Epidemiology - Part 2

Feb. 13, 2025

Recap question:

Feb. 13, 2025

Which of the following statements are true when using Dorfman's method?

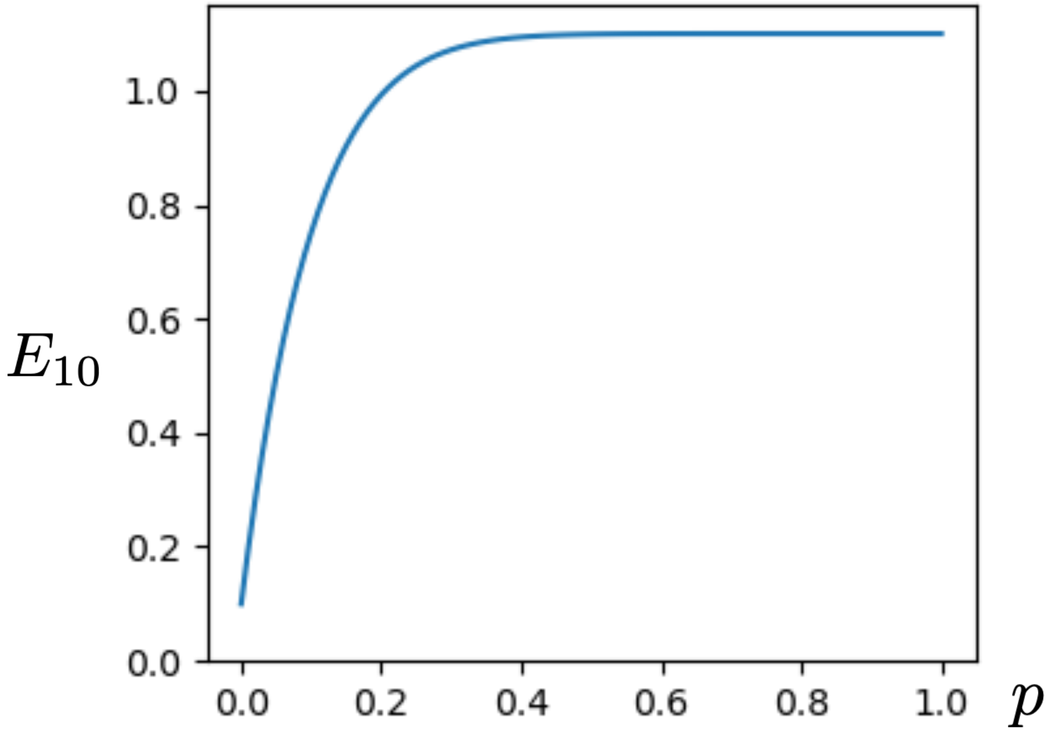
1. Sensitivity when using Dorfman's method is greater than the sensitivity of the underlying test
2. Sensitivity when using Dorfman's method is less than the sensitivity of the underlying test
3. Dorfman's method always uses fewer tests than testing each person individually

The sensitivity and specificity of Dorfman's method is

$$S_e' = S_e^2$$

The sensitivity is always at most 1, so squaring it gives a number that is not larger, which means that option 2 is correct and option 1 is incorrect.

Option 3 is also incorrect: if the prevalence is high enough, Dorfman's method does more work. Recall the graph for the expected number of tests per person



Epidemiology - Part 2

Feb. 13, 2025

By the end of this lecture, you will be able to:

1. Continue repeated testing
2. Reason about desired properties of a mathematical model
3. Reason about exponential growth
4. Define the SI model and the basic reproduction number R_0

Repeated testing

Repeated testing

Recall what we did for a single test: Consider one COVID test with 84% sensitivity and 99% specificity. (I = infected), and assume 10% of the population is infected

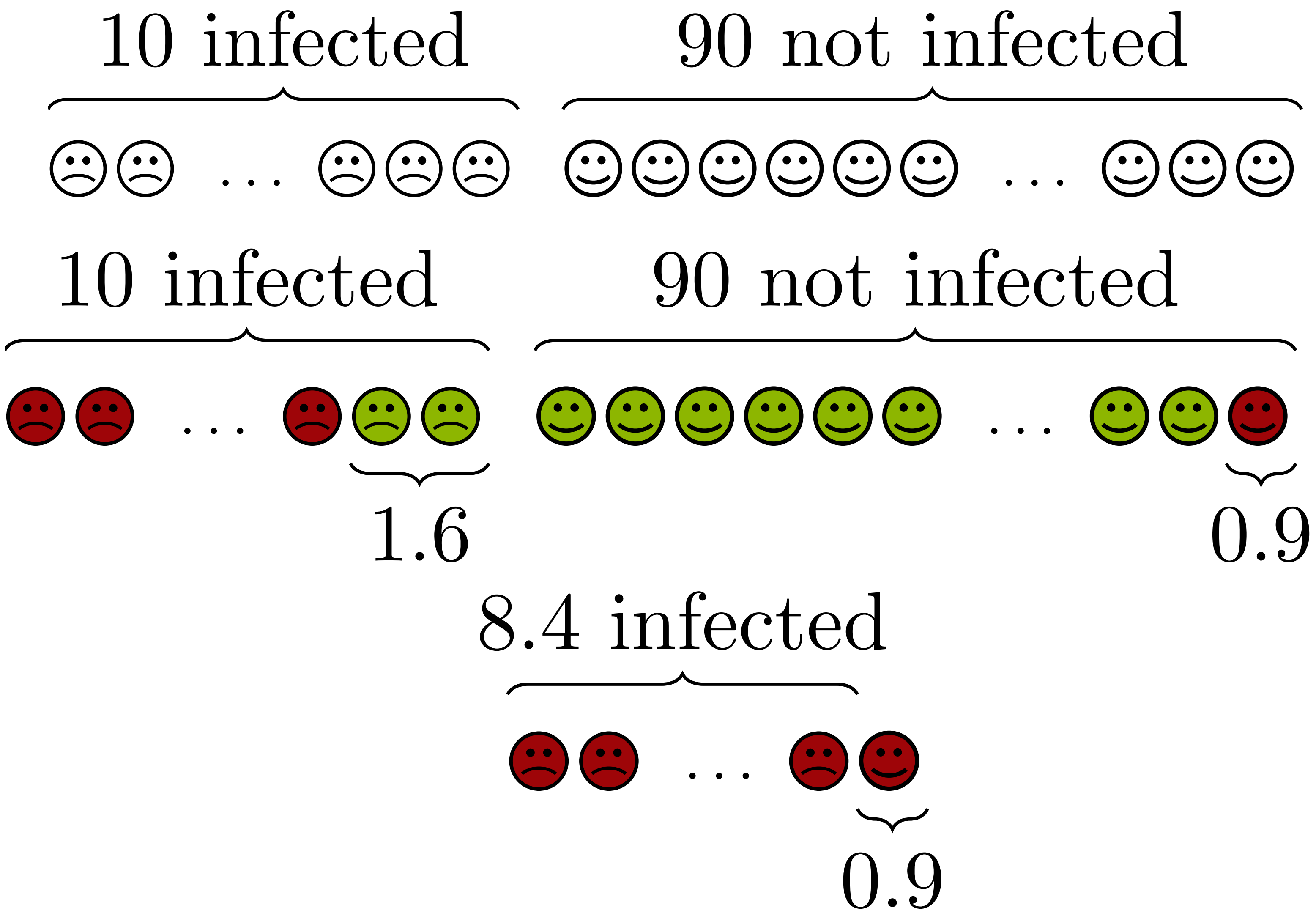
P(positive | I) = 0.84	True positive
P(negative | I) = 0.16	False negative
P(negative | nI) = 0.99	True negative
P(positive | nI) = 0.01	False positive

We used Bayes' rule to get the formula:

$$P(I|\text{positive}) = \frac{P(\text{positive}|I)P(I)}{P(\text{positive}|I)P(I) + P(\text{positive}|nI)P(nI)}$$

Example 1:

P(test positive | I) = 0.84	True positive
P(test negative | I) = 0.16	False negative
P(test negative | nI) = 0.99	True negative
P(test positive | nI) = 0.01	False positive



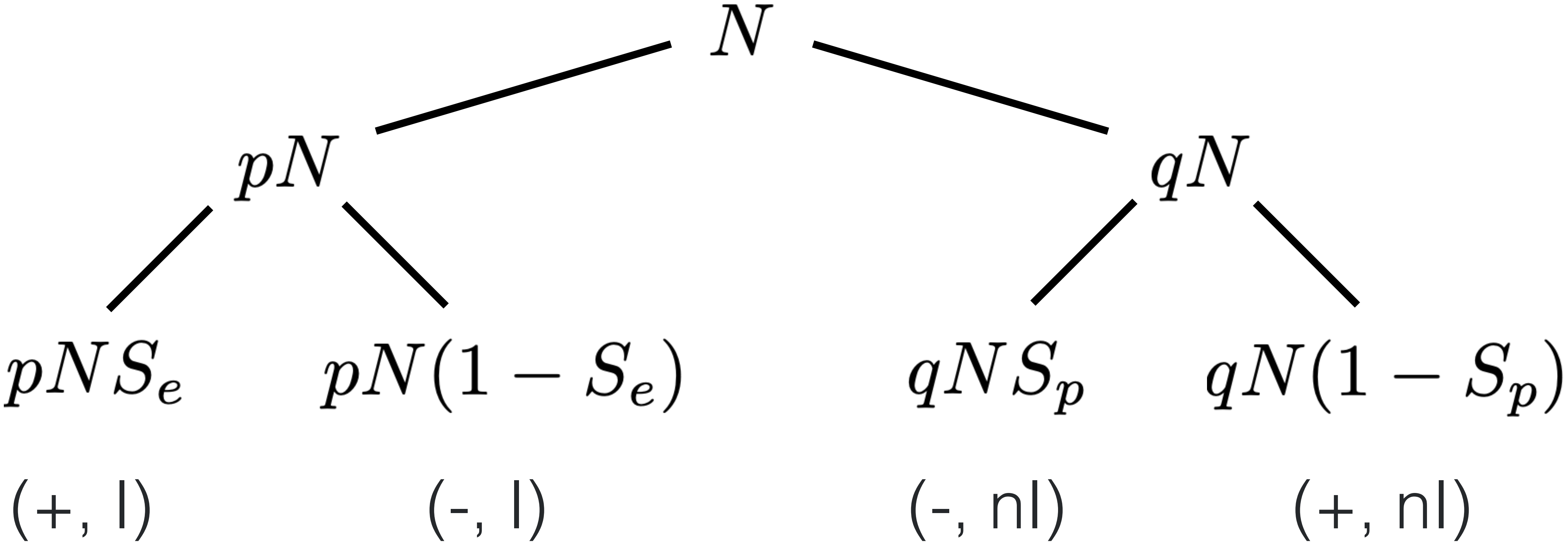
Repeated testing

Graphical illustration of the math:

S_e : sensitivity, S_p : specificity, N : population size

p : prevalence, $q = 1 - p$,

After the first test,

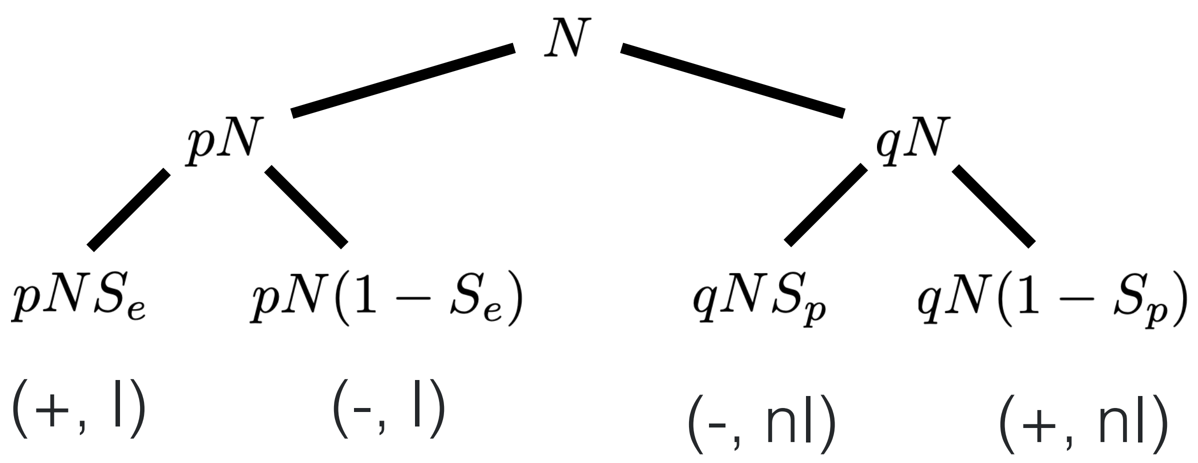


N

pN qN

pNS_e $pN(1 - S_e)$ qNS_p $qN(1 - S_p)$

(+, I) (-, I) (-, nI) (+, nI)



N

pN qN

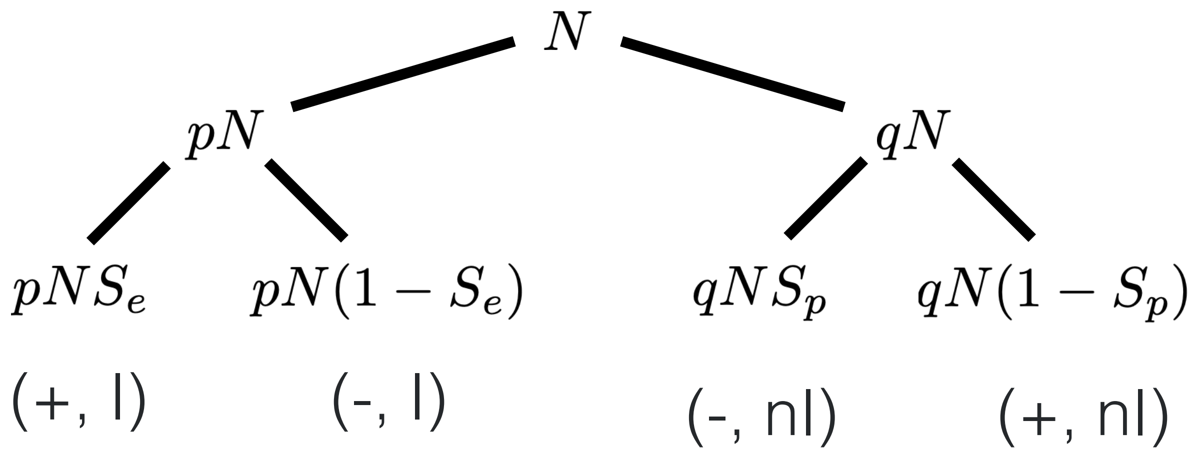
pNS_e $pN(1-S_e)$ qNS_p $qN(1-S_p)$

(+, I) (-, I) (-, nI) (+, nI)

$$N_+ = pNS_e + qN(1-S_p)$$

$$P(I|+) = \frac{pNS_e}{pNS_e + qN(1-S_p)} = \frac{pS_e}{pS_e + q(1-S_p)}$$

$$P(nI|+) = \frac{q(1-S_p)}{pS_e + q(1-S_p)}$$



N

pN qN

pNS_e $pN(1-S_e)$ qNS_p $qN(1-S_p)$

(+, I) (-, I) (-, nI) (+, nI)

$$N_+ = pNS_e + qN(1-S_p)$$

$$P(I|+) = \frac{pNS_e}{pNS_e + qN(1-S_p)} = \frac{pS_e}{pS_e + q(1-S_p)}$$

$$P(nI|+) = \frac{q(1-S_p)}{pS_e + q(1-S_p)}$$

For $S_e = 0.9,\ S_p = 0.95,\ p = 0.05,$

$P(I|+) \approx 0.49$

When testing again we do the same math, but only look at the individuals who tested positive in the first test, so the probability to be infected is now $p(I|+)$

