



**VIT<sup>®</sup>**

**Vellore Institute of Technology**

(Deemed to be University under section 3 of UGC Act, 1956)

Programme	:	<b>M.Tech Software Engineering</b>	Semester	:	<b>Fall2020</b>
Course	:	<b>Foundations of Data Analytics</b>	Code	:	<b>CSE3505</b>
Faculty	:	<b>Prof.Rajalakshmi R</b>	Slot	:	<b>D1</b>

## **PROJECT TITLE:**

**ESTIMATING THE RISK OF PARTYING IN THIS  
PANDEMIC**

## **TEAM MEMBERS:**

**P.Jyoshna – 17MIS1153**

## **INTRODUCTION:**

There is no doubt that, every now and then, one ought to celebrate life. This usually involves people coming together, talking, laughing, dancing, singing, shouting; simply put, it means throwing a party. With temperatures rising, summer offers all the more incentive to organize such a joyous event. Blinded by the light, it is easy to forget that we are, unfortunately, still in a pandemic. In this project we estimate and plot the risk of partying using R.

## **ABSTRACT:**

The risk of a party as the probability that somebody who is infected with the novel coronavirus and can spread it attends the gathering. The two major factors influencing this probability are the size of the party, that is, the number of people attending; and the prevalence of infected people in the relevant population. As we will see, the latter quantity is difficult to estimate. The probability of actually getting infected by a person who has the coronavirus depends further on a number of factors.

Specifically, we care about the number of currently infectious cases. This is different from the number of currently infected cases. Upon infection, it usually takes a while until one can infect others, with estimates ranging from 1 – 3 days before showing symptoms. The incubation period is the time it takes from getting infected to showing symptoms. It lasts about 5 days on average, with the vast majority of people showing symptoms within 12 days. Yet about a third to a

half of people can be infectious without showing any symptoms. Estimates suggest that one is infectious for about 8 – 10 days, but it can be longer.

Given the population of the following countries and the cumulative number of reported infections, we can compute the reported attack rate. Relating this to the estimate of the true attack rate gives us an indication of the extent that the reports undercount the actual infections.

### **SOURCE:**

<https://www.rivm.nl/documenten/wekelijkse-update-epidemiologische-situatie-covid-19-in-nederland>

## **LITERATURE REVIEW:**

### **1) Effective risk communication for the prevention and control of communicable diseases in Europe( [www.ecdc.europa.eu](http://www.ecdc.europa.eu))**

The literature indicates that the theories and practices of risk communication were extended markedly to the issue of communicable diseases around the mid-1990s, with initial studies building on health risk research by Covello, Paustenbach, Detlof von Winterfeldt and Slovic, amongst others. Nicoll & Murray and Reynolds & Seeger suggest that risk communication was embraced in the field of public health as ‘threats to health the likes of tuberculosis, HIV, influenza, anti-microbial resistance, chemical accidents and bio-terrorism’ were recognised as not only ubiquitous but likely to be increasingly common ‘with increased movements of people, animals and goods, climate change and industrialisation’ . In an era characterised by uncertainty, rapid change and globalisation, where national borders provide no barriers to the transmission of communicable diseases, and with new diseases emerging and the re-emergence of other diseases, the importance of models and guidelines for effective risk communication is clear. In addition, Reynolds & Seeger note that the perceived threat and intentional nature of an anthrax attack in the United States in 2001, followed soon after by the emergence and global spread of SARS, ‘placed significant pressure on the public health community to communicate effectively within context of immediacy, threat and high uncertainty’ . This environment highlighted the weaknesses of the risk communication guidelines and models of the time for responding effectively to the threats to health that communicable diseases present.

### **2) Ethical Guidance for Rationing Scarce Health-Related Resources in a Severe Influenza Pandemic(<https://www.health.state.mn.us/communities/ep/surge/crisis/litrev>)**

Pandemic influenza preparedness plans vary in their underlying assumptions about:

- (1) whether the pandemic will be mild, moderate, or severe.
- (2) what the supply and availability of scarce health-related resources will be during different times throughout the pandemic; and
- (3) what information will be available about who is at risk, who can benefit, and other operational matters.

