

# Memoryless excursions

## Notes on joint distributions

Nicky van Foreest

December 28, 2020

### CONTENTS

1	A confusing problem	1
1.1	Discrete memoryless r.v.s	3
1.2	Continuous memoryless r.v.s	16
1.3	The solution	30

### 1 A CONFUSING PROBLEM

Write  $M = \max\{X, Y\}$  and  $L = \min\{X, Y\}$  for given r.v.s  $X$  and  $Y$ . It is simple to see that when  $X$  and  $Y$  have the same distribution

$$E[L] + E[M] = E[L + M] = E[X + Y] = 2E[X]. \quad (1.1)$$

From this, we have that

$$E[M] = 2E[X] - E[L]. \quad (1.2)$$

Next, when  $X$  and  $Y$  are memoryless and iid, it is tempting to think that

$$E[M] = E[L] + E[X], \quad (1.3)$$

because of the following interpretation. There are two machines each working on a job in parallel. Let  $X$  and  $Y$  be the production times at either machine. Then, we first wait until the first job finishes; this time is evidently  $L = \min\{X, Y\}$ . Then, due to memoryless, the service time of the remaining job ‘restarts’; this takes an expected time  $E[X]$  to complete.

Using (1.2) and (1.3) we can solve for  $E[M]$  and  $E[L]$ . Adding the two equations and canceling  $E[L]$  at both sides gives  $2E[M] = 3E[X]$ , hence:

$$E[M] = \frac{3}{2} E[X], \quad (1.4)$$

$$E[L] = E[M] - E[X] = \frac{1}{2} E[X]. \quad (1.5)$$

However, as we will see in the first set of exercises below, this is *not true for discrete* memoryless r.v.s, i.e., when  $X, Y \sim \text{Geo}(p)$ , while it is true when  $X, Y \sim \text{Exp}(\lambda)$ , cf., BW Ex.5.6.5.<sup>1</sup> In the exercises below I try to help you find out why this is so, and how to adapt (1.2) and/or (1.3) so that they also hold when  $X, Y \sim \text{Geo}(p)$ .

Actually, these exercises serve a few much more general goals. They are meant to

1. become familiar with all concepts of Section 7.1 of BH.
2. provide a recap of concepts and results of earlier chapters in the book,
3. motivate to read (and memorize) the material of BH A1, A.2.5, A.6, A.7, A.8, A.9 and A.10.

Once you solved all these exercises below, the exercises of BH will be considerably much simpler. All exercises have solutions, and focus on computational steps. I will provide motivation and explanations in the online lectures (on youtube).

I tend to use the following step-wise approach:

1. To compute probabilities of functions of r.v.s: use the fundamental bridge and the formula

$$\begin{aligned}
 P\{g(X, Y) \in A\} &= E[I_{g(X, Y) \in A}] \\
 &= \iint I_{g(x, y) \in A} f_{X, Y}(x, y) dx dy \\
 &= \iint_{g^{-1}(A)} f_{X, Y}(x, y) dx dy, \\
 &= \iint_{\{(x, y); g(x, y) \in A\}} f_{X, Y}(x, y) dx dy,
 \end{aligned} \tag{1.6}$$

or the discrete variant of this, and then do the integrals. Conceptually this approach is easy, but carrying out the integration is not always simple. However, the advantage is that I don't need special tricks to solve the problem; it is just (quite a bit of) straightforward work. With practice, this approach always works.

2. The joint CDF  $F_{X, Y}$  follows from step 1, because  $F_{X, Y}(x, y) = P\{X \leq x, Y \leq y\}$ .
3. To get the joint density  $f_{X, Y}$  say, I often first compute the joint distribution  $F_{X, Y}$  and then take partial derivatives. This approach is 'trick-free', but requires at times the hard work of solving the integrals to find  $F_{X, Y}$ .

In the exercises below we first try to see which of (1.2) and (1.3) is (and is not) true for memoryless discrete r.v.s, then we consider memoryless continuous r.v.s. Once we have dealt with these two cases, we can resolve the confusing problem.

---

<sup>1</sup> BH: Blitzstein and Hwang

### 1.1 Discrete memoryless r.v.s

Let  $X, Y$  be iid.  $\sim \text{Geo}(p)$ .

**Ex 1.1.** As a recap of earlier chapters of BH, show that

$$\mathbb{P}\{X > j\} = q^{j+1}, \quad \mathbb{P}\{X \geq j\} = q^j, \quad \mathbb{E}[X] = \frac{q}{1-q}. \quad (1.7)$$

What is the domain of  $X$ ?