In other words, we replace:

$$N \rightarrow N' = N_+ \text{ and } p \rightarrow p' = p(I|+)$$

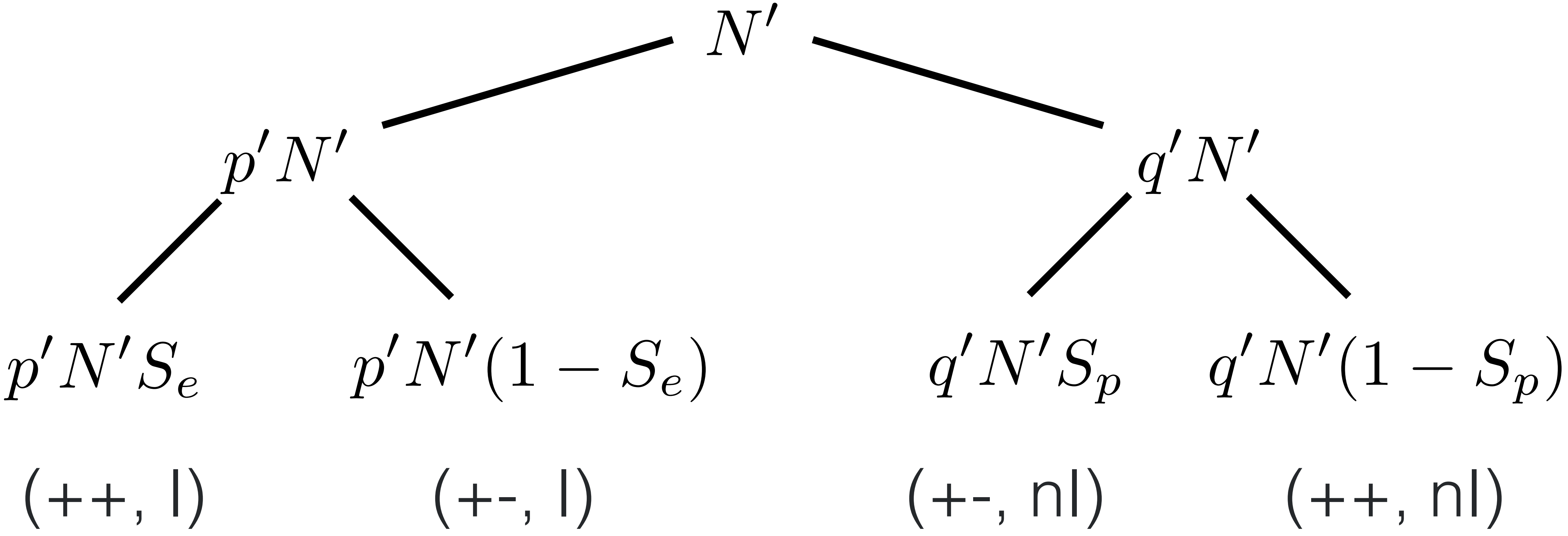
$$q \rightarrow q' = 1 - p'$$

When testing again we do the same math, but only look at the individuals who tested positive in the first test, so the probability to be infected is now $p(I|+)$

In other words, we replace:

$$N \to N' = N_+ \text{ and } p \to p' = p(I|+)$$

$$q \to q' = 1 - p'$$

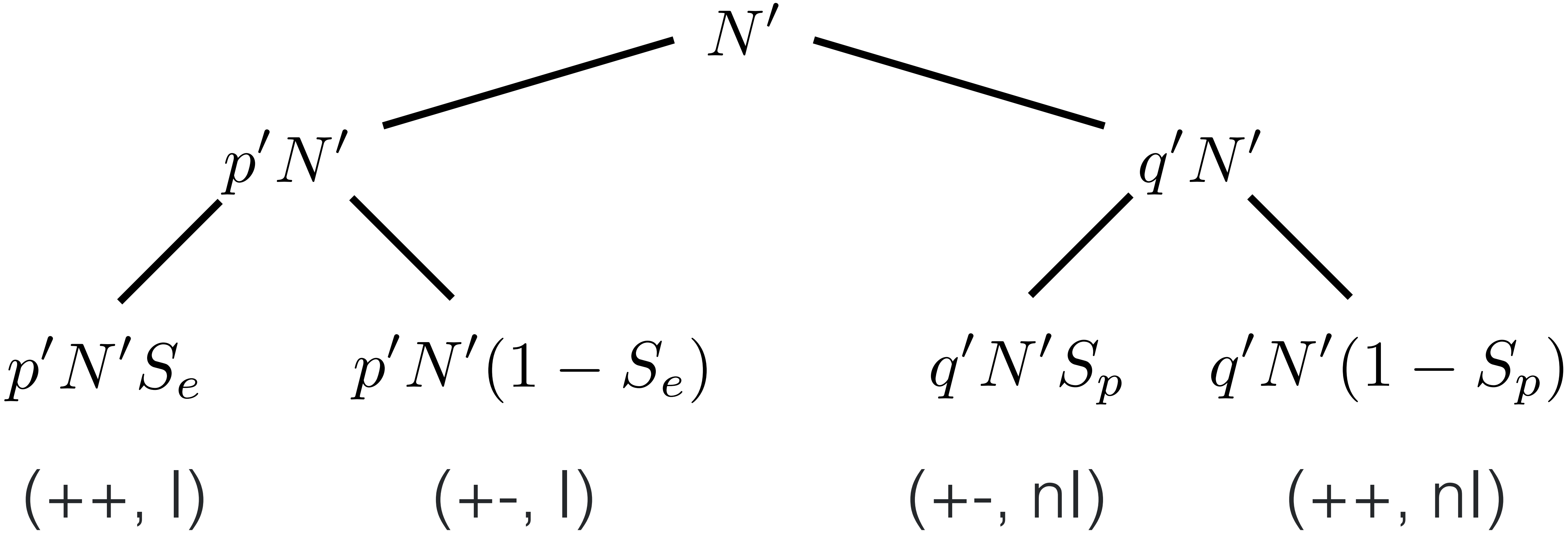


N'

$p'N'$ $\quad$ $q'N'$

$p'N'S_e$ $\quad$ $p'N'(1-S_e)$ $\quad$ $q'N'S_p$ $\quad$ $q'N'(1-S_p)$

(++, I) $\quad$ (+-, I) $\quad$ (+-, nI) $\quad$ (++, nI)



N'

$p'N'$ $q'N'$

$p'N'S_e$ $p'N'(1-S_e)$ $q'N'S_p$ $q'N'(1-S_p)$

(++, I) (+-, I) (+-, nI) (++, nI)

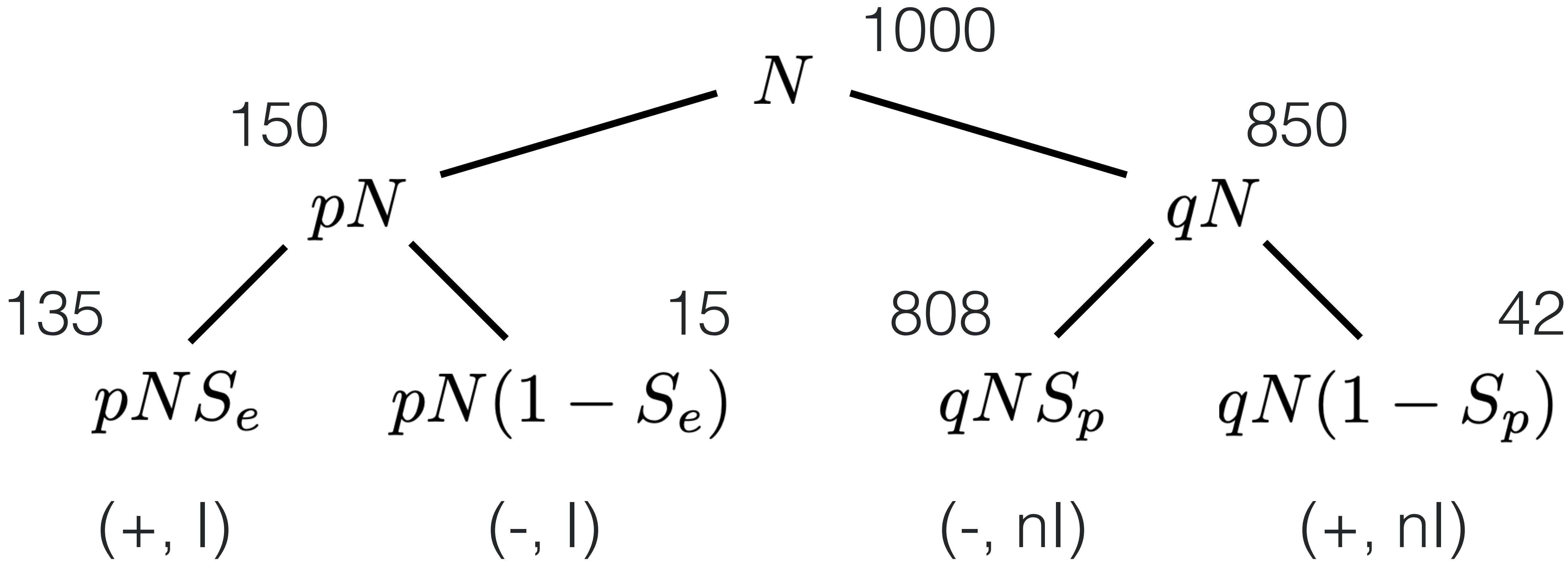
$$
\begin{aligned}
P(I|++) &= \frac{p'N'S_e}{p'N'S_e + q'N'(1-S_p)} = \frac{p'S_e}{p'S_e + q'(1-S_p)} \\
&= \frac{pS_e^2}{pS_e^2 + q(1-S_p)^2} \approx 0.94
\end{aligned}
$$

If we perform a third test, we test the samples whose first two test results are positive, and the same calculations give

$$p(I|{+}{+}{+}) = \frac{pS_e^3}{pS_e^3 + q(1 - S_p)^3} \approx 0.997$$

Estimating prevalence

$$S_e = 0.9,\ \ S_p = 0.95,\ \ p = 0.15$$



1000

N

150

pN

850

qN

135

pNS_e

15

$pN(1 - S_e)$

808

qNS_p

42

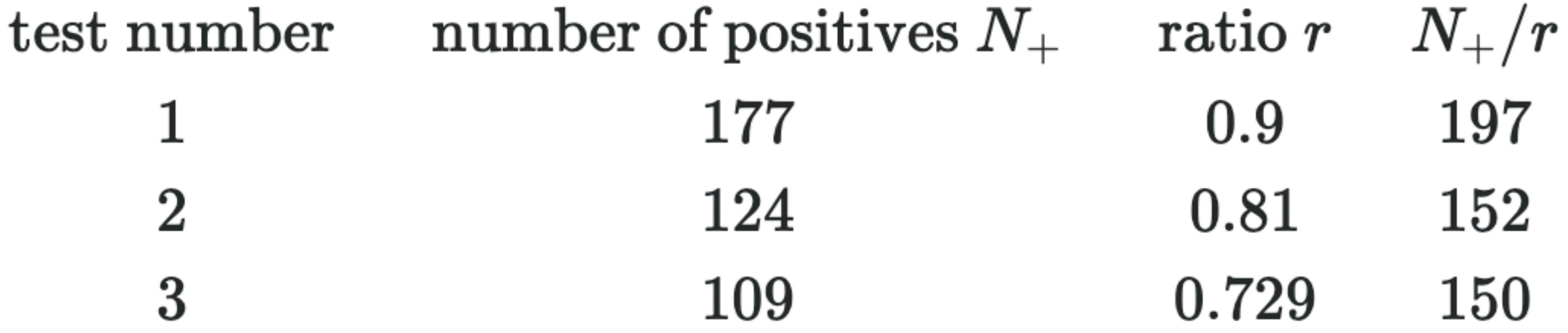
$qN(1 - S_p)$

(+, I)

(-, I)

(-, nI)

(+, nI)



test number	number of positives N_+	ratio r	N_+/r
1	177	0.9	197
2	124	0.81	152
3	109	0.729	150

If we neglect the number of false positives, the number of positives after k tests should be approximately S_e^k times the number of positive cases in the population, which can be used to estimate the total number of positives.

What should a model do?

Everything should be made as simple as possible, but not simpler.

\- Albert Einstein

What should a model do?