There is also considerable variety among pandemic plans in their attention to guidance for ethical decision-making; many pandemic plans contain no ethical guidance at all. This is due, at least in part, to the fact that the assumptions about the pandemic are left vague given all the uncertainties about what it will actually look like. While this ambiguity is fine for allowing flexibility in a planning document, it makes the development of ethical guidance more difficult as goals and strategies may change significantly depending upon the severity of the pandemic.

### **3) Pandemics: Risks, Impacts, and Mitigation**

**Nita Madhav, Ben Oppenheim, Mark Gallivan, Prime Mulembakani, Edward Rubin, and Nathan Wolfe.**

Preparing for a pandemic is challenging because of a multitude of factors, many of which are unique among natural disasters. Pandemics are rare events, and the risk of occurrence is influenced by anthropogenic changes in the natural environment. In addition, accountability for preparedness is diffuse, and many of the countries at greatest risk have the most limited capacity to manage and mitigate pandemic risk.

Unlike most other natural disasters, pandemics do not remain geographically contained, and damages can be mitigated significantly through prompt intervention. As a result, there are strong ethical and global health imperatives for building capacity to detect and respond to pandemic threats, particularly in countries with weak preparedness and high spark and spread risk.

Investments to improve pandemic preparedness may have fewer immediate benefits, particularly relative to other pressing health needs in countries with heavy burdens of endemic disease. Therefore, characterizing pandemic risk and identifying gaps in pandemic preparedness are essential for prioritizing

and targeting capacity-building efforts. Thinking about risks in terms of frequency and severity, notably using probabilistic modeling and EP curves, can quantify the potential pandemic risks facing each country and clarify the benefit-cost case for investing in pandemic preparedness.

No single, optimal response to a public health emergency exists; strategies must be tailored to the local context and to the severity and type of pandemic. However, overarching lessons emerge after multiple regional epidemics and global pandemics. For example, because of their high spark and spread risks, many LMICs would benefit most from building situational awareness and health care coordination capacity; public health response measures are far more cost-effective if they are initiated quickly and if scarce resources are targeted appropriately.

Building pandemic situational awareness is complex, requiring coordination across bureaucracies, across the public and private sectors, and across disciplines with different training and different norms (including epidemiology, clinical medicine, logistics, and disaster response). However, an appropriately sized and trained health workforce (encompassing doctors, nurses, epidemiologists, veterinarians, laboratorians, and others) that is supported by adequate coordination systems is a fundamental need—the World Health Organization has recommended a basic threshold of 23 skilled health professionals per 10,000 people.

Increasing the trained health workforce also will increase the capacity to detect whether any particular population (for example, human, farm animal, or wildlife) is suffering from a pathogen with high pandemic risk. Increasing the health workforce also will improve the overall resiliency of the health system, an improvement that can be applied to any emergency that results in morbidity and mortality shocks.

#### **4) Economics of COVID-19**

**(Abel Brodeur David Gray Anik Islam Suraiya Jabeen Bhuiyan)**

Before delving into the impacts of COVID-19 and government response on the economy, we documented the most popular data sources to measure COVID-19 known cases/deaths and social distancing. We first pointed out that COVID-19 cases and deaths suffer from measurement error due to many factors including testing capacity. Mobility measures using GPS coordinates from cell phones have been used extensively to measure social distancing.

However, there are certain caveats that apply, particularly in terms of privacy concerns and the representativeness of data. The paper also reviewed different research related to social distancing itself, particularly in regards to its determinants, its effectiveness in mitigating the spread of COVID-19, and its compliance. Going forward, social distancing and its measurements will continue to play a key role in academic research and policy development. 45 Going forward, the policy measures related to COVID-19 will continue to be an important area of research. These measures, which have varied both in terms of scope and implementation, are expected to yield a profound economic and social impact. This study tried to bring coherence to these issues by covering different public health and economic stimulus measures as well as providing a review of the literature on policy determinants, optimal lockdown measures, factors affecting lifting of lockdown, and combinations of fiscal and monetary policy measures aimed at ‘flattening the recession curve’.