**s.1.1.** Of course  $X \in \{0, 1, 2, \dots\}$ .

Here are a couple of tricks with recursions; this is how I rederive the properties of a geometric r.v.

$$\mathbb{P}\{X > 0\} = \mathbb{P}\{\text{failure}\} = q \quad (1.8)$$

$$\mathbb{P}\{X > j\} = q \mathbb{P}\{X > j-1\} \implies \mathbb{P}\{X > j\} = q^j \mathbb{P}\{X > 0\} = q^{j+1}. \quad (1.9)$$

$$\mathbb{P}\{X = j\} = \mathbb{P}\{X > j-1\} - \mathbb{P}\{X > j\} = q^j - q^{j+1} = (1-q)q^j = pq^j. \quad (1.10)$$

$$\mathbb{E}[X] = p \cdot 0 + q(1 + \mathbb{E}[X]) \implies \mathbb{E}[X] = q/(1-q) = q/p. \quad (1.11)$$

Now the same answers, but with the regular method.

$$\mathbb{P}\{X > j\} = \sum_{i=j+1}^{\infty} \mathbb{P}\{X = i\} \quad (1.12)$$

$$= p \sum_{i=j+1}^{\infty} q^i \quad (1.13)$$

$$= p \sum_{i=0}^{\infty} q^{j+1+i} \quad (1.14)$$

$$= pq^{j+1} \sum_{i=0}^{\infty} q^i \quad (1.15)$$

$$= pq^{j+1} \frac{1}{1-q} = pq^{j+1} \frac{1}{p} = q^{j+1}. \quad (1.16)$$

From this,

$$\mathbb{P}\{X \geq j\} = \mathbb{P}\{X > j-1\} = q^j, \quad (1.17)$$

$$(1.18)$$

The use of indicator variables is particularly important:

$$E[X] = \sum_{i=0}^{\infty} i P\{X = i\} \quad (1.19)$$

$$= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} I_{j < i} P\{X = i\} \quad (1.20)$$

$$= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} I_{j < i} P\{X = i\} \quad (1.21)$$

$$= \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} P\{X = i\} \quad (1.22)$$

$$= \sum_{j=0}^{\infty} P\{X > j\} \quad (1.23)$$

$$= \sum_{j=0}^{\infty} q^{j+1} \quad (1.24)$$

$$= q \sum_{j=0}^{\infty} q^j \quad (1.25)$$

$$= q/(1 - q) = q/p. \quad (1.26)$$

**Ex 1.2.** Check that  $X$ , hence  $Y$ , is memoryless.

**h.1.2.** What is the definition of memorylessness for discrete r.v.s?

**s.1.2.**

$$P\{X \geq n+m | X \geq m\} = \frac{P\{X \geq n+m, X \geq m\}}{P\{X \geq m\}} \quad (1.27)$$

$$= \frac{P\{X \geq n+m\}}{P\{X \geq m\}} \quad (1.28)$$

$$= \frac{q^{n+m}}{q^m} \quad (1.29)$$

$$= q^n = P\{X \geq n\}. \quad (1.30)$$

**Ex 1.3.** Show that  $P\{L \geq i\} = q^{2i}$  and conclude that  $L \sim \text{Geo}(1 - q^2)$ . Just to keep you sharp: what is the domain of  $L$ ?

**s.1.3.**

$$P\{L \geq k\} = \sum_{i,j} I_{\min\{i,j\} \geq k} P\{X = i, Y = j\} \quad (1.31)$$

$$= \sum_{i \geq k} \sum_{j \geq k} P\{X = i\} P\{Y = j\} \quad (1.32)$$

$$= P\{X \geq k\} P\{Y \geq k\} = q^k q^k = q^{2k}. \quad (1.33)$$

$P\{L > i\}$  has the same form as  $P\{X > i\}$ , but now with  $q^{2i}$  rather than  $q^i$ .

As henceforth all r.v.s we consider are non-negative, I'll drop the question to specify the domain of the r.v.s. However, in general you should not forget to specify this.

**Ex 1.4.** Show that  $E[L] = q^2/(1 - q^2)$ .

**h.1.4.**  $X \sim \text{Geo}(1 - q) \implies E[X] = q/(1 - q)$ . Now use that  $L \sim \text{Geo}(1 - q^2)$ .

**s.1.4.** Immediate from the hint and [\[1.3\]](#).

**Ex 1.5.** Show that

$$\mathbb{E}[L] + \mathbb{E}[X] = \frac{q}{1-q} \frac{1+2q}{1+q}. \quad (1.34)$$

**h.1.5.** Use that  $1 - q^2 = (1 - q)(1 + q)$ .

**s.1.5.**

$$\mathbb{E}[L] + \mathbb{E}[X] = \frac{q^2}{1-q^2} + \frac{q}{1-q} \quad (1.35)$$

$$= \frac{q}{1-q} \left( \frac{q}{1+q} + 1 \right) \quad (1.36)$$

$$(1.37)$$



**Ex 1.6.** Supposing that (1.2) is true, show that

$$E[M] = 2E[X] - EL = \frac{q}{1-q} \frac{2+q}{1+q}. \quad (1.38)$$

Show also that this is not the same as (1.34) (unless  $q = 0$ ). Conclude that this violates the simultaneous truth of (1.2) and (1.3).

**s.1.6.**

$$E[M] = 2E[X] - E[L] \quad (1.39)$$

$$= 2 \frac{q}{1-q} - \frac{q^2}{1-q^2} \quad (1.40)$$

$$= \frac{q}{1-q} \left( 2 - \frac{q}{1+q} \right) \quad (1.41)$$

$$= \frac{q}{1-q} \left( \frac{2+2q}{1+q} - \frac{q}{1+q} \right) \quad (1.42)$$

$$= \frac{q}{1-q} \frac{2+q}{1+q}. \quad (1.43)$$

So, something is wrong. I believe that (1.2) is correct, but then (1.3) must be wrong. But why? And, perhaps I am mistaken, and (1.3) is correct, and the other is not.