Provide a simplified explanation of complicated phenomena, that is still useful to predict the real world

Examples:

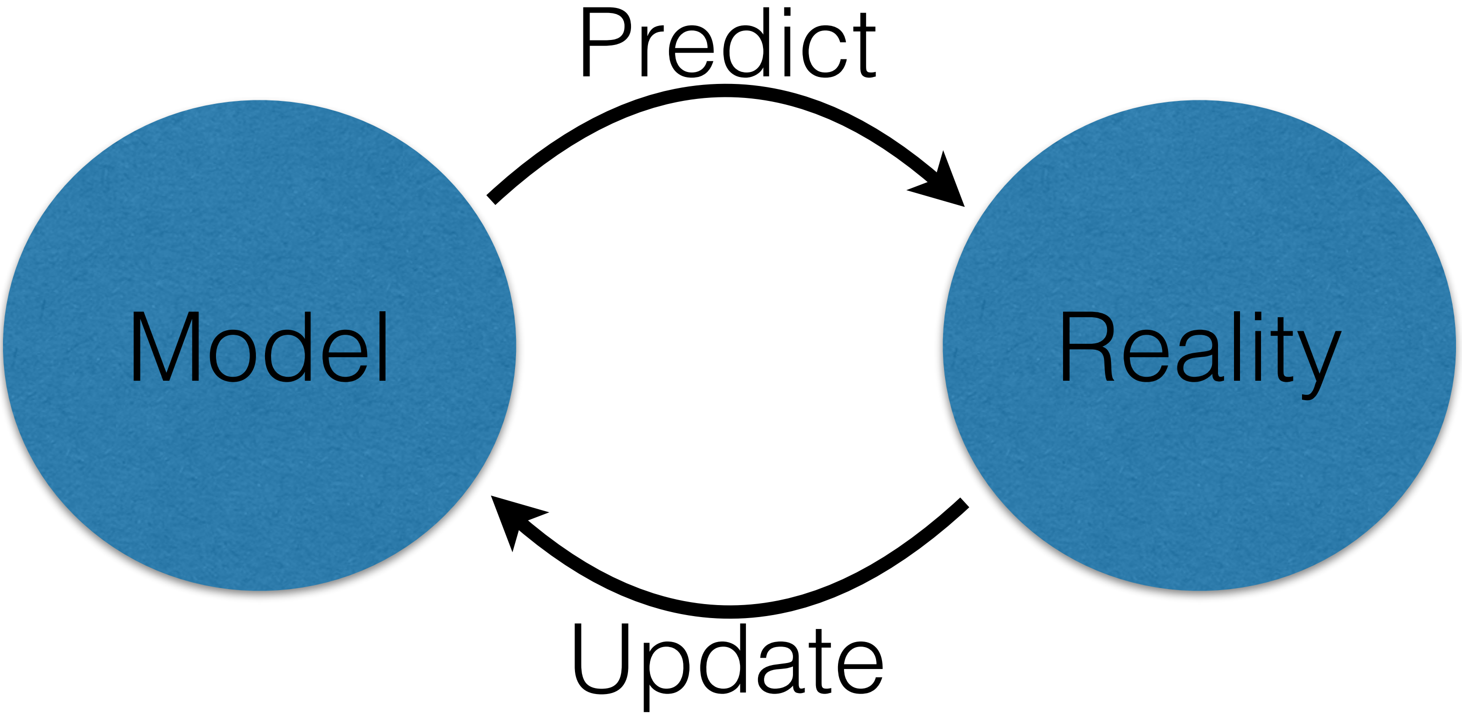
1. Newton's laws accurately predict many everyday phenomena, but has some limitations where theory of relativity or quantum mechanics have to be used
2. Climate models to study climate change
3. Economic models to study consumer behavior

What should a model do?

Models can change when phenomena are better understood, or new phenomena appear that cannot be explained with the old models

Example:

It took more than 200 years after Newton's laws were formulated for the theory of relativity to appear



What should a model do?

What kinds of predictions do we want to obtain from a model about epidemics?

What should a model do?

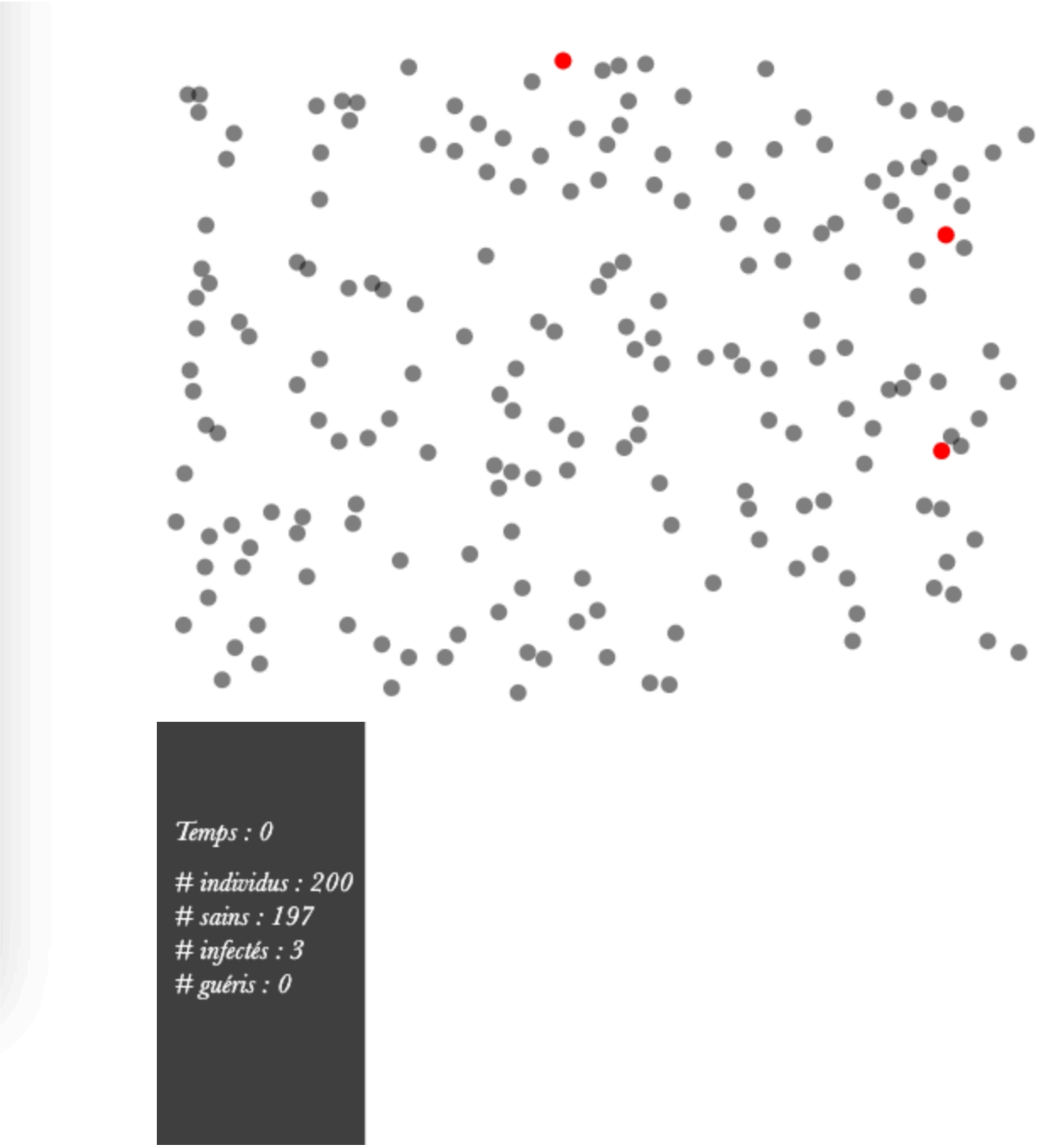
What kinds of predictions do we want to obtain from a model about epidemics?

1. Find out how many people are currently infected
2. Predict how this number will change in the long run: will everyone be infected at some point or will the disease die out?
3. Predict how different interventions change the rate of spread (vaccination, mask mandates, …)
4. …

Simplest model for epidemics

1. People move around in a city/country/continent
2. At the start, there is one infected person
3. Every time an infected person meets someone, the new person becomes infected with probability p

Simplest model for epidemics



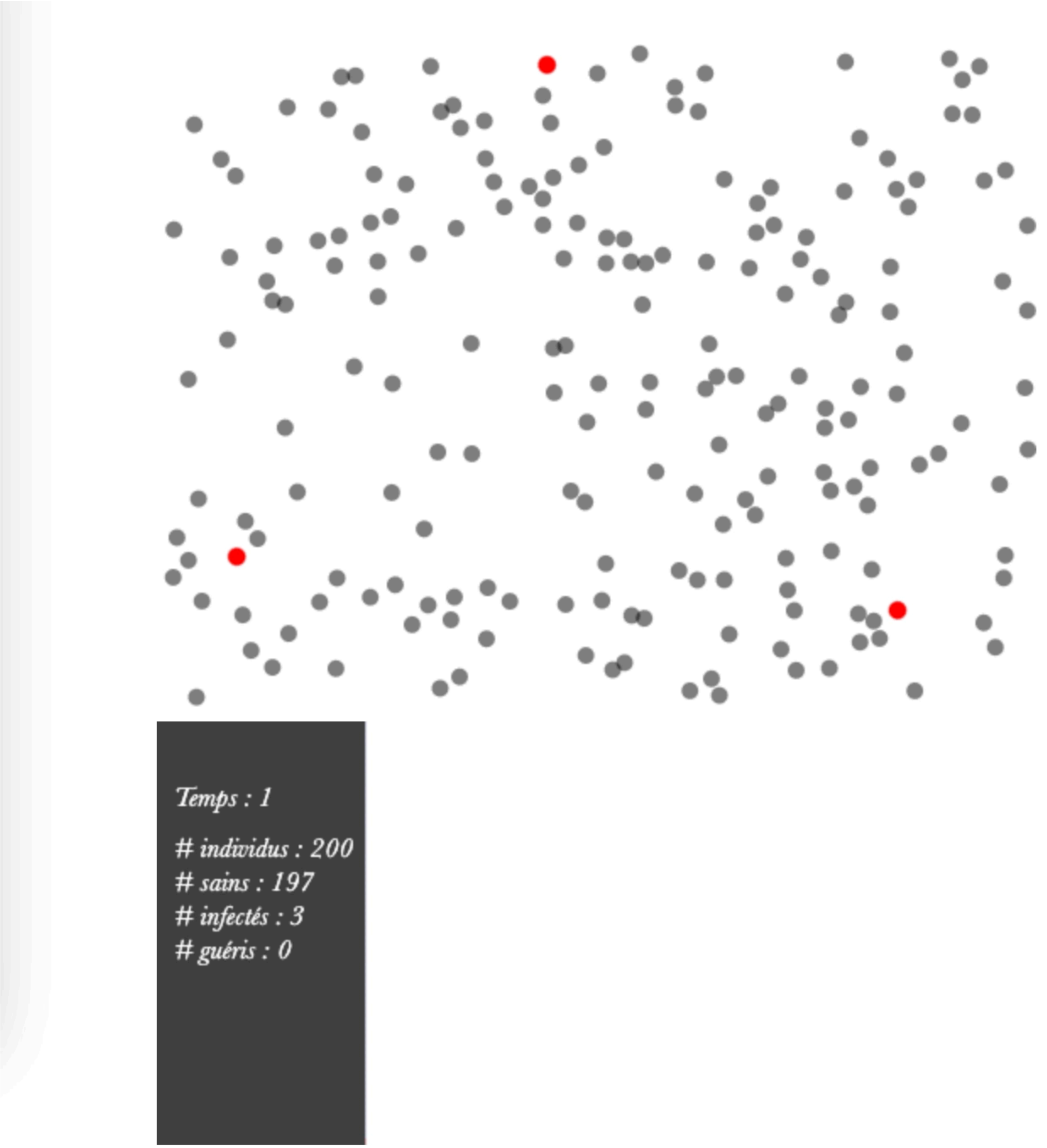
Temps : 0

 # individus : 200
 # sains : 197
 # infectés : 3
 # guéris : 0

Red = infected
Gray = not infected

https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-1.html?lang=fr

Simplest model for epidemics



Temps : 1

 # individus : 200
 # sains : 197
 # infectés : 3
 # guéris : 0

Red = infected
Gray = not infected
Green = recovered
and immune

https://images.math.cnrs.fr/Modelisation-d-une-epidemie-partie-1.html?lang=fr

Goal for the rest of today: derive expressions for the number of infected as time goes on.

Turns out to be related to exponential growth.