## **5) Clinical Presentation, and Management of the Pandemic    Coronavirus Disease 2019 (COVID-19)**

**Pramath Kakodkar, Nagham Kaka, MN Baig**

This COVID-19 pandemic is a reminder of the volatility in the ongoing planning to manage the primary and secondary infection of SARS-CoV-2. This planning can be improved by accurate modeling of current data and by eliminating the misinformation in our era of data surplus. Additional variables that can strengthen countermeasure to this pandemic are rapidly updating surveillance data, availability of robust accredited information, and a multidisciplinary approach that bridges the gap in knowledge between basic sciences and clinical sciences.

This literature review comprehensively summarizes the most relevant study relating to the individual parameters that influence the clinical course and management of COVID-19. Due to the lack of available and validated therapeutics, most of the countermeasures rely on the usage of public health containment and quarantine approaches. Primary learning points from this COVID-19 pandemic are to uphold transparency to prevent delays in threat identification.

## **CODE:**

```
library('dplyr')
```

```
library('COVID19')
```

```
get_undercount <- function(country, attack_rate) {
```

```
  dat <- covid19(country = country, end = '2020-05-04', verbose = FALSE)
```

```
  dat %>%
```

```
    group_by(id) %>%
```

```
    summarize(
```

```
      population = head(population, 1),
```

```
      total_cases = tail(confirmed, 1)
```

```
    ) %>%
```

```
    mutate(
```

```
      attack_rate = attack_rate,
```

```
      reported_attack_rate = 100 * total_cases / population,
```

```
      undercount_factor = attack_rate / reported_attack_rate
```

```
    )
```

```
  }
```



```
get_undercount(c('Germany', 'Spain', 'United
Kingdom', 'India', 'Turkey', 'Pakistan', 'Bangladesh', 'Australia', 'Canada', 'Denmark', 'Fin
land'), c(0.85, 5.5, 5.10, 4.58, 7.62, 2.90, 2.56, 2.47, 1.35, 3.97, 3.43))
```

```
# A tibble: 11 x 6
  id      population total_cases attack_rate reported_attack_rate undercount_factor
  <chr>      <int>      <int>      <dbl>      <dbl>      <dbl>
1 AUS      24982688      7014      0.85      0.0281      30.3
2 BGD     161356039     10143      5.5      0.00629     875.
3 CAN      37057765      60761      5.1      0.164       31.1
4 DEU      82905782     165094      4.58      0.199       23.0
5 DNK       5793636      9670      7.62      0.167       45.7
6 ESP     46796540     218011      2.9      0.466        6.22
7 FIN       5515525      5327      2.56      0.0966      26.5
8 GBR     66460344     190584      2.47      0.287        8.61
9 IND     1352617328     46434      1.35      0.00343     393.
10 PAK     212215030     22049      3.97      0.0104     382.
11 TUR      82319724     127659      3.43      0.155       22.1
```

## Estimating the risk of a party:

We estimated that INDIA had 5 times  $20.4 = 102$  true infectious cases per 100,000 inhabitants. Assuming that the probability of infection is the same for all citizens, this results in  $102 / 100,000 = 0.00102$ , which gives a 0.102% or 1 in 980 chance that a *single* party guest has the virus and can spread it.

A party with just one guest would be — *intimate*. So let's invite a few others. What are the chances that *at least one* of them can spread the virus? We compute this by first computing the complement, that is, the probability that *no* party guest is infectious.

The chance that any one person from INDIA is not infectious is  $1 - 0.00102 = 0.9990$ , or 99.90%. With our assumption of guests being an independent random sample from the population, the probability that none of the  $n$  guests can spread the virus is  $0.9990^n$ .

The functions below compute the probability that no party guests can spread the virus, as well as the probability that *at least one* of the guests is infectious; the latter probability is simply the complement of the former.

```
# Probability that no guest is infectious
```

```
prob_virus_free <- function(n, true_relative_cases = 20.4 * 5) {  
  prob_virus <- true_relative_cases / 100000  
  (1 - prob_virus)^n  
}
```

```
# Probability that at least one guest is infectious
```

```
party_risk <- function(n, true_relative_cases = 20.4 * 5) {  
  1 - prob_virus_free(n, true_relative_cases)  
}
```

```
plot_risk <- function(ns, true_relative_cases = 20.4 * 5, ...) {  
  plot(  
    ns, 1 - prob_virus_free(ns, true_relative_cases = true_relative_cases),  
    type = 'l', axes = FALSE,  
    xlab = 'Party size', ylab = 'Probability of at least one infection',
```

```
  xaxs = 'i', yaxs = 'i', ...  
)  
axis(1)  
axis(2, las = 2)  
}
```

