Introduction to Stochastic Processes With R, by Robert P. Dobrow

Fernando Náufel

21/04/2022

Contents

Presentation

Notes about (Dobrow 2016).

CHAPTER 1

Introduction and review

1.1 Deterministic and Stochastic Models

Example 1.2 (SIR model)

- $S_t =$ number of susceptible people at time t.
- $I_t =$ number of newly infected people at time t.
- $R_t =$ number of recovered people at time t.
- $z =$ probability that a susceptible individual becomes infected once in contact with an infected person.
- Assume every susceptible person comes in contact with every infected person.

- Probability p that a susceptible individual becomes infected at a point in time:

$$p_t = 1 - (1 - z)^{I_{t-1}}$$

The book has the exponent as I_t, but the correct exponent is I_{t-1}.

In this discrete time model, we compute the number of newly infected people as a function of the number of people infected at the previous time step.

- This is because $(1-z)^{I_{t-1}}$ is the probability that the person has contact with all I_{t-1} infected people and does not become infected.
- The number of newly infected people I_t follows a binomial distribution with $n = S_{t-1}$ and probability of success p_t:

$$I_t \sim \text{Bin}(S_{t-1}, p_t)$$

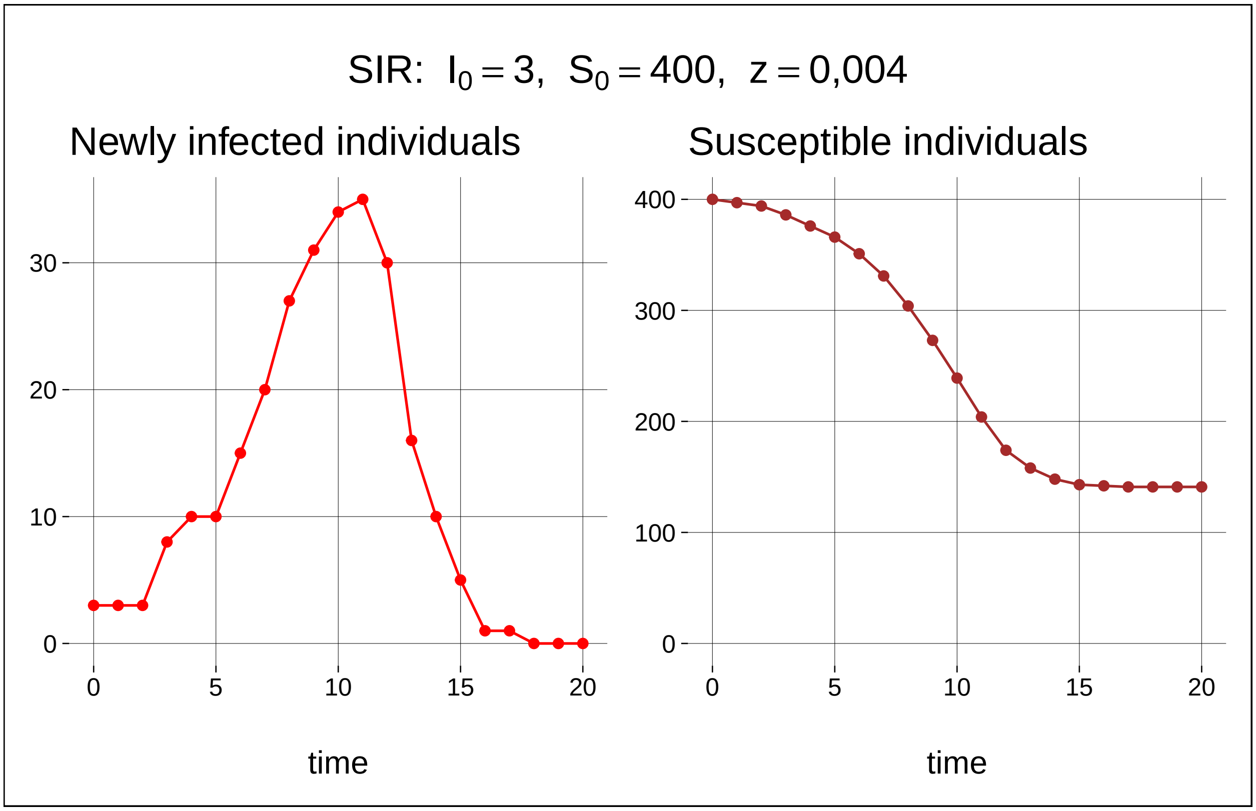
so the PMF is

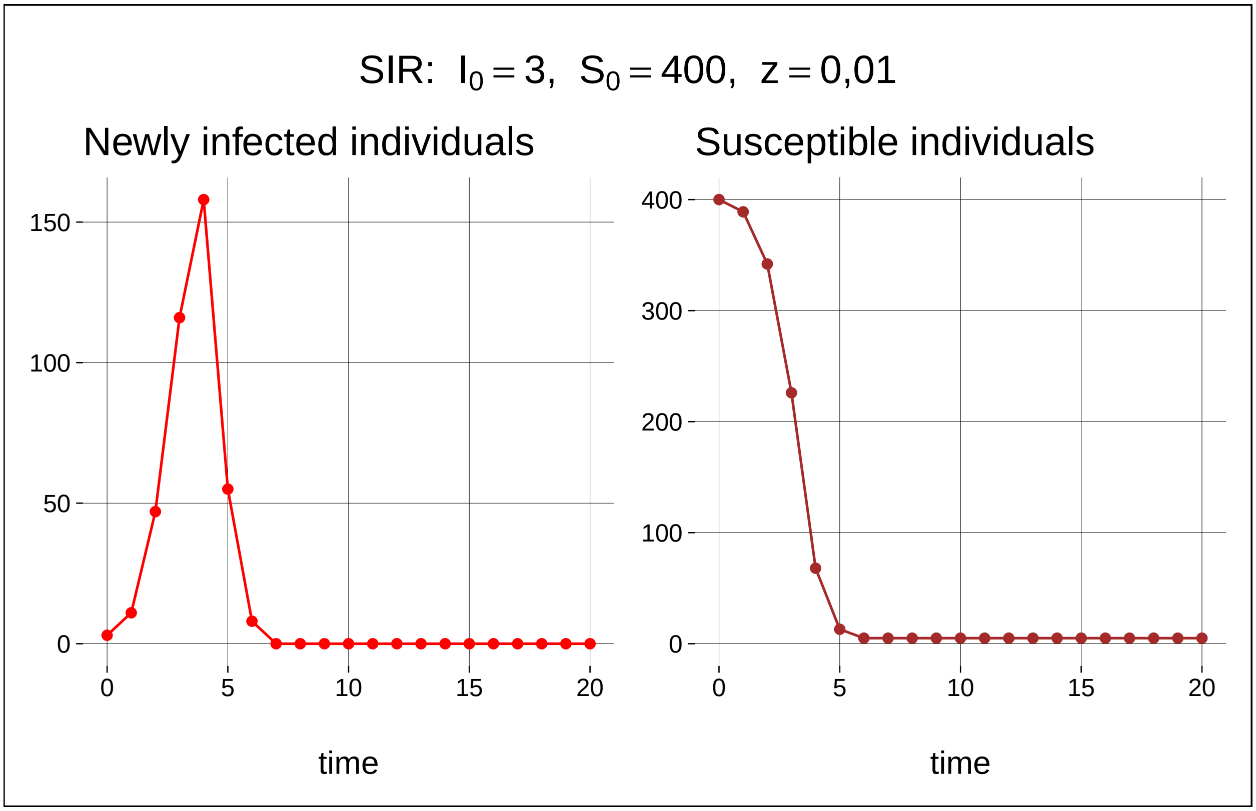
$$P(I_t = k) = \binom{S_{t-1}}{k} p_t^k \, (1-p_t)^{S_{t-1}-k}$$