Exponential growth

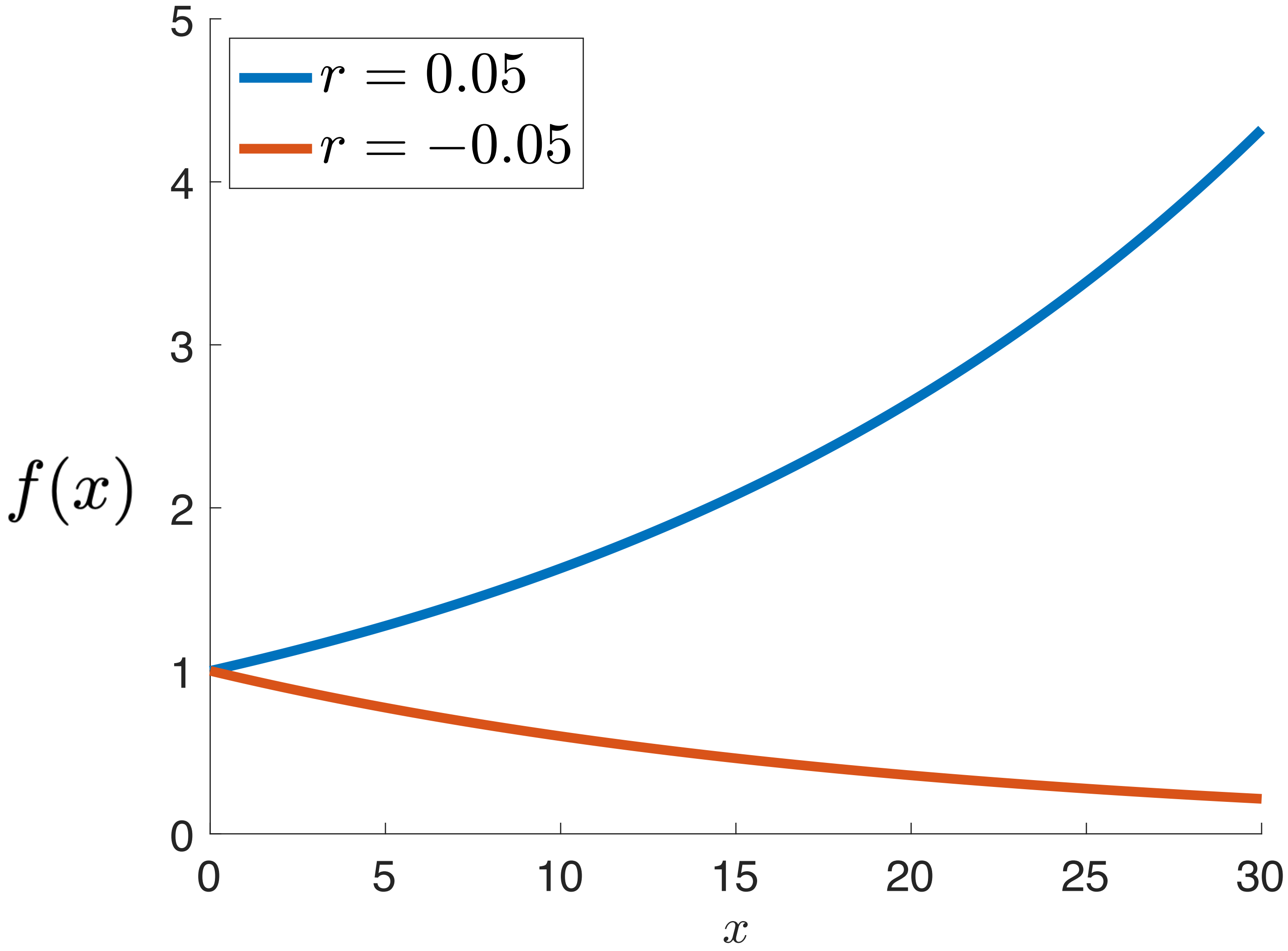
$$f(x) = a(1 + r)^x$$

$f(x)$: exponential growth function

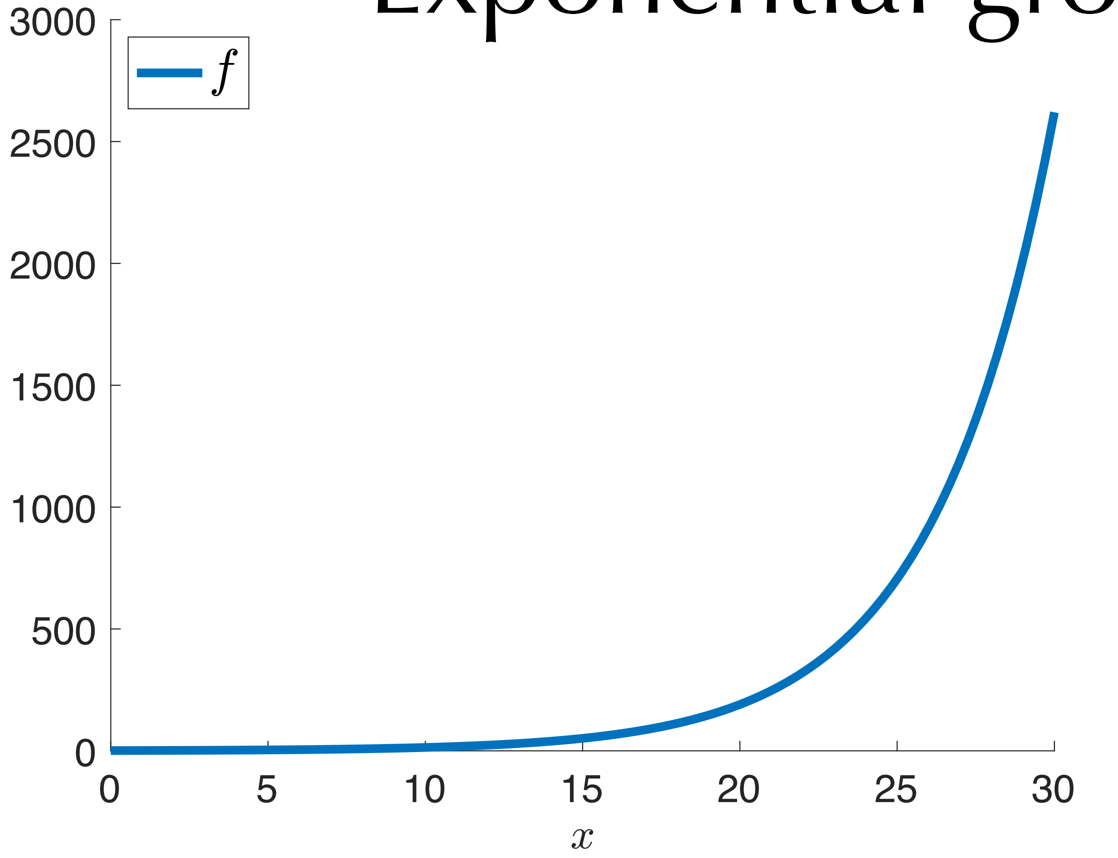
a : initial amount

r : growth rate

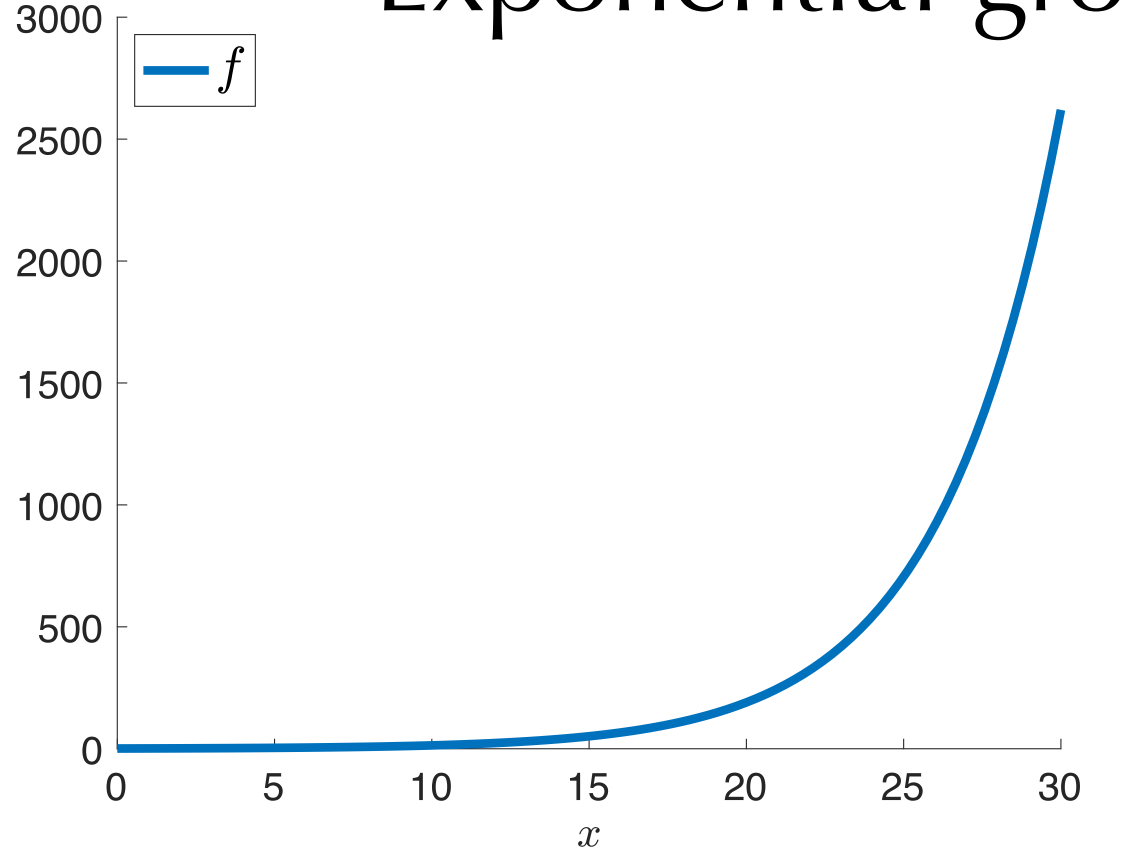
x : # of time intervals

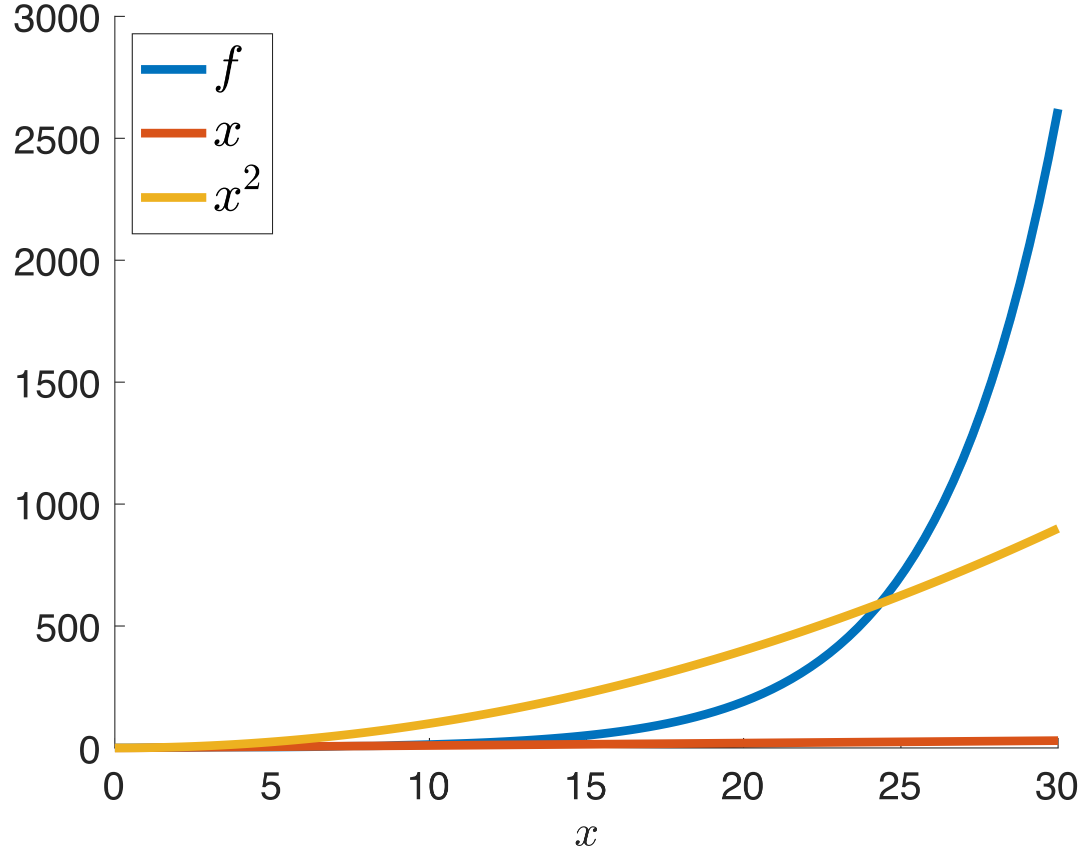


Exponential growth $f(x) = a(1 + r)^x$

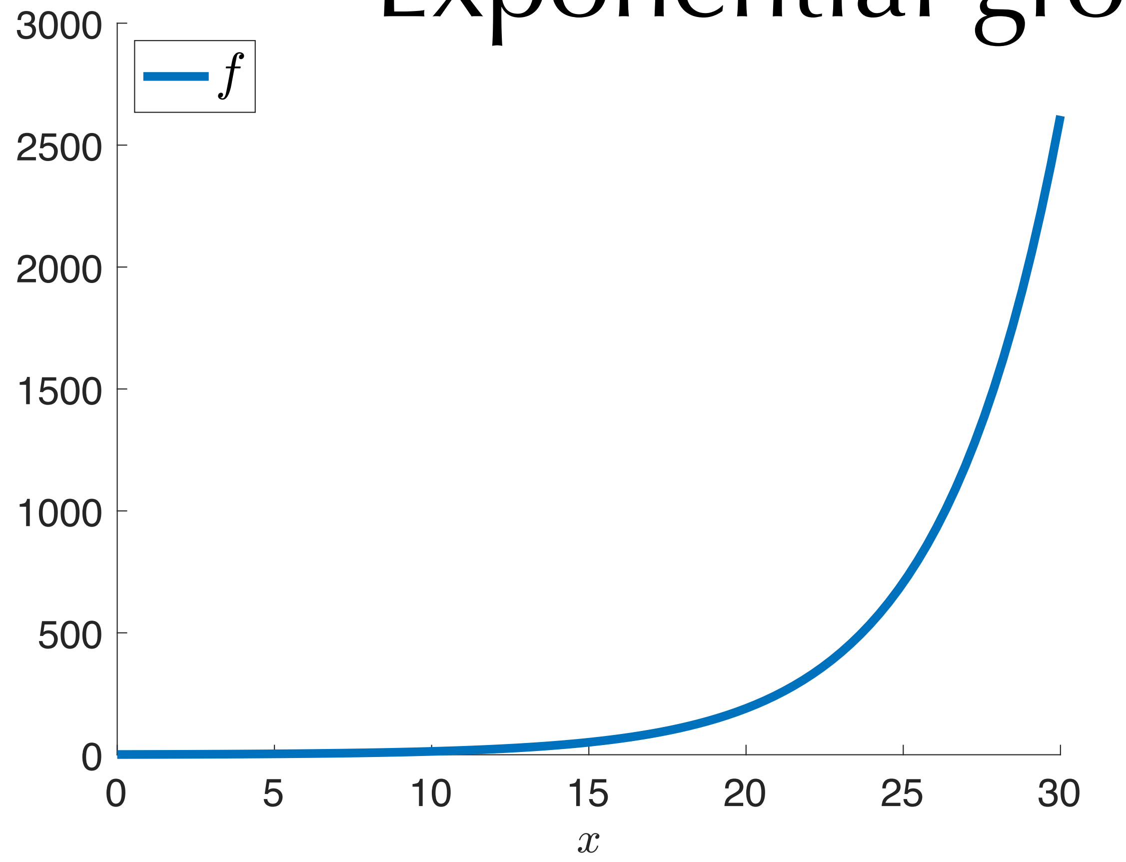


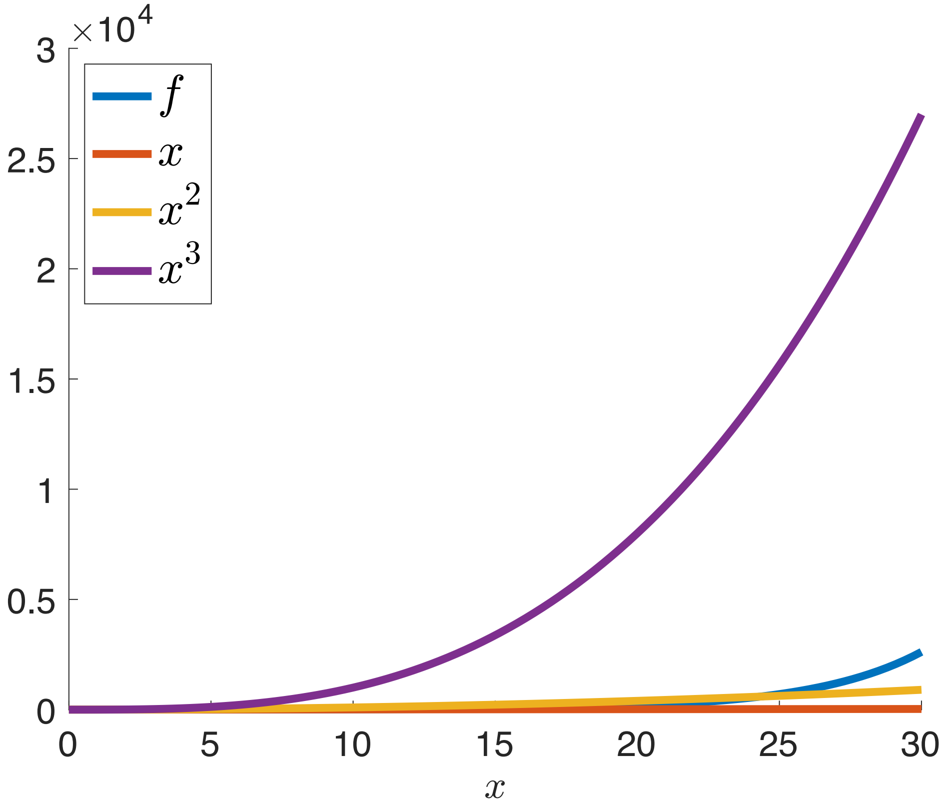
Exponential growth $f(x) = a(1+r)^x$

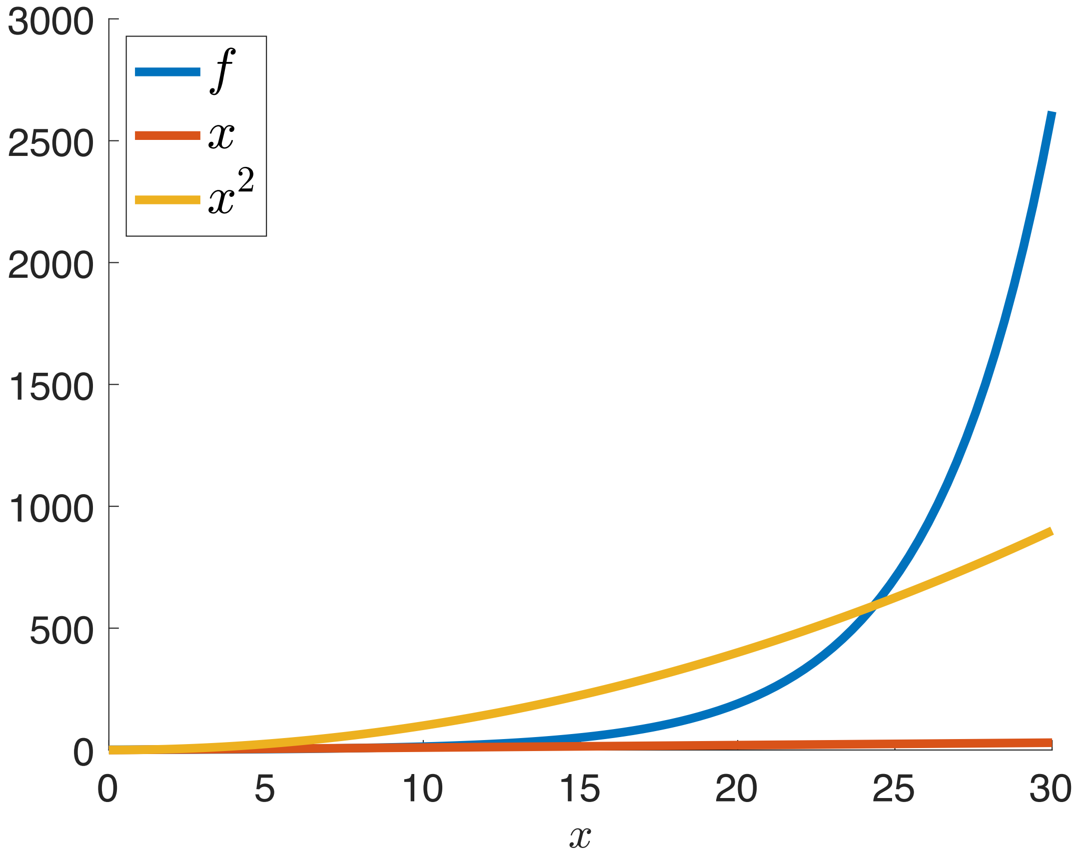




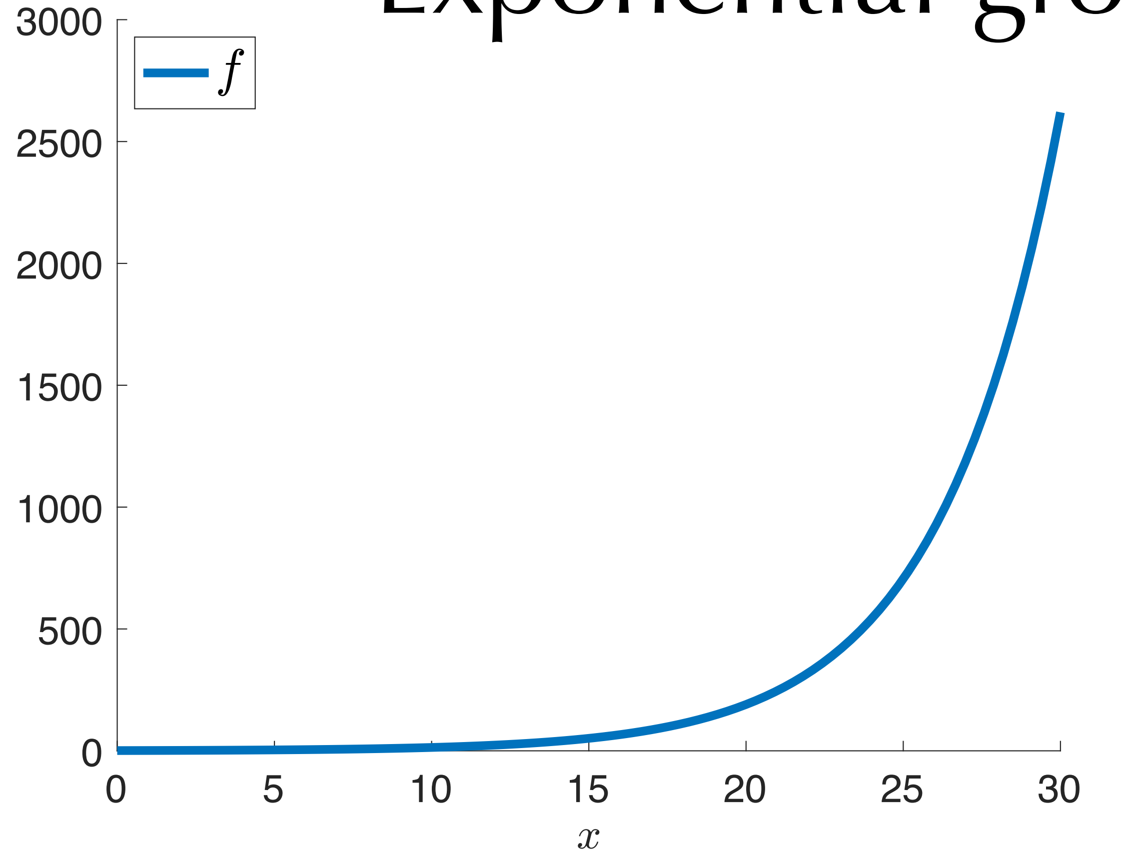
Exponential growth $f(x) = a(1+r)^x$

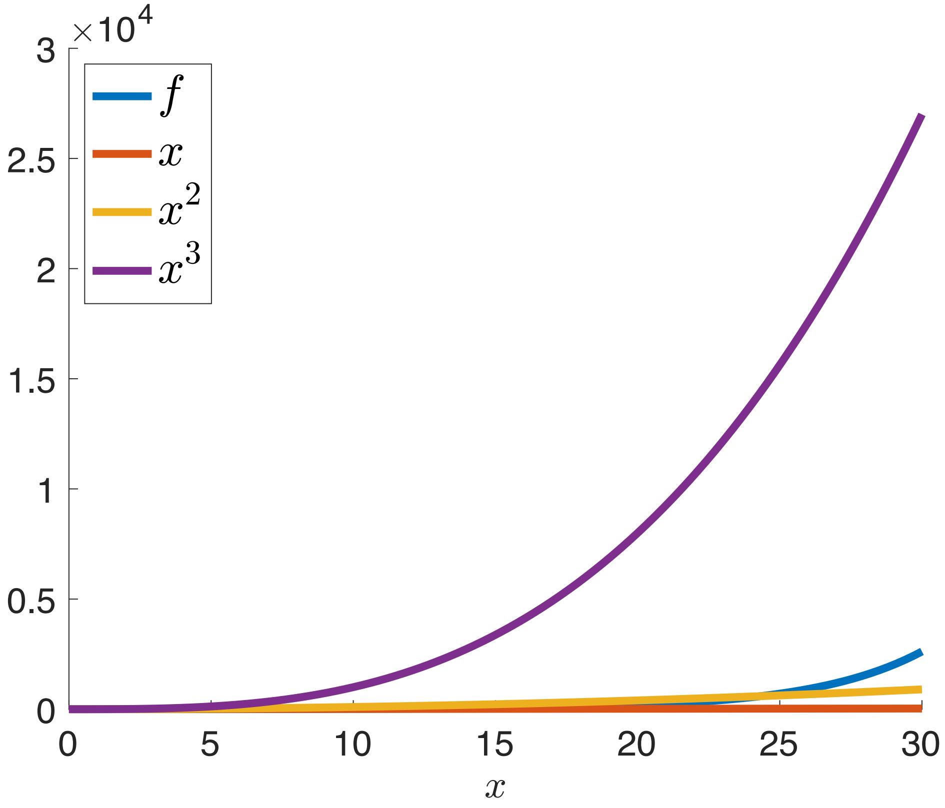


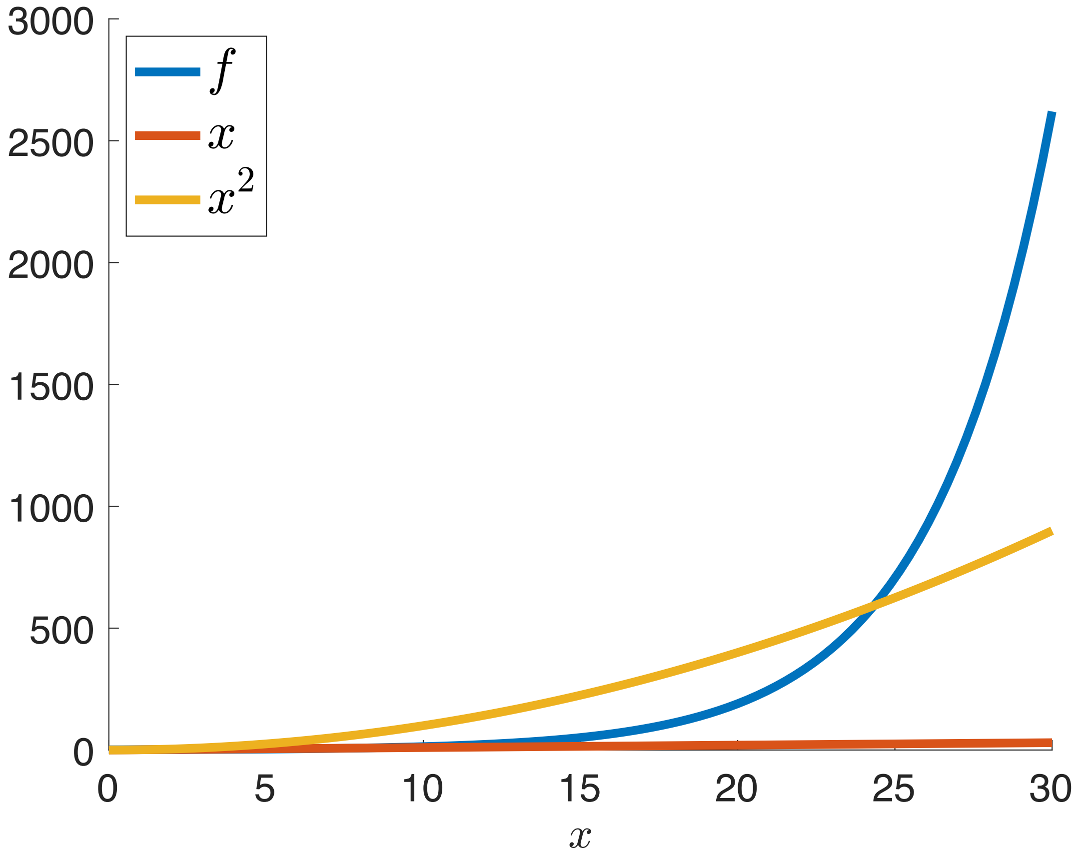


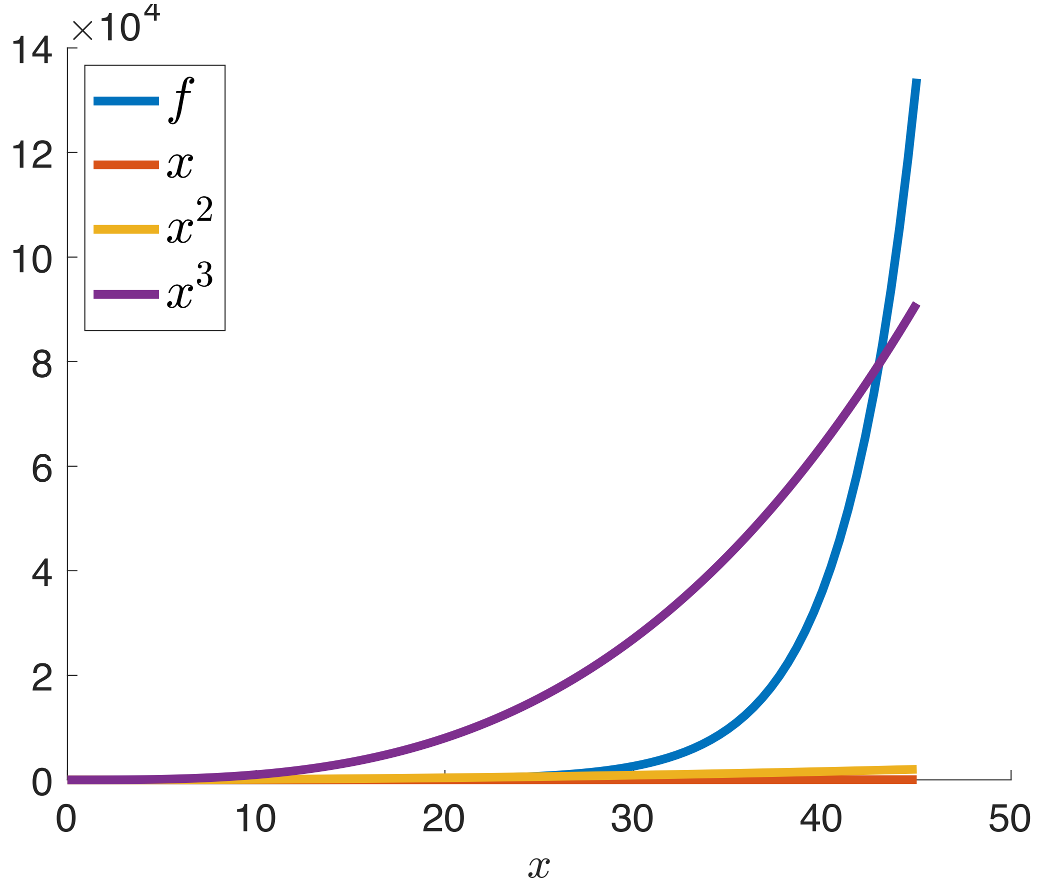


Exponential growth $f(x) = a(1+r)^x$

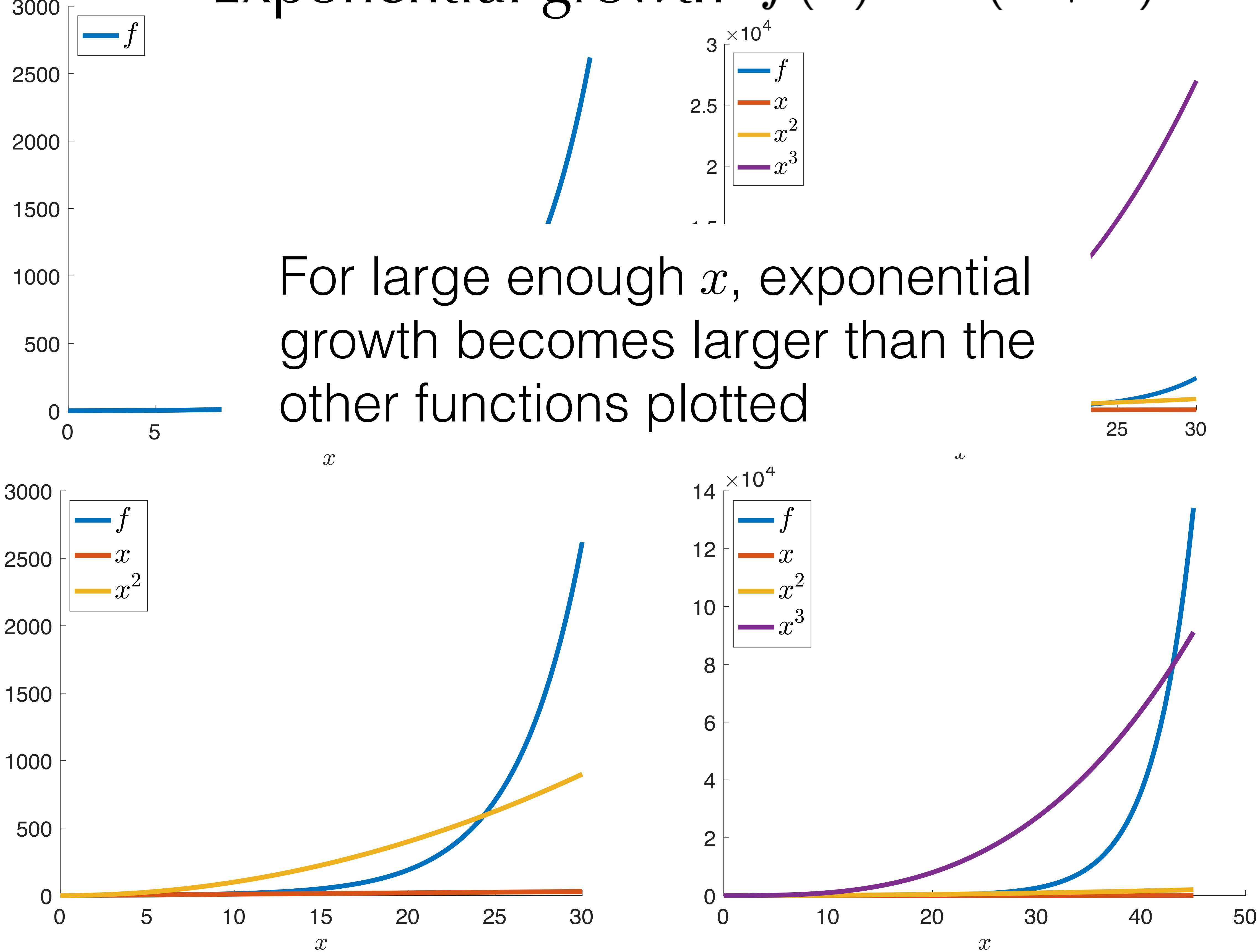








Exponential growth $f(x) = a(1+r)^x$



For large enough x, exponential growth becomes larger than the other functions plotted

$$f(x) = a(1+r)^x$$

Example 1:

Suppose that the initial balance of your savings account is $1,000. The savings account has a yearly compound interest rate of 1%. Without any new deposit, what is the balance of the account after 20 years?

$$f(20) = 1000(1+0.01)^{20} = 1220.2$$

Example 2:

Suppose that the initial balance of your savings account is $1,000. The savings account has a yearly compound interest rate of 1%. Without any new deposit, when will the balance reach $2,000?

$$
\begin{aligned}
f(x) = 1000(1+0.01)^x &= 2000 \\
1.01^x &= 2 \\
x \log 1.01 &= \log 2 \\
x &= \frac{\log 2}{\log 1.01} \approx 69.7
\end{aligned}
$$

Example 3:

There is an infectious disease, for which each infected person spreads the disease to exactly one healthy person every day. Suppose that there is only one infected patient on day 0 and the patients will not recover once they are infected. How many people will be infected after 32 days?

SI model

The SI model is the simplest form of all disease models. Individuals are born into the simulation with no immunity (susceptible). Once infected and with no treatment, individuals stay infected and infectious throughout their life, and remain in contact with the susceptible population.

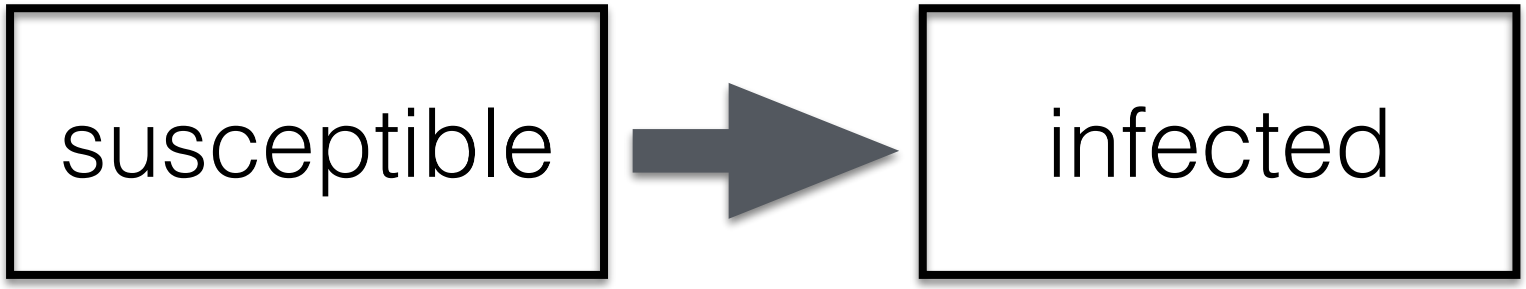
SI model

Each time an infected person meets a susceptible, there is a chance for the disease to spread to the non-infected. The rate of transmission (β) is defined as the average number of infection-producing contacts per unit time.

SI model

Each time an infected person meets a susceptible, there is a chance for the disease to spread to the non-infected. The rate of transmission (β) is defined as the average number of infection-producing contacts per unit time.

If there are N people in total, I infected and S susceptible, each susceptible has $\beta I/N$ infection-producing contacts per unit time. In total, there are therefore $\beta IS/N$ infections per unit time.

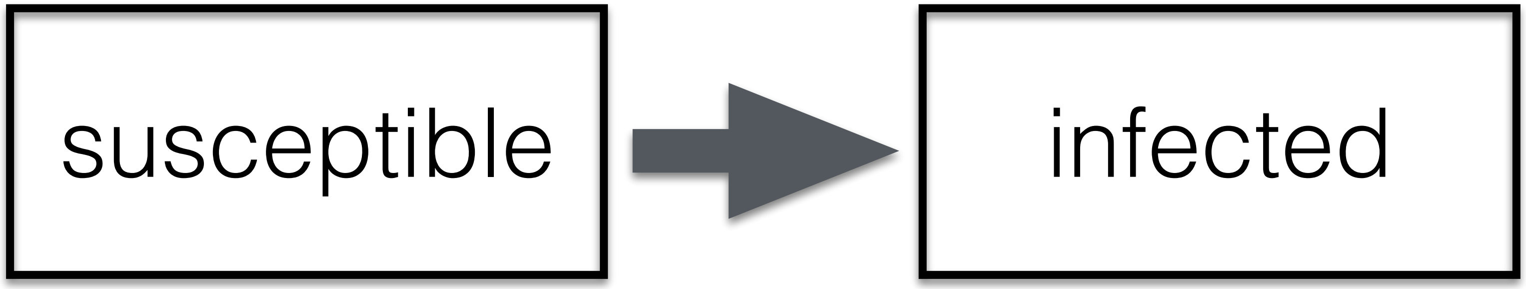


Susceptible
on day $t+1$

$$S_{t+1} = S_t - \beta \frac{S_t I_t}{N}$$

Susceptible
on day t

Newly infected



susceptible

infected

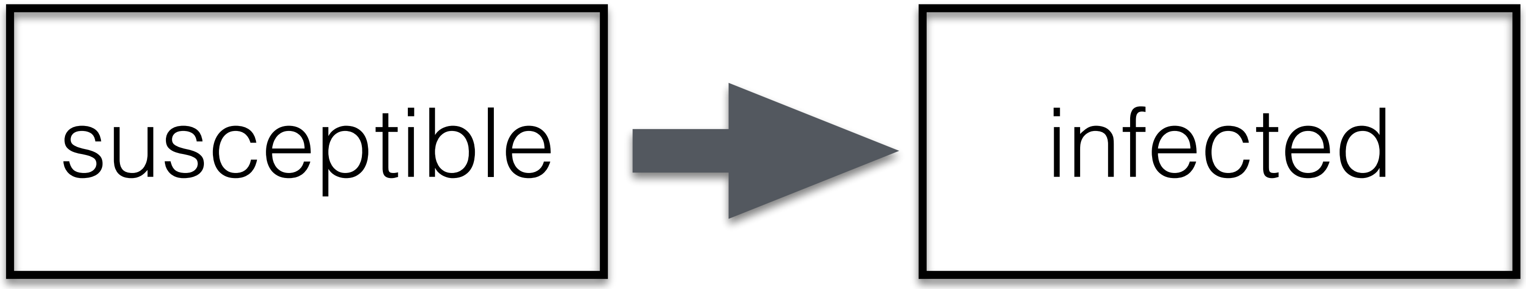
$$S_{t+1} = S_t - \beta\frac{S_t I_t}{N}$$