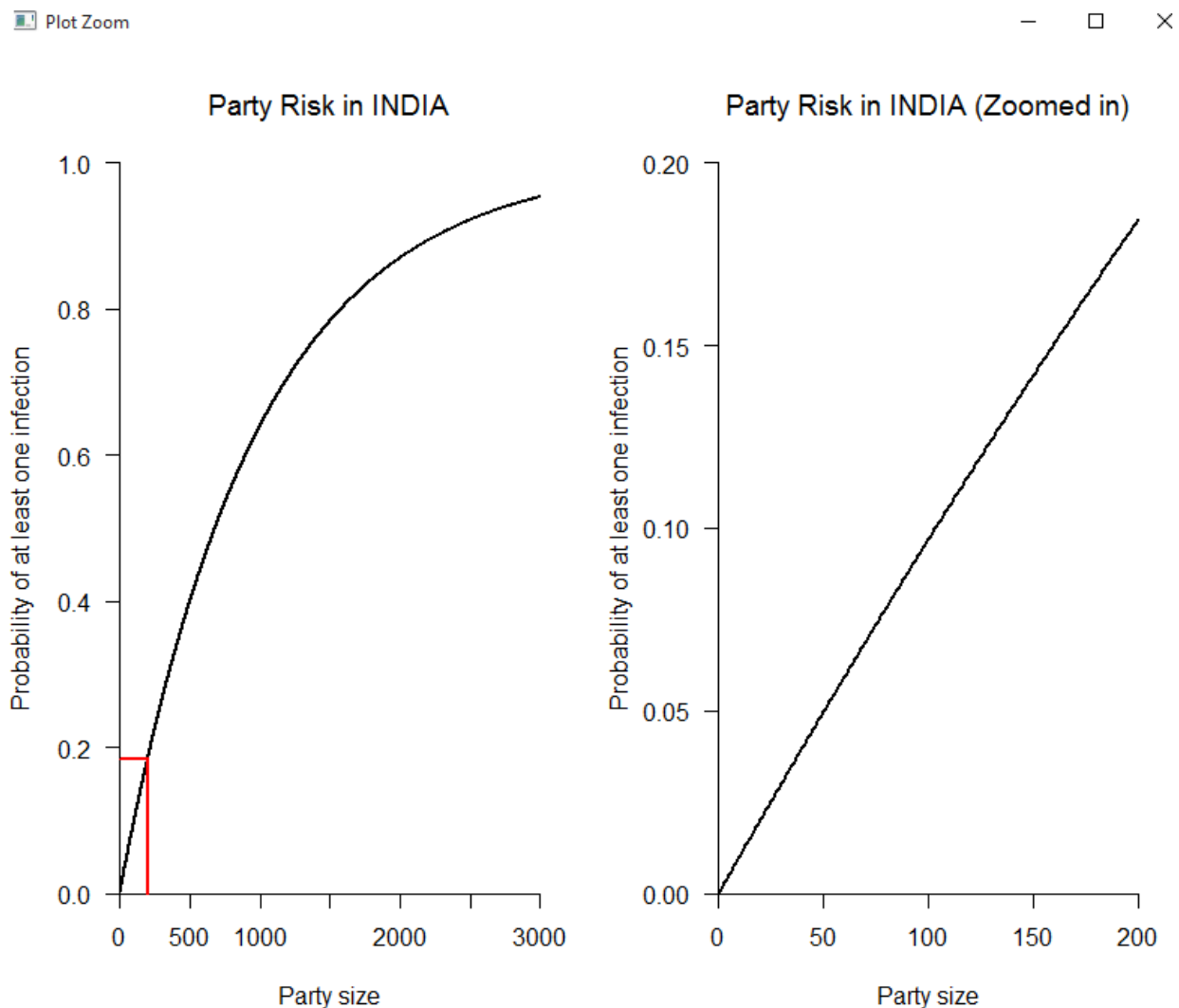
```
par(mfrow = c(1, 2))  
n1 <- 3000  
n2 <- 200  
y2 <- 1 - prob_virus_free(n2)
```

```
plot_risk(  
  seq(0, n1), lwd = 2, main = 'Party Risk in INDIA',  
  xlim = c(0, n1), ylim = c(0, 1), font.main = 1.5  
)  
lines(c(n2, n2), c(0, y2), col = 'red', lwd = 2)  
lines(c(0, n2), c(y2, y2), col = 'red', lwd = 2)
```

```
plot_risk(  
  seq(0, n2), lwd = 2, main = 'Party Risk in INDIA (Zoomed in)',
```

```
ylim = c(0, 0.20), xlim = c(0, n2), font.main = 1.5
```