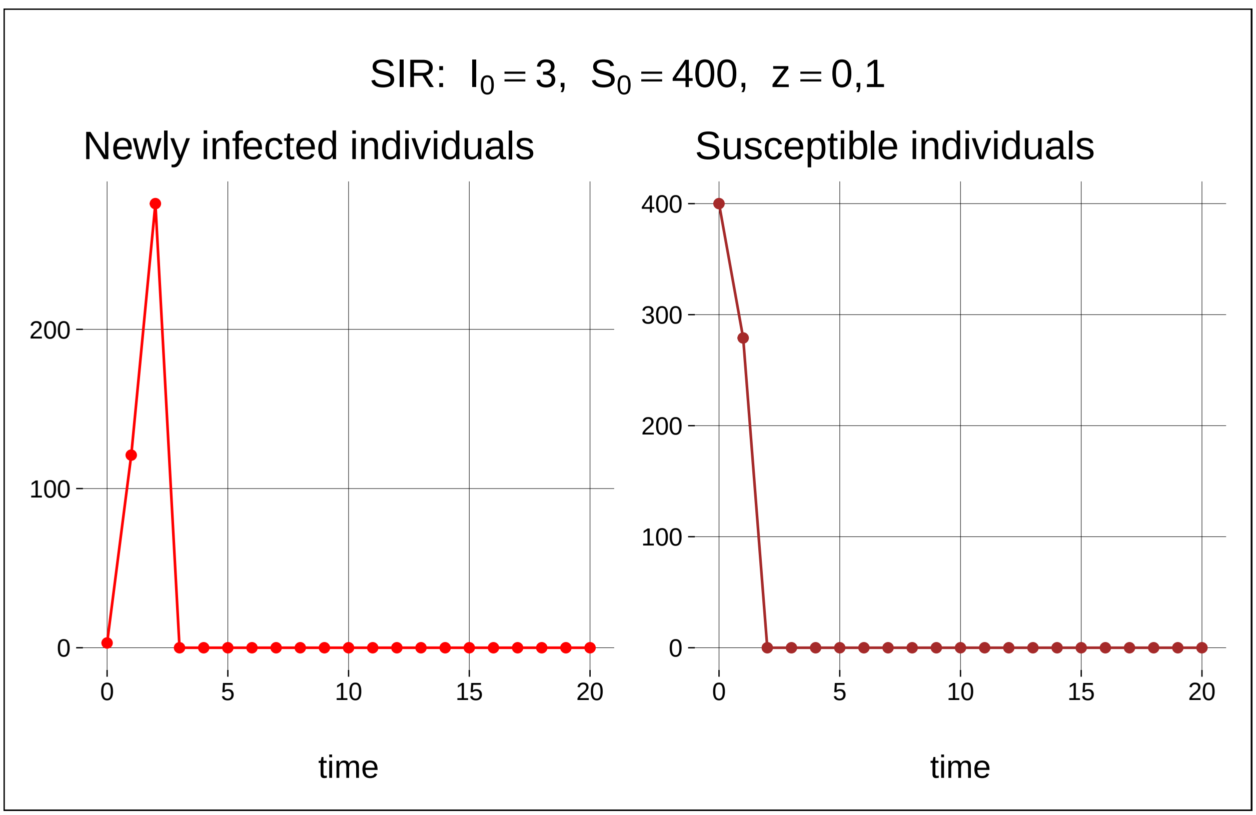
- Then the number of susceptible people at time t is