Newly infected

$$I_{t+1} = I_t + \beta\frac{S_t I_t}{N} = I_t + \beta I_t(1 - \frac{I_t}{N})$$

Infected
on day $t+1$

Infected
on day t



susceptible

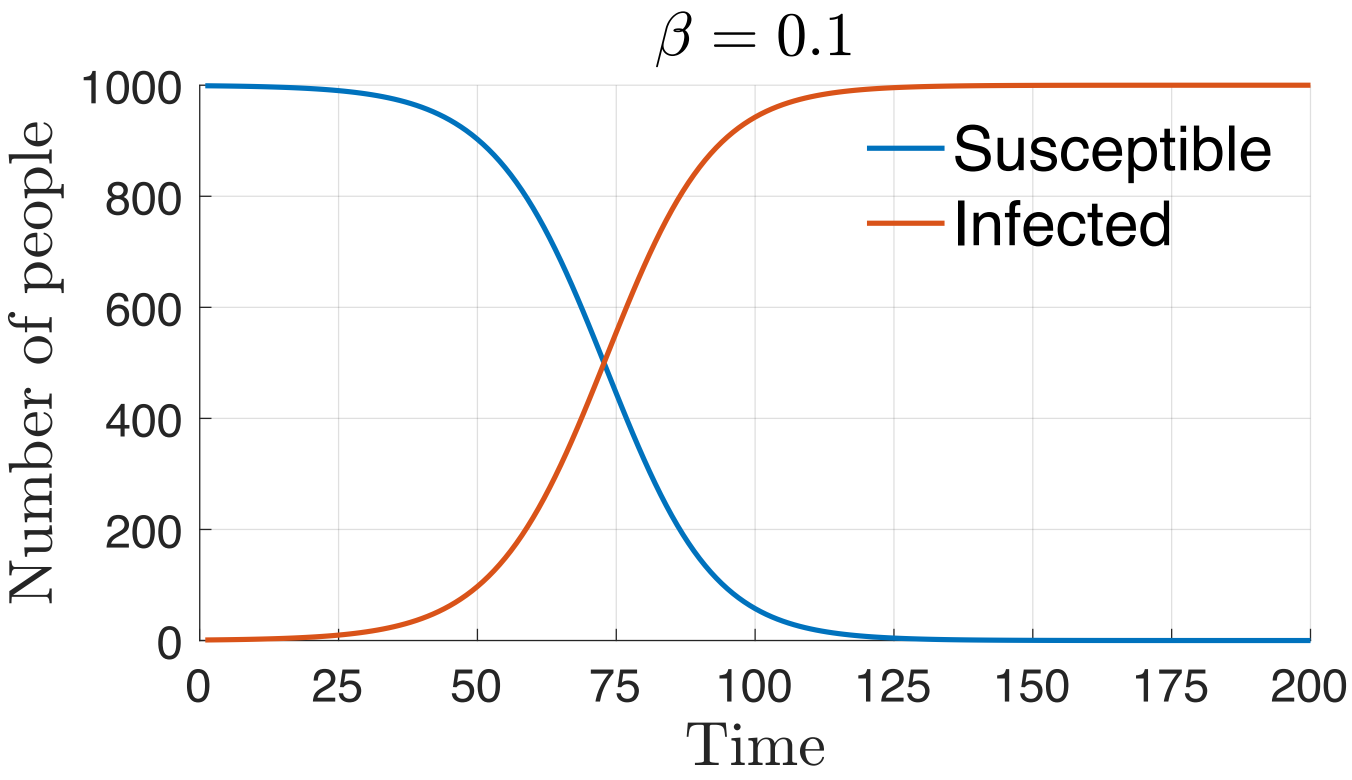
infected

$$S_{t+1} = S_t - \beta \frac{S_t I_t}{N}$$

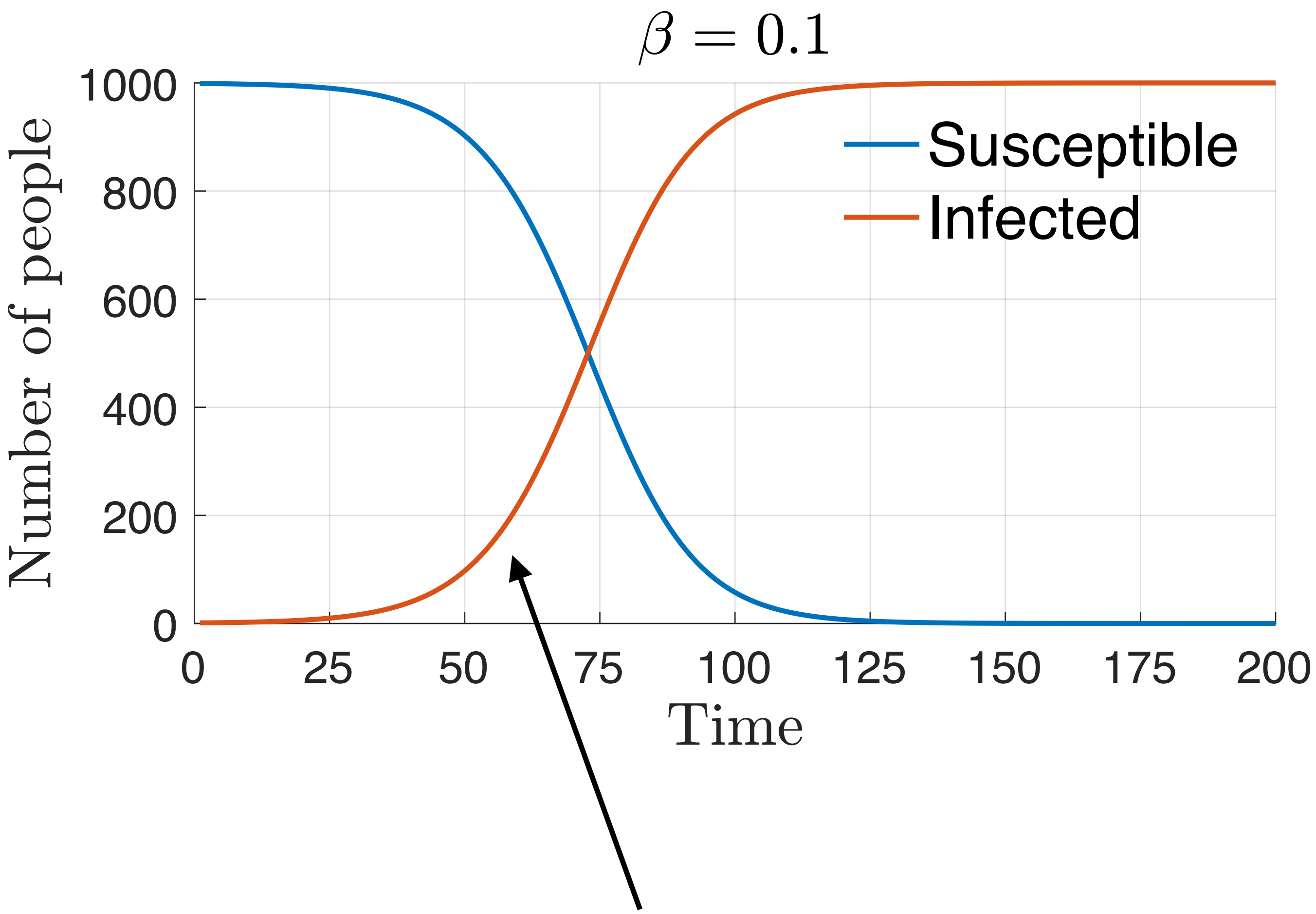
$$I_{t+1} = I_t + \beta \frac{S_t I_t}{N} = I_t + \beta I_t (1 - \frac{I_t}{N})$$

To compute I_t , plug in I_0, S_0 and update according to the second formula

Example: In a closed environment with population size N = 1,000, the transmission rate β is 0.1 per day, and there is only one infected person on day 1.

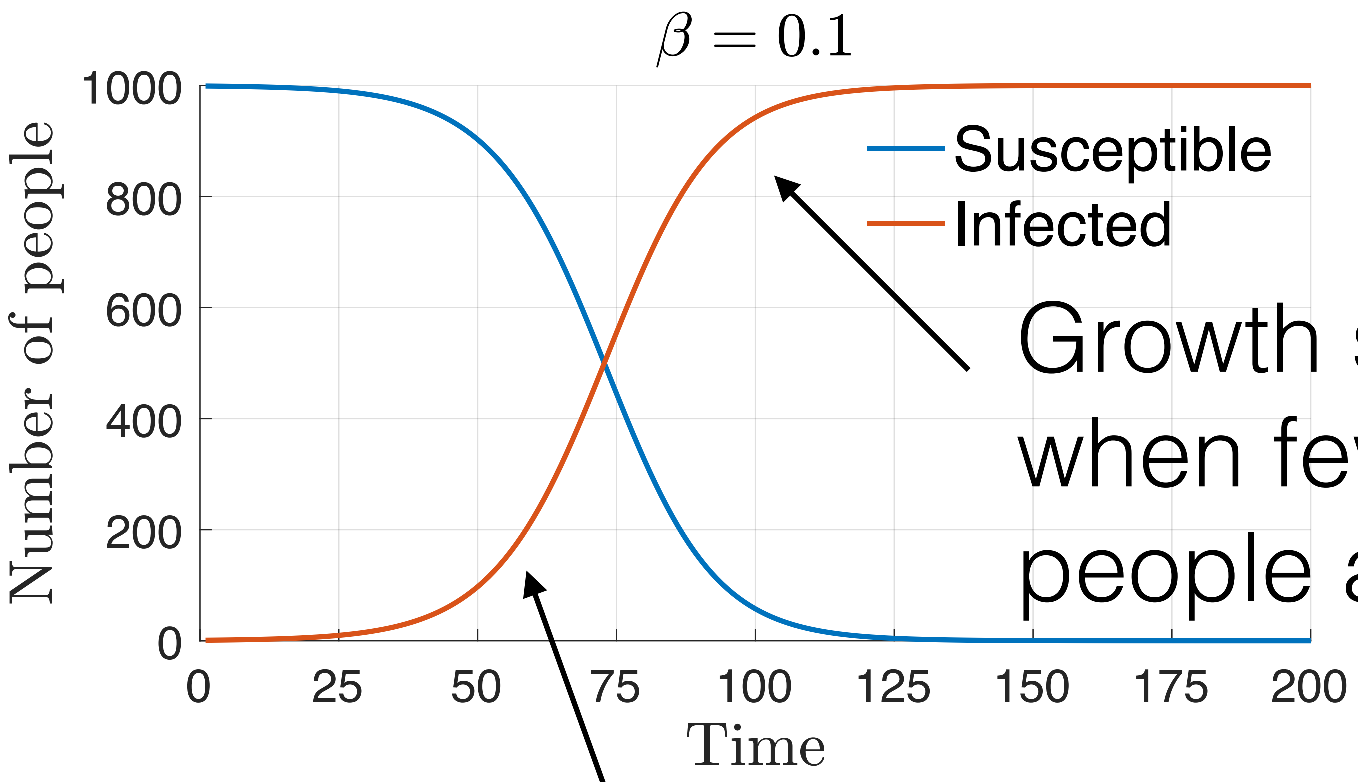


Example: In a closed environment with population size N = 1,000, the transmission rate β is 0.1 per day, and there is only one infected person on day 1.



Exponential growth at first, with rate β

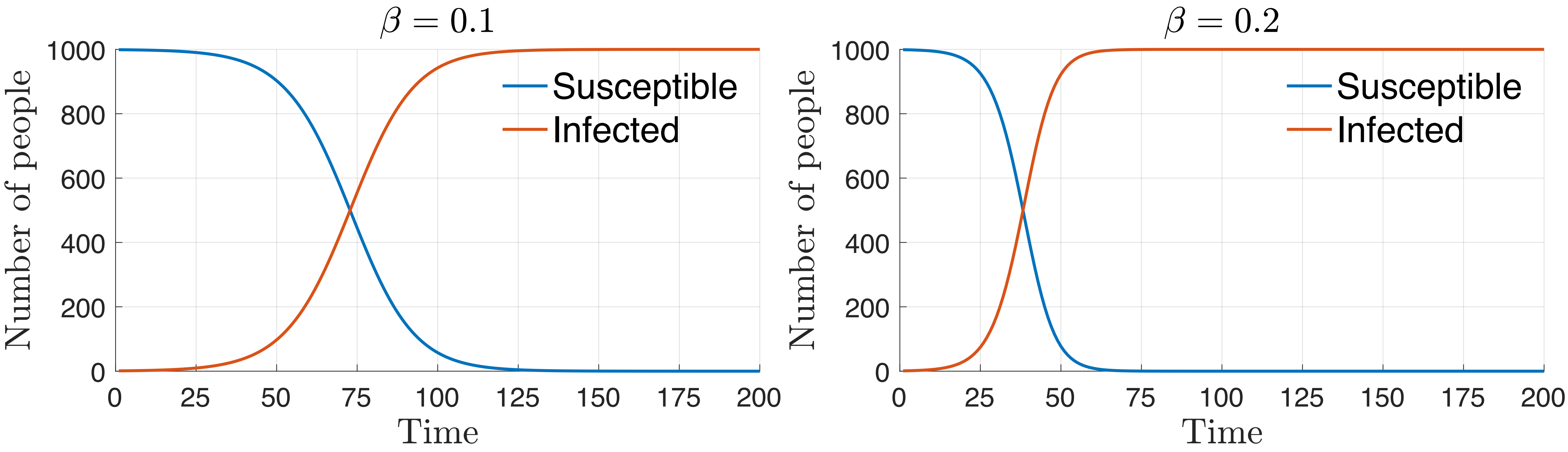
Example: In a closed environment with population size N = 1,000, the transmission rate β is 0.1 per day, and there is only one infected person on day 1.



Growth slows down when fewer susceptible people are left

Exponential growth at first, with rate β

Example: In a closed environment with population size N = 1,000, the transmission rate β is 0.1 per day, and there is only one infected person on day 1.



Critique of the SI model

Critique of the SI model

1. Does not allow people to recover, become immune or pass away
2. Does not take into account how people move (e.g. grocery stores and offices are hubs of people)
3. …

Models with recovery

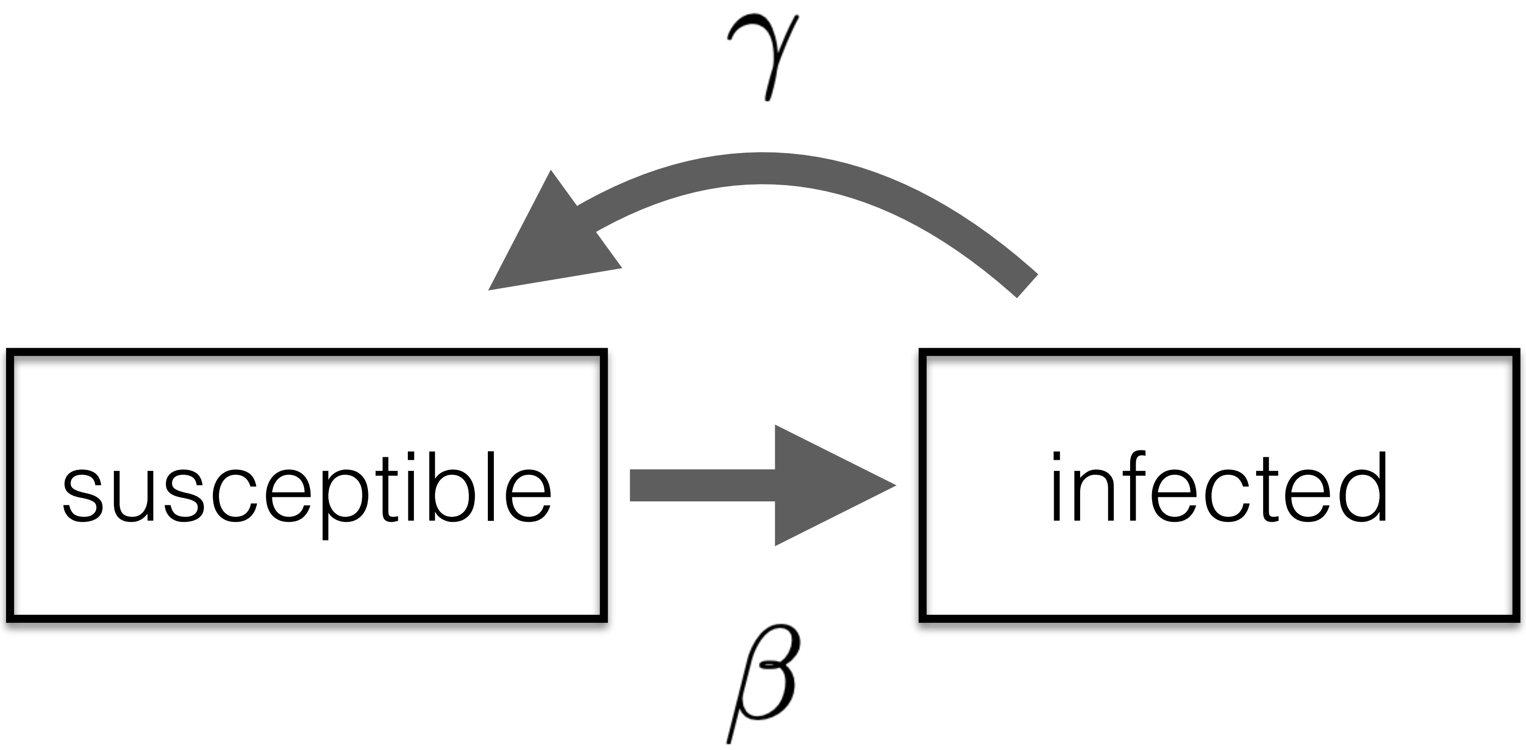
For other diseases, people recover and do not stay infected forever. This model is appropriate for diseases that commonly have repeated infections, for example, the common cold.