In fact, to convince myself that (1.3) is indeed wrong, I pursued three ideas to check (1.38), hence (1.2)

**Ex 1.7.** Idea 1: Show that for the PMF of  $M$ :

$$p_M(k) = \mathbb{P}\{M = k\} = 2pq^k(1 - q^k) + p^2q^{2k}. \quad (1.44)$$

Then use this to show (1.38).

**s.1.7.**

$$\mathbb{P}\{M = k\} = \mathbb{P}\{\max\{X, Y\} = k\} \quad (1.45)$$

$$= p^2 \sum_{ij} I_{\max\{i,j\}=k} q^i q^j \quad (1.46)$$

$$= 2p^2 \sum_{ij} I_{i=k} I_{j < k} q^i q^j + p^2 \sum_{ij} I_{i=j=k} q^i q^j \quad (1.47)$$

$$= 2p^2 q^k \sum_{j < k} q^j + p^2 q^{2k} \quad (1.48)$$

$$= 2p^2 q^k \frac{1 - q^k}{1 - q} + p^2 q^{2k} \quad (1.49)$$

$$= 2pq^k(1 - q^k) + p^2 q^{2k} \quad (1.50)$$

$$= 2pq^k + (p^2 - 2p)q^{2k} \quad (1.51)$$

$$= 2\mathbb{P}\{X = k\} - \mathbb{P}\{L = k\}, \quad (1.52)$$

where I use that  $p^2 - 2p = p(p - 2) = (1 - q)(1 - q - 2) = -(1 - q)(1 + q) = -(1 - q^2)$ .

$$\mathbb{E}[M] = \sum_k k \mathbb{P}\{M = k\} \quad (1.53)$$

$$= \sum_k k(2\mathbb{P}\{X = k\} - \mathbb{P}\{L = k\}) \quad (1.54)$$

$$= 2\mathbb{E}[X] - \mathbb{E}[L]. \quad (1.55)$$

**Ex 1.8.** Idea 2: Use  $P\{M \leq k\} = (P\{X \leq k\})^2$  to check (1.44).

**h.1.8.**  $P\{M = k\} = P\{M \leq k\} - P\{M \leq k - 1\}$ .

**s.1.8.**

$$P\{M \leq k\} = P\{X \leq k, Y \leq k\} \quad (1.56)$$

$$= P\{X \leq k\} P\{Y \leq k\} \quad (1.57)$$

$$= (1 - P\{X > k\})(1 - P\{Y > k\}) \quad (1.58)$$

$$= (1 - q^{k+1})^2. \quad (1.59)$$

$$P\{M = k\} = P\{M \leq k\} - P\{M \leq k - 1\} \quad (1.60)$$

$$= 1 - 2q^{k+1} + q^{2k+2} - (1 - 2q^k + q^{2k}) \quad (1.61)$$

$$= 2q^k(1 - q) + q^{2k}(q^2 - 1) \quad (1.62)$$

$$= 2P\{X = k\} - q^{2k}(1 - q^2) \quad (1.63)$$

$$= 2P\{X = k\} - P\{L = k\}. \quad (1.64)$$

**Ex 1.9.** Idea 3a: show that the joint PMF

$$p_{L,M}(i,j) = P\{L=i, M=j\} = 2p^2q^{i+j}I_{j>i} + p^2q^{2i}I_{i=j}. \quad (1.65)$$

**s.1.9.**

$$P\{L=i, M=j\} = 2P\{X=i, Y=j\}I_{j>i} + P\{X=Y=i\}I_{i=j} \quad (1.66)$$

$$= 2p^2q^{i+j}I_{j>i} + p^2q^{2i}I_{i=j}. \quad (1.67)$$

**Ex 1.10.** Idea 3b: Use [1.9] to compute the marginal PMFs of  $M$  and of  $L$ .

**h.1.10.** For the first, marginalize out  $L$ , for the second, marginalize out  $M$ .

**s.1.10.**

$$P\{M = j\} = \sum_i P\{L = i, M = j\} \quad (1.68)$$

$$= \sum_i (2p^2 q^{i+j} I_{j>i} + p^2 q^{2i} I_{i=j}) \quad (1.69)$$

$$= 2p^2 q^j \sum_{i=0}^{j-1} q^i + p^2 q^{2j} \quad (1.70)$$

$$= 2p q^j (1 - q^j) + p^2 q^{2j} \quad (1.71)$$

$$= 2p q^j + (p^2 - 2p) q^{2j}, \quad (1.72)$$

which is equal to (1.51).

$$P\{L = i\} = \sum_j P\{L = i, M = j\} \quad (1.73)$$

$$= \sum_j (2p^2 q^{i+j} I_{j>i} + p^2 q^{2i} I_{i=j}) \quad (1.74)$$

$$= 2p^2 q^i \sum_{j=i+1}^{\infty} q^j + p^2 q^{2i} \quad (1.75)$$

$$= 2p^2 q^{2i+1} \sum_{j=0}^{\infty} q^j + p^2 q^{2i} \quad (1.76)$$

$$= 2p q^{2i+1} + p^2 q^{2i} \quad (1.77)$$

$$= p q^{2i} (2q + p) \quad (1.78)$$

$$= (1 - q) q^{2i} (q + 1), \quad p + q = 1, \quad (1.79)$$

$$= (1 - q^2) q^{2i}, \quad (1.80)$$

and this we found earlier.