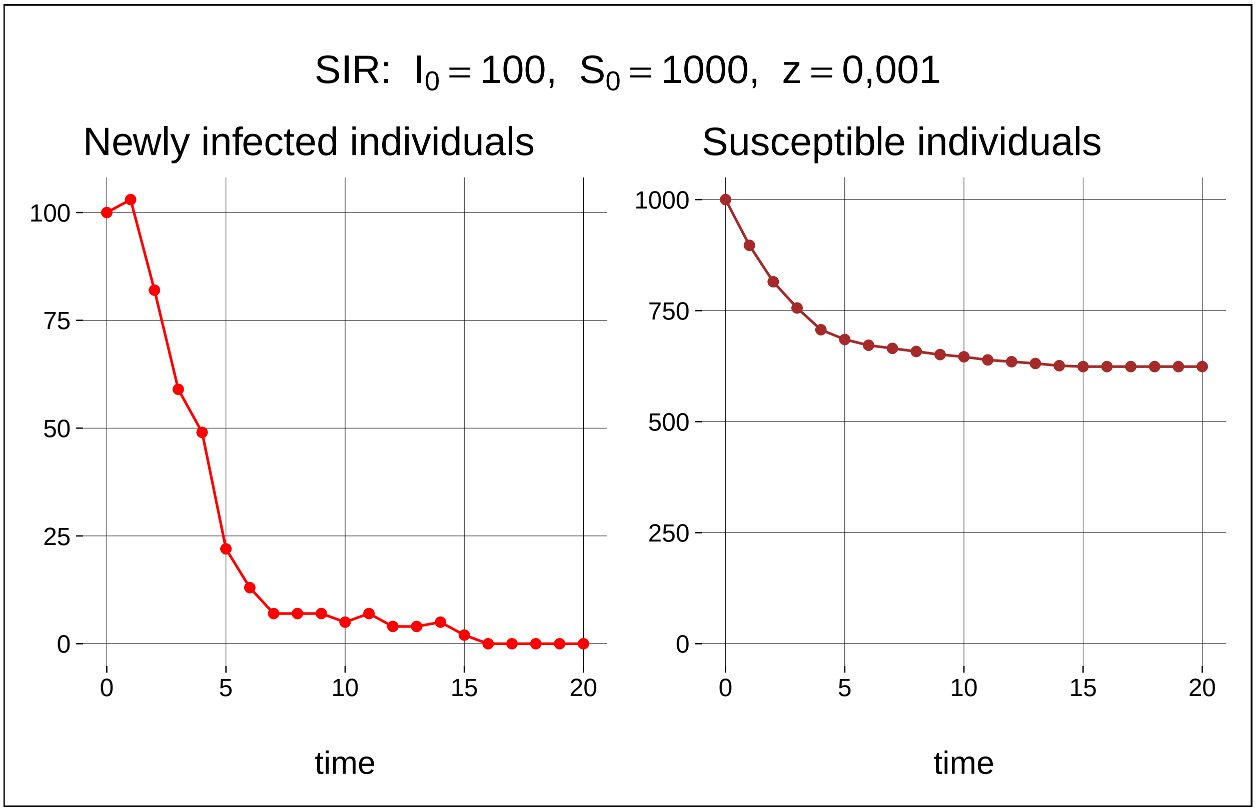
$$S_t = S_{t-1} - I_t$$

- This simplified example does not take into account the transition from infected to recovered.
- I_t is the number of *newly* infected people at time t.
- In the computation, only newly infected people are contagious. It's as if people remain infected and contagious for one time step only.


SIR: I0 = 3, S0 = 400, z = 0,004
Newly infected individuals
Susceptible individuals
0
10
20
30
0
100
200
300
400
0
5
10
15
20
time
time


SIR: $I_0 = 3$, $S_0 = 400$, $z = 0{,}01$
Newly infected individuals
Susceptible individuals
150
100
50
0
400
300
200
100
0
0 5 10 15 20
time
time


SIR: $I_0 = 3$, $S_0 = 400$, $z = 0{,}1$
Newly infected individuals
Susceptible individuals
200
100
0
400
300
200
100
0
0 5 10 15 20
time
time



1.2

What is a stochastic process?

Example 1.6 (Random walk and gambler's ruin)

...