Rate of transmission (β): the average number of infection-producing contacts per unit time

Rate of recovery (γ): the inverse of the average duration of infection

SIS model

In the SIS model, the infected individuals can return to the susceptible state after infection. This model is appropriate for diseases that commonly have repeated infections, for example, the common cold.



γ

susceptible

infected

β

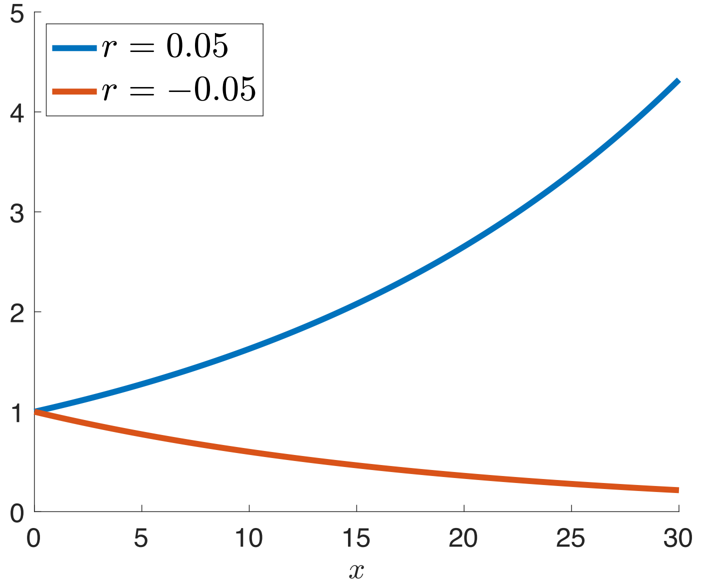
Basic reproduction number (R_0)

$$R_0 = \beta/\gamma$$

Rate of recovery (γ): the inverse of the duration of infection

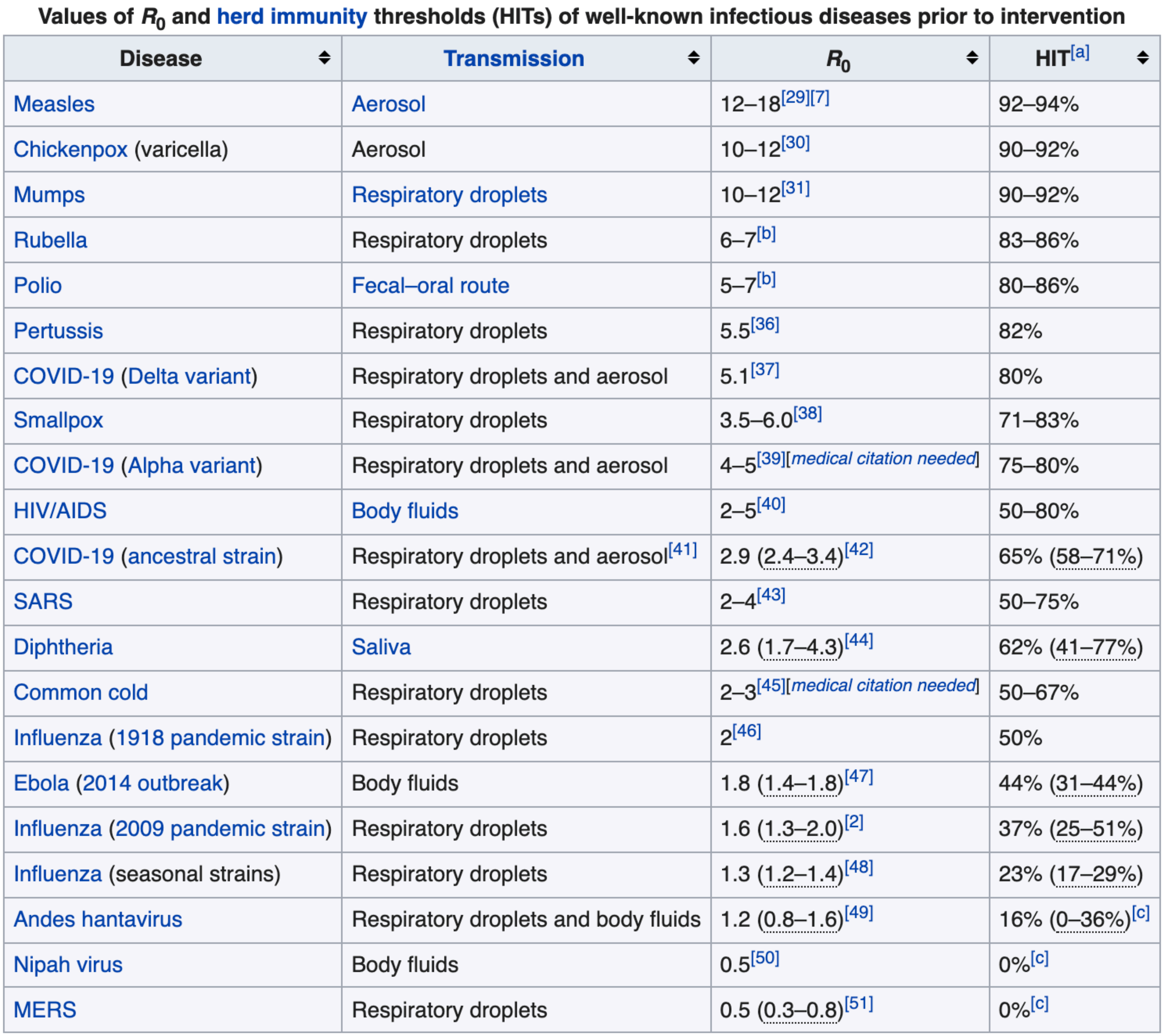
The rate for the exponential growth now becomes proportional to $\beta - \gamma$

$$f(x) = a(1+r)^x$$



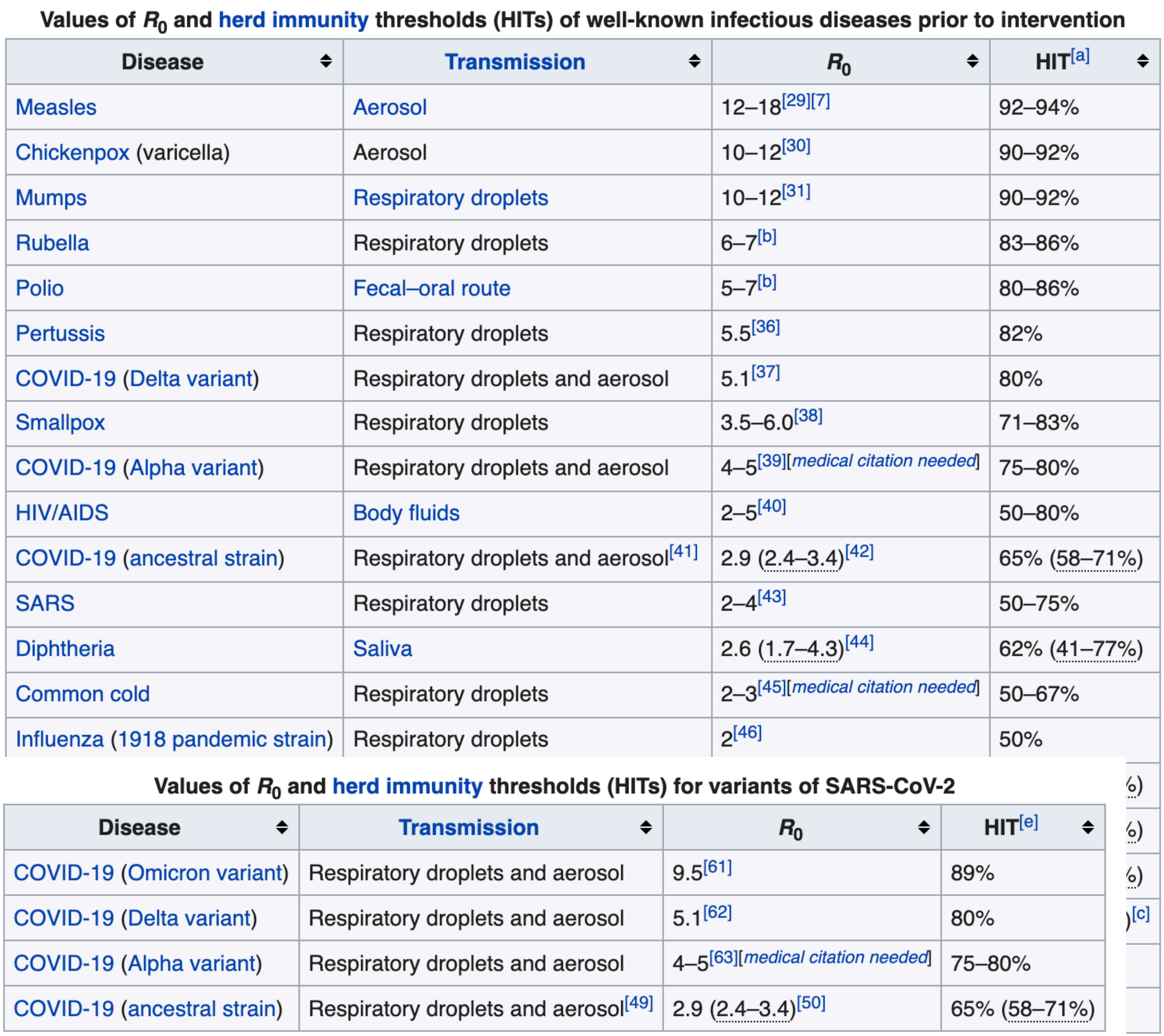
$R_0 < 1$: the disease will die out

$R_0 > 1$: the disease will spread



Values of R_0 and herd immunity thresholds (HITs) of well-known infectious diseases prior to intervention

Disease	Transmission	R_0	HIT[a]
Measles	Aerosol	12–18[29][7]	92–94%
Chickenpox (varicella)	Aerosol	10–12[30]	90–92%
Mumps	Respiratory droplets	10–12[31]	90–92%
Rubella	Respiratory droplets	6–7[b]	83–86%
Polio	Fecal–oral route	5–7[b]	80–86%
Pertussis	Respiratory droplets	5.5[36]	82%
COVID-19 (Delta variant)	Respiratory droplets and aerosol	5.1[37]	80%
Smallpox	Respiratory droplets	3.5–6.0[38]	71–83%
COVID-19 (Alpha variant)	Respiratory droplets and aerosol	4–5[39][*medical citation needed*]	75–80%
HIV/AIDS	Body fluids	2–5[40]	50–80%
COVID-19 (ancestral strain)	Respiratory droplets and aerosol[41]	2.9 (2.4–3.4)[42]	65% (58–71%)
SARS	Respiratory droplets	2–4[43]	50–75%
Diphtheria	Saliva	2.6 (1.7–4.3)[44]	62% (41–77%)
Common cold	Respiratory droplets	2–3[45][*medical citation needed*]	50–67%
Influenza (1918 pandemic strain)	Respiratory droplets	2[46]	50%
Ebola (2014 outbreak)	Body fluids	1.8 (1.4–1.8)[47]	44% (31–44%)
Influenza (2009 pandemic strain)	Respiratory droplets	1.6 (1.3–2.0)[2]	37% (25–51%)
Influenza (seasonal strains)	Respiratory droplets	1.3 (1.2–1.4)[48]	23% (17–29%)
Andes hantavirus	Respiratory droplets and body fluids	1.2 (0.8–1.6)[49]	16% (0–36%)[c]
Nipah virus	Body fluids	0.5[50]	0%[c]
MERS	Respiratory droplets	0.5 (0.3–0.8)[51]	0%[c]



Values of R_0 and herd immunity thresholds (HITs) of well-known infectious diseases prior to intervention

Disease	Transmission	R_0	HIT[a]
Measles	Aerosol	12–18[29][7]	92–94%
Chickenpox (varicella)	Aerosol	10–12[30]	90–92%
Mumps	Respiratory droplets	10–12[31]	90–92%
Rubella	Respiratory droplets	6–7[b]	83–86%
Polio	Fecal–oral route	5–7[b]	80–86%
Pertussis	Respiratory droplets	5.5[36]	82%
COVID-19 (Delta variant)	Respiratory droplets and aerosol	5.1[37]	80%
Smallpox	Respiratory droplets	3.5–6.0[38]	71–83%
COVID-19 (Alpha variant)	Respiratory droplets and aerosol	4–5[39][*medical citation needed*]	75–80%
HIV/AIDS	Body fluids	2–5[40]	50–80%
COVID-19 (ancestral strain)	Respiratory droplets and aerosol[41]	2.9 (2.4–3.4)[42]	65% (58–71%)
SARS	Respiratory droplets	2–4[43]	50–75%
Diphtheria	Saliva	2.6 (1.7–4.3)[44]	62% (41–77%)
Common cold	Respiratory droplets	2–3[45][*medical citation needed*]	50–67%
Influenza (1918 pandemic strain)	Respiratory droplets	2[46]	50%

Values of R_0 and herd immunity thresholds (HITs) for variants of SARS-CoV-2

Disease	Transmission	R_0	HIT[e]
COVID-19 (Omicron variant)	Respiratory droplets and aerosol	9.5[61]	89%
COVID-19 (Delta variant)	Respiratory droplets and aerosol	5.1[62]	80%
COVID-19 (Alpha variant)	Respiratory droplets and aerosol	4–5[63][*medical citation needed*]	75–80%
COVID-19 (ancestral strain)	Respiratory droplets and aerosol[49]	2.9 (2.4–3.4)[50]	65% (58–71%)

References

1. https://en.wikipedia.org/wiki/Basic_reproduction_number
2. https://docs.idmod.org/projects/emod-hiv/en/latest/model-si.html