Finally, BH Ex.5.6.5 use that the r.v.  $M - L$  is independent of  $L$  when  $X, Y \sim \text{Exp}(\lambda)$  and independent. But this leads to our confusing problem, so perhaps my intuition about memorylessness is not ok. To check this, I solved the following exercise.

**Ex 1.11.** Show that the events  $\{L = i\}$  and  $\{M > L\}$  are independent when  $X, Y \sim \text{Geo}(p)$ .

**h.1.11.** Show that  $P\{L = i | M > L\} = P\{L = i\}$ ; recall, to show independence we need the probability (measure)  $P\{\cdot\}$ .

Use the expression  $P\{g(X, Y) \in A\} = \sum \sum_{g(i,j) \in A} p_{X,Y}(i, j)$

**s.1.11.**

$$P\{L = i | M > L\} = \frac{P\{L = i, M > L\}}{P\{M > L\}}. \quad (1.81)$$

$$P\{L = i, M > L\} = 2P\{X = i, Y > X\} \quad (1.82)$$

$$= 2P\{X = i\} P\{Y > i\} \quad (1.83)$$

$$= 2pq^i \cdot q^{i+1} \quad (1.84)$$

$$= 2pq^{2i+1}. \quad (1.85)$$

I use three (related) ways to compute  $P\{M > L\}$ . In the first I convert the event  $\{M > L\}$  directly into a statement about  $X$  and  $Y$ .

$$P\{M > L\} = P\{X > Y\} + P\{Y > X\} \quad (1.86)$$

$$= 2P\{X > Y\} \quad (1.87)$$

$$= 2 \sum_j P\{X > j, Y = j\} \quad (1.88)$$

$$= 2 \sum_j P\{X > j\} P\{Y = j\} \quad (1.89)$$

$$= 2p \sum_j q^{j+1} q^j \quad (1.90)$$

$$= 2p \sum_j q^{2j+1} \quad (1.91)$$

$$= 2pq \sum_j q^{2j} \quad (1.92)$$

$$= 2pq/(1 - q^2). \quad (1.93)$$

In the second, I rewrite the event  $\{M > L\}$  as a function of  $X$  and  $Y$ , and then use the formula of the hint:

$$P\{M > L\} = P\{\max\{X, Y\} > \min\{X, Y\}\} \quad (1.94)$$

$$= p^2 \sum_{i,j} I_{\max\{i,j\} > \min\{i,j\}} p_{XY}(i, j) \quad (1.95)$$

$$= p^2 \sum_{i,j} I_{\max\{i,j\} > \min\{i,j\}} p_X(i) p_Y(j) \quad (1.96)$$

$$= p^2 \sum_{i,j} I_{\max\{i,j\} > \min\{i,j\}} q^i q^j \quad (1.97)$$

$$= 2p^2 \sum_i \sum_{j>i} q^i q^j \quad (1.98)$$

$$= 2p^2 \sum_i q^i \sum_{j>i} q^j \quad (1.99)$$

$$= 2p^2 \sum_i q^i \sum_{j=0}^{\infty} q^{i+1+j} \quad (1.100)$$

$$= 2p^2 \sum_i q^{2i+1} \sum_{j=0}^{\infty} q^j \quad (1.101)$$

$$= 2p \sum_i q^{2i+1} \quad (1.102)$$

$$= 2pq \sum_i q^{2i} \quad (1.103)$$

$$= 2pq/(1 - q^2). \quad (1.104)$$

In the third, I use the probability  $P\{L = i, M > L\}$  and take the sum over  $i$ :

$$P\{M > L\} = \sum_i P\{L = i, M > L\} = \sum_i 2pq^{2i+1} = 2pq \sum_i 2pq^{2i} = \frac{2pq}{1 - q^2}. \quad (1.105)$$

Hence,

$$P\{L = i | M > L\} = \frac{P\{L = i, M > L\}}{P\{M > L\}} \quad (1.106)$$

$$= \frac{2pq^{2i+1}}{2pq/(1 - q^2)} \quad (1.107)$$

$$= q^{2i}(1 - q^2) \quad (1.108)$$

$$= P\{L = i\}, \quad (1.109)$$

where the last equation follows since  $L \sim \text{Geo}(1 - q^2)$ .

What have we achieved up to now? With respect to our problem, we are sure that (1.2) is correct when  $X, Y \sim \text{Geo}(p)$  and (1.3) is not. From the technical point of view, we practiced with joint, marginal, and conditional PMFs, and we have seen how to use indicator functions and the fundamental bridge.

So, this made me wonder why is (1.3) incorrect when  $X, Y \sim \text{Geo}(p)$ , but it is correct when  $X, Y \sim \text{Exp}(\lambda)$ . To understand this discrepancy better, I decided to analyze the latter case in as much detail as I could think of, hoping that this would provide me with a lead. The work will be pretty technical at times, but there is no way around when dealing with joint densities, etc.

## 1.2 *Continuous memoryless r.v.s*

In this section we analyze the correctness of (1.2) and (1.3) for continuous memoryless r.v.s, i.e., exponentially distributed r.v.s. This will offer a good opportunity to practice with several concepts: joint, marginal and conditional distributions of continuous r.v.s, limits, and moment-generating functions. We don't only solve the problem of Section 1, in passing we will obtain a lot of necessary and useful practice.