)



The left panel shows how the party risk — the probability that at least one infectious person is the party — increases with  $n$ . In particular, to have near certainty that at least one infectious person shows up requires a very large party. The right panel zooms in on reasonably party sizes. Most parties that are thrown

indoors probably do not exceed 100 attendants, depending on how rich and reckless the host is.

The estimate of the party risk based on this simple calculation are: there is a 2.02, a 4.97, and a 14.19 chance that at least one guest can spread the coronavirus for parties of size 20, 50, and 150, respectively.

### **Uniform infection probability:**

At the same time, however, our calculations assume that the risk of getting the coronavirus is evenly spread across the population. We used this fact when estimating the probability that any one person has the coronavirus as the total number of cases divided by the population size.

The probability of infection is not, however, evenly distributed. For example, Pollán et al. (2020) report a seroprevalence of people aged 65 or more of about 60, while people aged between 20 and 34 showed a seroprevalence of 4.4 between April 27<sup>th</sup> to May 11<sup>th</sup>. These days, however, there is a substantial rise in young people who become infected, in the United States and likely also in Europe. Because young people are less likely to develop symptoms, the virus can spread largely undetected.

Moreover, it seems to me that people who would join a party are in general more *adventurous*. This would increase the chances of an infectious person attending a party; thus, our calculation above may in fact underestimate the party risk.

At the same time, one would hope that people who show symptoms avoid parties. If all guests do this, then only pre-symptomatic or asymptomatic spread can occur, which would reduce the party risk by a half up to two thirds. On the flip side, people who show symptoms might get tested for COVID-19 and, upon receiving a negative test, consider it safe to attend a party. This might be foolish, however;

recent estimates suggest that tests miss about 20% infections for people who show symptoms. For people without symptoms, the test performs even worse.

For parties taking place in summer, it is not unlikely that many guests engaged in holiday travel in the days or weeks before the date of the party. Since travel increases the chances of infection, this would further increase the chances that at least one party guest has contracted the coronavirus.

### **Sensitivity analysis:**

We have assumed that the reported number of infected cases undercounts the true number of infectious cases by a factor of 5. In particular, we used the reported cases of 20.4 per 100,000 and applied a correction factor of 5. In the following, we visualize the party risk as a function of the estimated true infectious cases per 100,000 inhabitants .

Between July 8th and July 21st, 20.4 cases per 100,000 inhabitants were reported. A correction factor of 5 would bolster this to 102 cases, a correction factor of 10 to 204 cases, and so on; thus, one can back calculate the correction factor from the estimated true number of cases.

The code below visualizes the probability that at least one party guest has the coronavirus and can spread it as a function of the estimated true number of infectious cases per 100,000 inhabitants and the size of the party.

### **CODE:**

```
library('RColorBrewer')
```

```

# Calculates the party size that results in 'prob_virus_free'
# for a given 'true_relative_cases'

get_party_size <- function(prob_virus_free, true_relative_cases) {
  log(prob_virus_free) / log(1 - true_relative_cases / 100000)
}

plot_total_risk <- function(party_sizes, true_relative_cases, ...) {

  plot(
    true_relative_cases, ns, type = 'n', xaxs = 'i',
    yaxs = 'i', axes = FALSE, ylab = 'Party Size',
    xlab = 'True Number of Infectious Cases per 100,000 Inhabitants', ...
  )

  ticks <- seq(0, 1000, 100)

  axis(1, at = ticks)
  axis(2, at = ticks, las = 2)

