Abbildung 2 (links) für den 5. März angegeben.

⁸https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/_inhalt.html