First we need to recall some basic facts about the exponential distribution.

**Ex 1.12.** Show that  $X$ , hence  $Y$ , is memoryless.

**s.1.12.** See BH.



**Ex 1.13.** Show that  $E[X] = 1/\lambda$ .

**h.1.13.** Use partial integration.

**s.1.13.** It is essential that you know all elements of this solution by heart.

$$E[X] = \int_0^{\infty} x f_X(x) dx \quad (1.110)$$

$$= \int_0^{\infty} x \lambda e^{-\lambda x} dx \quad (1.111)$$

$$= -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \quad (1.112)$$

$$= -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}. \quad (1.113)$$

Substitution is also a very important technique to solve such integrals. Here we go again:

$$E[X] = \int_0^{\infty} x f_X(x) dx \quad (1.114)$$

$$= \int_0^{\infty} x \lambda e^{-\lambda x} dx \quad (1.115)$$

$$= \frac{1}{\lambda} \int_0^{\infty} u e^{-u} du, \quad (1.116)$$

by the substitution  $u = \lambda x \Rightarrow du = d(\lambda x) \Rightarrow du = \lambda dx \Rightarrow dx = du/\lambda$ . With partial integration (do it!), the integral evaluates to 1.

Now we can shift our attention to the r.v.s  $L$  and  $M$ .

**Ex 1.14.** Show that the density  $f_L(x)$  of  $L$  is equal to  $2\lambda e^{-\lambda x}$ . Use this to see that  $L \sim \text{Exp}(2\lambda)$  and that  $E[L] = 1/(2\lambda)$ . Realize that with this we have an independent check of (1.5).

**h.1.14.** First compute the distribution  $F_L(x)$  of  $L$ .

**s.1.14.**

$$P\{L > x\} = P\{X > x, Y > x\} = (P\{X > x\})^2 = (1 - P\{X \leq x\})^2 = (e^{-\lambda x})^2 = e^{-2\lambda x}. \quad (1.117)$$

Hence,  $F_L(x) = 1 - e^{-2\lambda x}$ , hence  $L \sim \text{Exp}(2\lambda)$ . But then,  $f_L(x) = F'_L(x) = 2\lambda e^{-2\lambda x}$ .

With infinitesimals we can get the same result, but mind, this technique requires some serious thought:

$$f_L(x)dx = P\{L \in [x, x + dx]\} \quad (1.118)$$

$$= P\{X \in [x, x + dx], Y > x\} + P\{Y \in [x, x + dx], X > x\} \quad (1.119)$$

$$= f_X(x)(1 - F_Y(x)) + f_Y(x)(1 - F_X(x))dx \quad (1.120)$$

$$= 2f_X(x)(1 - F_Y(x))dx \quad (1.121)$$

$$= 2\lambda e^{-\lambda x} e^{-\lambda x} dx = 2\lambda e^{-2\lambda x} dx. \quad (1.122)$$

Thus,

$$E[L] = \int_0^\infty x f_L(x) dx \quad (1.123)$$

$$= \int_0^\infty x 2\lambda e^{-2\lambda x} dx. \quad (1.124)$$

Substitution of  $u = 2\lambda x$  and using the solution of [1.13] gives the answer.

**Ex 1.15.** Compute  $F_M(x) = P\{M \leq x\}$ . With this, show that  $f_M(x) = 2(1 - e^{-\lambda x})\lambda e^{-\lambda x}$ , and then use this to compute  $E[M]$  to show that (1.4) holds.

**s.1.15.** Here are two related ways. This is one:

$$P\{M \leq x\} = P\{X \leq x, Y \leq x\} = (F_X(x))^2, \quad (1.125)$$

since  $X, Y$  iid.

The second is based on (1.6):

$$P\{M \leq u\} = E[I_{M \leq u}] \quad (1.126)$$

$$= \int_0^\infty \int_0^\infty I_{x \leq u, y \leq u} f_{XY}(x, y) dx dy \quad (1.127)$$

$$= \int_0^\infty \int_0^\infty I_{x \leq u, y \leq u} f_X(x) f_Y(y) dx dy \quad (1.128)$$

$$= \int_0^u f_X(x) dx \int_0^u f_Y(y) dy \quad (1.129)$$

$$= F_X(u) F_Y(u) = (F_X(u))^2. \quad (1.130)$$

We get the same result.

Now the density:

$$f_M(u) = F'_M(u) = 2F_X(u)f_X(u) = 2(1 - e^{-\lambda u})\lambda e^{-\lambda u}. \quad (1.131)$$

Here is again my favorite method to get the density

$$f_M(u) du = P\{M \in [u, u + du]\} = 2P\{X \in [u, u + du], Y \leq u\} = 2f_X(u)F_X(u) du. \quad (1.132)$$

We use the density to compute the expectation of  $M$ :

$$E[M] = \int_0^\infty x f_M(x) dx = \quad (1.133)$$

$$= 2\lambda \int_0^\infty x(1 - e^{-\lambda x})e^{-\lambda x} dx = \quad (1.134)$$

$$= 2\lambda \int_0^\infty x e^{-\lambda x} dx - 2\lambda \int_0^\infty x e^{-2\lambda x} dx \quad (1.135)$$

$$= 2E[X] - E[L], \quad (1.136)$$

where the last equality follows from the previous exercises.