```

```
abline(h = ticks, col = 'gray86')
```

```
abline(v = ticks, col = 'gray86')
```

```
probs_virus <- seq(0.01, 0.99, 0.01)
```

```
party_sizes <- sapply(1 - probs_virus, get_party_size, true_relative_cases)
```

```
cols <- rev(heat.colors(50))
```

```
cols <- colorRampPalette(cols, bias = 2)(99)[-1]
```

```
ix <- c(seq(1, 55, 1), seq(60, 95, 5))
```

```
show_text <- c(seq(1, 10, 1), seq(15, 55, 5))
```

```
for (i in ix) {
```

```
  y <- party_sizes[, i]
```

```
  lines(true_relative_cases, y, col = cols[i], lwd = 2.5)
```

```
  j <- which.min(true_relative_cases[-1] + y[-1])
```

```
  if (i %in% show_text) {
```

```
    text(
```

```
      true_relative_cases[j] + 1, y[j] + 1, paste0(probs_virus[i] * 100, '%'),
```



```

        cex = 0.80
    )
}
}
}

ns <- seq(0, 300)

true_relative_cases <- seq(0, 1500, length.out = length(ns))

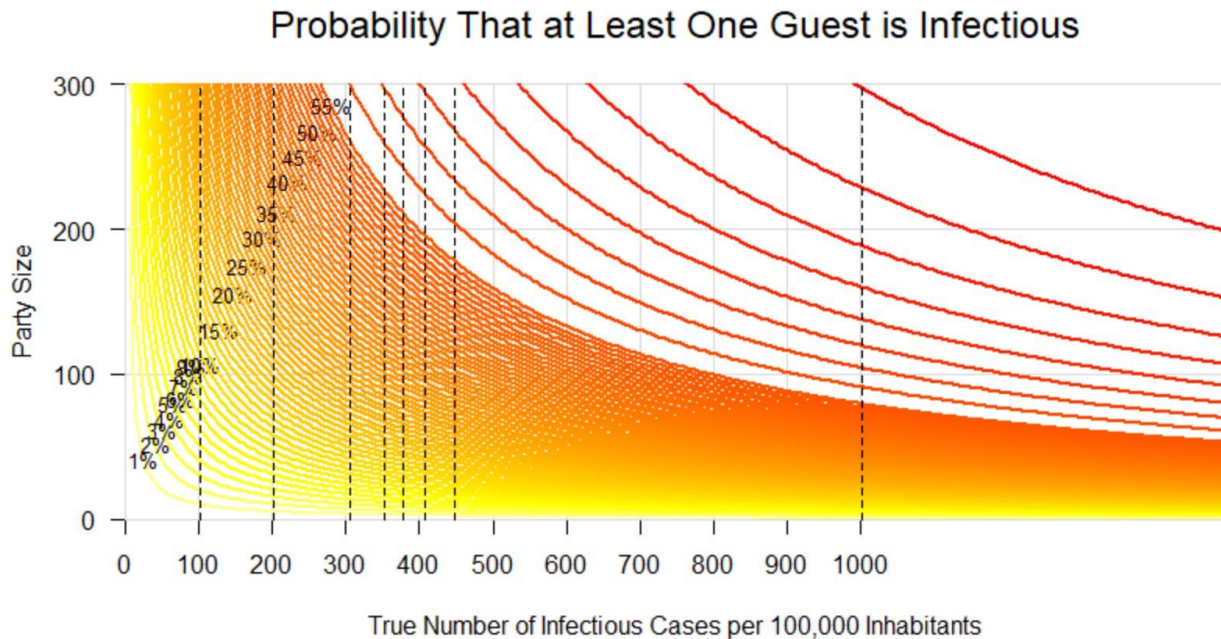
plot_total_risk(
  ns, true_relative_cases,
  main = 'Probability That at Least One Guest is Infectious',
  font.main = 1, cex.main = 1.5
)

lines(c(20.4 * 5, 20.4 * 5), c(0, 1000), lty = 2)
lines(c(40.1 * 5, 40.1 * 5), c(0, 1000), lty = 2)
lines(c(60.9 * 5, 60.9 * 5), c(0, 1000), lty = 2)
lines(c(70.3 * 5, 70.3 * 5), c(0, 1000), lty = 2)
lines(c(75.4 * 5, 75.4 * 5), c(0, 1000), lty = 2)
lines(c(81.7 * 5, 81.7 * 5), c(0, 1000), lty = 2)