Appendix B: Probability review

B.4 Common probability distributions

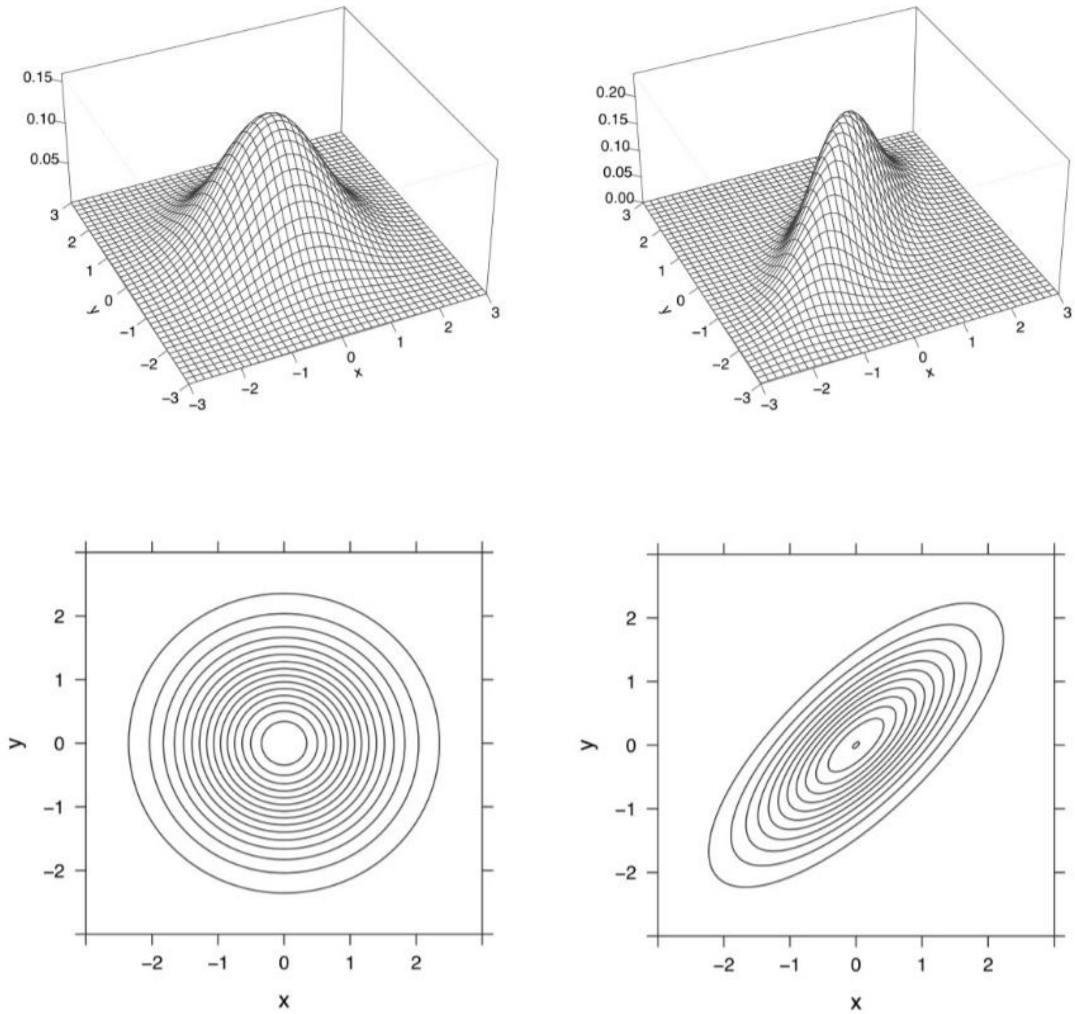
Bivariate normal

- The bivariate normal is defined here through its PDF — which is not given in its general form, but only in the case where $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(0,1)$:

$$f(x,y) = \frac{1}{2\pi\tau} \cdot \exp\left(-\frac{1}{2\tau^2} \cdot (x^2 - 2\rho xy + y^2)\right)$$

  with $\tau = \sqrt{1-\rho^2}$, where ρ is the correlation between X and Y.
- If the marginal distributions of X and Y are given, ρ is still free to vary in $[-1, 1]$.
- So, there are five parameters: μ_X, σ_X, μ_Y, σ_Y, and ρ.
- This figure from (Blitzstein and Hwang 2019) shows two bivariate normals with the same marginal distributions (the standard univariate normal) but different correlations ρ:

```
knitr::include_graphics('images/bvn.jpg')
```



- The question is

> Are both of these considered standard bivariate normals?

- From wikipedia[1]:

  > A real random vector $\mathrm{x} = (X_1, \ldots, X_k)^\intercal$ is called a standard normal random vector if all of its components X_k are independent and each is a zero-mean unit-variance normally distributed random variable, i.e. if $X_k \sim \mathcal{N}(0,1)$ for all k.

- For any RVs X and Y, independence implies $\rho = 0$. So, according to this definition, the standard bivariate normal has PDF

[1] Citing Lapidoth, Amos (2009). *A Foundation in Digital Communication.* Cambridge University Press. ISBN 978-0-521-19395-5.

$$f(x, y) = \frac{1}{2\pi} \cdot \text{exp} \left(-\frac{1}{2} \cdot (x^2 + y^2) \right)$$

and corresponds only to the graphs on the left.

- In the general case, for $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ and with $\rho \neq 0$, the PDF is

$$f(x, y) = \\ \frac{1}{2\pi\sigma_X\sigma_Y\tau} \cdot \text{exp} \left(-\frac{1}{2\tau^2} \left[\left(\frac{x - \mu_X}{\sigma_X} \right)^2 - 2\rho \left(\frac{x - \mu_X}{\sigma_X} \right) \left(\frac{y - \mu_Y}{\sigma_Y} \right) + \left(\frac{y - \mu_Y}{\sigma_Y} \right)^2 \right] \right) \tag{1.1}$$

with $\tau = \sqrt{1 - \rho^2}$ as before.

- The PDF of the conditional distribution of X given $Y = y$ is

$$f_{X|Y}(x, y) = \frac{f(x, y)}{f_Y(y)}$$

where f_Y is the marginal PDF

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y)\, dx \\ &= \frac{1}{\sigma_Y\sqrt{2\pi}} \cdot \text{exp} \left(-(y - \mu_Y)^2 / 2\sigma_Y^2 \right) \end{aligned}$$

yielding

$$\begin{aligned} f_{X|Y}(x, y) &= \frac{f(x, y)}{f_Y(y)} \\ &= \frac{\frac{1}{2\pi\sigma_X\sigma_Y\tau} \cdot \text{exp} \left(-\frac{1}{2\tau^2} \left[\left(\frac{x-\mu_X}{\sigma_X} \right)^2 - 2\rho \left(\frac{x-\mu_X}{\sigma_X} \right) \left(\frac{y-\mu_Y}{\sigma_Y} \right) + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 \right] \right)}{\frac{1}{\sigma_Y\sqrt{2\pi}} \cdot \text{exp} \left(-(y - \mu_Y)^2 / 2\sigma_Y^2 \right)} \end{aligned}$$

References

Blitzstein, Joseph K., and Jessica Hwang. 2019. *Introduction to Probability, Second Edition*. CRC Press.

Dobrow, Robert P. 2016. *Introduction to Stochastic Processes with R*. John Wiley & Sons.