**Ex 1.16.** We can also compute  $f_M(y)$  (and  $f_L(x)$ ) from the joint density  $f_{L,M}(x, y)$  by marginalization. For this, use (1.6) to rewrite

$$F_{L,M}(x, y) = P\{L \leq x, M \leq y\}$$

in terms of  $F_X$  and  $F_Y$ . Once you checked your answer, you might see that you forgot an important condition in the earlier exercises..

**s.1.16.** First the joint distribution.

$$F_{L,M}(x, y) = P\{L \leq x, M \leq y\} \quad (1.137)$$

$$= 2P\{X \leq x, x < Y \leq y\} \quad (1.138)$$

$$= 2P\{X \leq x\} P\{x < Y \leq y\} \quad (1.139)$$

$$= 2F_X(x)(F_Y(y) - F_Y(x))I_{y>x}. \quad (1.140)$$

Note the indicator! Taking partial derivatives,

$$f_{L,M}(x, y) = \partial_x \partial_y F_{L,M}(x, y) \quad (1.141)$$

$$= \partial_x \partial_y 2F_X(x)(F_Y(y) - F_Y(x))I_{y>x} \quad (1.142)$$

$$= 2\partial_x F_X(x) \partial_y (F_Y(y) - F_Y(x))I_{y>x} \quad (1.143)$$

$$= 2f_X(x)f_Y(y)I_{y>x} \quad (1.144)$$

$$(1.145)$$

Now,

$$f_M(y) = \int_0^\infty f_{L,M}(x, y) dx \quad (1.146)$$

$$= 2 \int_0^y f_X(x) f_Y(y) dx \quad (1.147)$$

$$= 2f_Y(y) \int_0^y f_X(x) dx \quad (1.148)$$

$$= 2f_Y(y)F_X(y), \quad (1.149)$$

$$f_L(x) = \int_0^\infty f_{L,M}(x, y) dy \quad (1.150)$$

$$= 2f_X(x) \int_x^\infty f_Y(y) dy \quad (1.151)$$

$$= 2f_X(x)(1 - F_Y(x)). \quad (1.152)$$

In Chapter 8 you will learn how to work with transformations. Do this exercise when you read Chapter 8.

**Ex 1.17.** Find a function  $g$  that maps  $X, Y$  to  $L, M$ . Then compute  $f_{L,M}$  by using change of variables.

**s.1.17.** We need  $g$  first, and then we compute the Jacobian.

$$(u, v) = g(x, y) = (\min\{x, y\}, \max\{x, y\}). \quad (1.153)$$

Hence,  $g^{-1}(u, v) = \{(u, v), (v, u)\}$ ; note that this occurs because  $g$  is not one-to-one. However, this is not a real problem because we can easily keep track of ‘each branch’ of  $g^{-1}$ .

Next, noting that  $\partial_x \min\{x, y\} = I_{x \leq y}$ ,

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} I_{x \leq y} & I_{y < x} \\ I_{x > y} & I_{y \geq x} \end{vmatrix} = |I_{x \leq y} - I_{x > y}| = 1. \quad (1.154)$$

Now with the change of variable formula (in which I drop the Jacobian right away as it is 1),

$$f_{L,M}(u, v) = f_{X,Y}(u, v) + f_{X,Y}(v, u) = 2f_{X,Y}(u, v)I_{u \leq v}. \quad (1.155)$$

We can conclude that (1.4) and (1.5) are correct for  $X, Y \sim \text{Exp}(\lambda)$ . But, I have yet another way to check this, namely, from (1.3) I see that  $E[M - L] = E[X]$ . Let's try to verify that also. For this, we can first compute the conditional density  $f_{M-L|L}(y|x)$ ; once you did the next exercise you will directly see how to compute  $E[M - L]$ . Besides this, we also have to practice with conditional densities.<sup>2</sup>

**Ex 1.18.** Show that  $f_{M-L|L}(y|x) = \lambda e^{-\lambda y}$ , i.e.,  $M - L|L \sim \text{Exp}(\lambda)$ .

**h.1.18.** Use Bayes' rule, and compute  $f_{L,M-L}(x, y)$  first.

**s.1.18.**

$$f_{M-L|L}(y|x) = \frac{f_{L,M-L}(x, y)}{f_L(x)}. \quad (1.156)$$

Clearly, we must first find  $f_{L,M-L}(x, y)$ .

First I'll use the standard method. This is a bit tricky, because we have to take care of the limits in the integration. In itself the idea is not difficult, i.e., compute  $F_{L,M-L}(x, y)$ , then take partial derivatives, but the technical details require attention.