```

```
lines(c(89.4 * 5, 89.4 * 5), c(0, 1000), lty = 2)
```

```
lines(c(200.6 * 5, 200.6 * 5), c(0, 1000), lty = 2)
```



Each colored line represents a combination of true number of infectious cases and party size that yields the same party risk. For example, attending a party of size 20 when the true number of infectious cases per 100,000 inhabitants is 50 yields a party risk of 1%, but so would, roughly, attending a party of size 10 when the true relative number of infectious cases is 100. Thus, there is a trade-off between the size of the party and the true number of infectious cases.

While we have computed the party risk for a single party, this risk naturally increases when you attend multiple ones. Suppose you have been invited to parties of size 20, 35, and 50 which will take place in the next month. Let's for simplicity assume that all guests are different each time. Let's further assume that the number of infectious cases stays constant over the next month; with a current estimate of  $R_t = 1.29$ , this seems unlikely. Together, these assumptions allow us to calculate

the *total* party risk as the party risk of attending a single party of size 105, which gives 10.16%. It seems that, in this case, fortune does not favor the bold.

## **ASSUMPTIONS**

The analysis shown is a very rough back-of-the-envelope calculation. We have made a number of crucial assumptions to arrive at some numbers

### **Independence:**

First, and most critically, we have assumed that party guests are a randomly and independently drawn from the population. It is this assumption that allowed us to compute the joint probability that none of the party guests have the virus by multiplying the probabilities of any individual being virus-free.

This means we are sampling not from the whole population, as our simple calculation assumes, but from some particular subpopulation that is well connected. Since the party guests likely share social circles or even households, the effective party size — in terms of being relevant for virus transmission — is smaller than the actual party size; this is because these individuals share the same risks. A party with 20 married couples seems safer than a party with 40 singles. This would suggest that we overestimate the risk of a party.

## Uniform infection probability:

At the same time, however, our calculations assume that the risk of getting the coronavirus is evenly spread across the population. We used this fact when estimating the probability that any one person has the coronavirus as the total number of cases divided by the population size.

The probability of infection is not, however, evenly distributed. These days, however, there is a substantial rise in young people who become infected, in the United States and likely also in Europe. Because young people are **less likely to develop symptoms**, the virus can spread largely undetected.

Moreover, it seems to me that people who would join a party are in general more adventurous. This would increase the chances of an infectious person attending a party; thus, our calculation above may in fact underestimate the party risk.

At the same time, one would hope that people who show symptoms avoid parties. If all guests do this, then only pre-symptomatic or asymptomatic spread can occur, which would reduce the party risk by a half up to two thirds. On the flip side, people who show symptoms might get tested for COVID-19 and, upon receiving a negative test, consider it safe to attend a party. For people without symptoms, the test performs even worse.

For parties taking place in summer, it is not unlikely that many guests engaged in holiday travel in the days or weeks before the date of the party. Since travel increases the chances of infection, this would further increase the chances that at least one party guest has contracted the coronavirus.

## **Estimating true infections:**

We have assumed that the number of new cases in the last two weeks equals the number of currently infectious cases. This is certainly an approximation. Ideally, we would have a geographically explicit model which, at any point in time and space, provides an estimate of the number of infectious cases. To my knowledge, we are currently lacking such a model.

If the 178 people who tested positive all self-isolate or, worse, end up in hospital, this clearly curbs virus spread compared to when they would continue to roam around. The former seems more likely. Moreover, these reported cases are likely not independent either, with outbreaks usually being localized. Similar to the fact that party guests know each other, the fact that reported cases cluster would lead us to overestimate the extent of virus spread.

At the same time, in the Netherlands only those that show symptoms can get tested. Since about a third to a half are asymptomatic or pre-symptomatic in the sense that they spread the virus considerably before symptom onset, the reported number of cases likely gives an undercount of infectious people.

All these complications can be summarized, roughly, in the correction factor, which gives the extent to which we believe that the reported number of cases deviates from the true number of infectious cases. We have first focused on a factor of 5, but then assessed the robustness of our results in a sensitivity analysis.