$$P\{L \leq x, M - L \leq y\} = 2P\{X \leq x, Y - X \leq y\} \quad (1.157)$$

$$= 2 \int_0^\infty \int_0^\infty I_{u \leq x, v-u \leq y, u \leq v} f_{X,Y}(u, v) du dv \quad (1.158)$$

$$= 2 \int_0^\infty \int_0^\infty I_{u \leq x, v-u \leq y, u \leq v} \lambda^2 e^{-\lambda u} e^{-\lambda v} du dv \quad (1.159)$$

$$= 2 \int_0^x \int_0^\infty I_{u \leq v \leq u+y} \lambda^2 e^{-\lambda u} e^{-\lambda v} du dv \quad (1.160)$$

$$= 2 \int_0^x \lambda e^{-\lambda u} \int_u^{u+y} \lambda e^{-\lambda v} dv du \quad (1.161)$$

$$= 2 \int_0^x \lambda e^{-\lambda u} (-e^{-\lambda v}) \Big|_u^{u+y} du \quad (1.162)$$

$$= 2 \int_0^x \lambda e^{-\lambda u} (e^{-\lambda u} - e^{-\lambda(u+y)}) du \quad (1.163)$$

$$= 2\lambda \int_0^x e^{-2\lambda u} du - 2\lambda \int_0^x e^{-\lambda(2u+y)} du \quad (1.164)$$

$$= 2\lambda \int_0^x e^{-2\lambda u} du - 2\lambda e^{-\lambda y} \int_0^x e^{-2\lambda u} du \quad (1.165)$$

$$= (1 - e^{-\lambda y}) 2\lambda \int_0^x e^{-2\lambda u} du \quad (1.166)$$

$$= (1 - e^{-\lambda y}) (-e^{-2\lambda u}) \Big|_0^x \quad (1.167)$$

$$= (1 - e^{-\lambda y}) (1 - e^{-2\lambda x}). \quad (1.168)$$

<sup>2</sup> And in BH Chapter 9 of the book you will learn that  $E[M - L|L] = E[M - L]$ , and for this we also need  $f_{M-L|L}$ .

For the density, we need to take the partial derivatives with respect to  $x$  and  $y$ :

$$f_{L,M-L}(x,y) = \partial_x \partial_y F_{L,M-L}(x,y) \quad (1.169)$$

$$= \partial_x \partial_y (1 - e^{-\lambda y})(1 - e^{-2\lambda x}) \quad (1.170)$$

$$= \lambda e^{-\lambda y} 2\lambda e^{-2\lambda x} \quad (1.171)$$

$$= f_Y(x)f_L(y). \quad (1.172)$$

If we plug this into (1.155) we see that

$$f_{M-L|L}(y|x) = f_Y(y) = \lambda e^{-\lambda y} \quad (1.173)$$

We arrive at the same result as in (1.154).

**Ex 1.19.** Use the solution of [1.18] to show that  $E[M - L] = E[X]$ .

**h.1.19.** Use the expression of  $f_{L,M-L}(x, y)$  to find the density of  $f_{M-L}(y)$ .

**s.1.19.**

$$E[M - L] = \int_0^\infty x f_{M-L}(x) dx. \quad (1.174)$$

Next,

$$f_{M-L}(y) = \int_0^\infty f_{L,M-L}(x, y) dx = \int_0^\infty f_L(x) f_Y(y) dx = f_Y(y) \int_0^\infty f_L(x) dx = f_Y(y). \quad (1.175)$$

Hence,

$$E[M - L] = \int_0^\infty y f_Y(y) dy = E[Y] = E[X]. \quad X, Y \text{ iid.} \quad (1.176)$$



**Ex 1.20.** Here is another opportunity to check our work, but now with changing variables, cf. Chapter 8. Find  $f_{L,M-L}(x,y)$  by means of changing of variables from  $L,M$  to  $L,M-L$ .

**s.1.20.** The function  $h(x,y) = (x,y-x)I_{x \leq y}$  maps  $(L,M)$  to  $(L,M-L)$ . Then,  $h^{-1}(u,v) = (u,u+v)$ . For the Jacobian

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} I_{x \leq y} & 0 \\ -I_{x \leq y} & I_{x \leq y} \end{vmatrix} = I_{x \leq y}. \quad (1.177)$$

And with this,

$$f_{L,M-L}(u,v) = f_{L,M}(u,u+v)I_{u \leq u+v} \quad (1.178)$$

$$= f_{L,M}(u,u+v) \quad (1.179)$$

$$= 2f_X(u)f_Y(u+v) \quad (1.180)$$

$$= 2\lambda^2 e^{-2u} e^{-\lambda v} \quad (1.181)$$

$$= f_L(u)f_Y(v). \quad (1.182)$$

We did a number of checks for the case  $X, Y, \text{iid.}, \sim \text{Exp}(\lambda)$ . But I have yet another idea to see whether the results we have obtained so far are consistent. For this I use BH Ex 5.43 in which it is shown that the geometric distribution is the discrete analog of the exponential distribution. Now I want to see how, by taking proper limits, the results of this section can be obtained as limiting cases of those of Section 1.1.

More specically, I want to see intuitively<sup>3</sup> how  $X \sim \text{Geo}(\lambda/n)$  approaches  $Y \sim \text{Exp}(\lambda)$  as  $n \rightarrow \infty$ .

**Ex 1.21.** Divide the interval  $[0, \infty)$  into many small intervals of length  $1/n$ . Let  $X \sim \text{Geo}(\lambda/n)$  for some  $\lambda > 0$  and  $n \gg 0$ , take some  $x \geq 0$ , let  $i$  be such that  $x \in [i/n, (i+1)/n)$ . Then show that

$$P\{X/n \approx x\} \approx \frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right)^{xn}. \quad (1.183)$$

**s.1.21.** First,

$$P\{X/n \approx x\} = P\{X/n \in [i/n, (i+1)/n)\} = P\{X \in [i, i+1)\} = pq^i \approx pq^{nx}. \quad (1.184)$$

Now take  $p = \lambda/n$  and  $q = 1 - p = 1 - \lambda/n$ , then,

$$P\{X/n \approx x\} \approx \frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right)^{xn}. \quad (1.185)$$

---

<sup>3</sup> A formal proof is too hard for this course.