- ➔ For example, for a party size of 50, the chances that at least one guest is infectious is 1.84% or a correction factor of 1.824 (corresponding to the official RIVM estimate), 4.97% for a factor of 5.24, and 18.49% for a factor of 20.

## **Estimating the probability of infection:**

We have defined the party risk as the probability that at least one party guest has the coronavirus and is infectious. If this person does not spread the virus to other guests, no harm is done.

The probability of getting infected is a function of the time one is exposed to the virus, and the amount of virus one is exposed to. Estimates suggest that about 1,000 SARS-CoV-2 infectious virus particles suffice for an infection. With breathing, about 20 viral particles diffuse into the environment per minute; this increases to 200 for speaking; coughing or sneezing can release 200,000,000(!) virus particles. These do not all fall to the ground, but instead can remain suspended in the air and fill the whole room; thus, physical distancing alone might not be enough indoors. It seems reasonable to assume that, when party guests are crowded in a room for a number of hours, many of them stand a good chance of getting infected if any one of the guests is infectious.

It is different outdoors. A Japanese study found that virus transmission inside was about 19 times more likely than outside. Analyzing 318 outbreaks in China between January 4th and February 11th, it was found that only a single one occurred outdoors. This suggests that parties outdoors should be much safer than parties indoors. Yet outdoor parties feature elements unlike other outdoor events; for example, there are areas — such as bars or public toilets — which could become spots for virus transmission. Our simple calculations suggest, with a correction factor of 5, that the probability that at least one person out of 150 has the coronavirus is 14.19%. While, in contrast to an indoor setting, the infected person is unlikely to infect the majority of the other guests, it seems likely that at least some guests will get the virus.

## **CONCLUSION:**

My decision to throw (or attend) a party requires not only an estimate of how likely it is that the virus spreads at the gathering; it also requires an assessment of how much I actually care.

As argued above, it is almost certain that the virus spreads to other guests if one guest arrives infected. Noting that all guests are young, one might be tempted to argue that the cost of virus spread is low. In fact, people who party might even be helping — heroically — to build herd immunity!

First, while the proportion of infected people who die is very small for young people — Salje et al. (2020) estimate it to be 0.0045% for people in their twenties and 0.015% for people in their 30s. For some people, recovery can be lengthy — much longer than the two weeks we previously believed it would take. Known as “mild” cases, they might not be so mild after all. Moreover, the potential strange neurological effects of a coronavirus infection are becoming increasingly apparent.

## **REFERENCES:**

- Flaxman, Mishra, Gandy et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 3164.
- Ghinai, I., Woods, S., Ritger, K. A., McPherson, T. D., Black, S. R., Sparrow, L., ... & Arwady, M. A. (2020). Community Transmission of SARS-CoV-2 at Two Family Gatherings-Chicago, Illinois, February-March 2020. *MMWR. Morbidity and mortality weekly report*, 69(15), 446.
- Havers, F. P., Reed, C., Lim, T. W., Montgomery, J. M., Klena, J. D., Hall, A. J., ... & Krapivunaya, I. (2020). Seroprevalence of Antibodies to SARS-CoV-2 in Six Sites in the United States, March 23-May 3, 2020. *JAMA Internal Medicine*.
- He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., ... & Mo, X. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 26(5), 672-675.
- Klompas, M., Baker, M. A., & Rhee, C. (2020). Airborne Transmission of SARS-CoV-2: Theoretical Considerations and Available Evidence. *JAMA*.
- Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., & Lessler, J. (2020). Variation in false-negative rate of reverse transcriptase polymerase

chain reaction–based SARS-CoV-2 tests by time since exposure. *Annals of Internal Medicine*.

- Lachmann, M., & Fox, S. (2020). When thinking about reopening schools, an important factor to consider is the rate of community transmission. *Santa Fe Institute Transmission*.
- Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., ... & Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of Internal Medicine*, 172(9), 577-582.
- Morawska, L., & Cao, J. (2020). Airborne transmission of SARS-CoV-2: The world should face the reality. *Environment International*, 105730.
- Nishiura, H., Oshitani, H., Kobayashi, T., Saito, T., Sunagawa, T., Matsui, T., ... & Suzuki, M. (2020). Closed environments facilitate secondary transmission of coronavirus disease 2019 (COVID-19).