From BH.A.2.5,  $\lim_{n \rightarrow \infty} (1 - \lambda/n)^n = e^{-\lambda}$ , which is a standard limit but, interestingly, not entirely trivial to prove.<sup>4</sup> Next, by introducing the highly intuitive definition  $dx = \lim_{n \rightarrow \infty} 1/n$ , it follows from (1.182) that

$$\mathbb{P}\{X/n \approx x\} \rightarrow \lambda e^{-\lambda x} dx, \quad \text{as } n \rightarrow \infty. \quad (1.186)$$

In fact, working with infinitesimals such as  $dx$  is one of my favorite methods, for instance to derive densities. The intuition is that  $dx$  is some really small number but not zero. In a tiny bit more formal words,  $dx < 1/n$  for any  $n$ , but  $dx \neq 0$ . For details, check out the web; search on nonstandard calculus. However, while the method is very intuitive, it's easy to mess up. (Power tools are nice, but often only when you know how to handle them and practice a lot.)wreck havoc. If you don't want to learn this, no problem, it will not be part of the course proper.

In fact, if you don't like infinitesimals, it is also possible obtain this limit with moment-generating functions (and we can practice with this function again).

**Ex 1.22.** Derive the moment-generating function  $M_{X/n}(s)$  of  $X/n$  when  $X \sim \text{Geo}(p)$ . Then, let  $p = \lambda/n$ , and show that  $\lim_{n \rightarrow \infty} M_{X/n}(s) = M_Y(s)$ , where  $Y \sim \text{Exp}(\lambda)$ .

**h.1.22.** Use that, as in the Poisson distribution,  $e^{-\lambda} = \sum_{i=0}^{\infty} \lambda^i / i!$ . In fact, this precisely Taylor's expansion of  $e^{s/n}$ .

**s.1.22.**

$$M_{X/n}(s) = \mathbb{E}\left[e^{sX/n}\right] = \sum_i e^{si/n} p q^i = p \sum_i (q e^{s/n})^i = \frac{p}{1 - q e^{s/n}}. \quad (1.187)$$

With  $p = \lambda/n$  this becomes

$$M_{X/n}(s) = \frac{\lambda/n}{1 - (1 - \lambda/n)(1 + s/n + 1/n^2 \times (\dots))} \quad (1.188)$$

$$= \frac{\lambda/n}{\lambda/n - s/n + 1/n^2 \times (\dots)} \quad (1.189)$$

$$= \frac{\lambda}{\lambda - s + 1/n \times (\dots)} \quad (1.190)$$

$$\rightarrow \frac{\lambda}{\lambda - s}, \quad \text{as } n \rightarrow \infty, \quad (1.191)$$

where we write  $1/n^2 \times (\dots)$  for all terms that will disappear when we take the limit  $n \rightarrow \infty$ . This is just handy notation to hide details in which we are not interested.

<sup>4</sup> If you like mathematics, check the neat proof in Rudin's Principles of mathematical analysis.

**Ex 1.23.** Check that  $\lim_{n \rightarrow \infty} E[L/n] = 1/(2\lambda)$  when  $X \sim \text{Geo}(\lambda/n)$ . Conclude that we retrieve (1.5).

**h.1.23.** Use [1.4].

**s.1.23.**

$$E[L/n] = \frac{1}{n} E[L] = \frac{1}{n} \frac{q^2}{1-q^2} \quad (1.192)$$

$$= \frac{1}{n} \frac{(1-\lambda/n)^2}{1-(1-\lambda/n)^2} \quad (1.193)$$

$$= \frac{1}{n} \frac{1-2\lambda/n+(\lambda/n)^2}{2\lambda/n+(\lambda/n)^2} \quad (1.194)$$

$$= \frac{1-2\lambda/n+(\lambda/n)^2}{2\lambda+\lambda^2/n} \quad (1.195)$$

$$\rightarrow \frac{1}{2\lambda}. \quad (1.196)$$

**Ex 1.24.** Here is yet another check on the correctness of  $f_M(x)$ . Show that the PMF  $P\{M = k\}$  of (1.44) converges to  $f_M(x)$  of [1.15] when  $n \rightarrow \infty$ . Take  $k$  suitable.

**s.1.24.** Take  $p = \lambda/n$ ,  $q = 1 - \lambda/n$ , and  $k \approx xn$ , hence  $k/n \approx x$ . Then,

$$P\{M/n = k/n\} = 2pq^{k/n}(1 - q^{k/n}) + p^2q^{2k/n} \quad (1.197)$$

$$= 2\frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right)^{k/n}\left(1 - \left(1 - \frac{\lambda}{n}\right)^{k/n}\right) + \frac{\lambda^2}{n^2}\left(1 - \frac{\lambda}{n}\right)^{2k/n} \quad (1.198)$$

$$\rightarrow 2\lambda dx e^{-\lambda x}(1 - e^{-\lambda x}) + \lambda^2 dx^2 e^{-2\lambda x}. \quad (1.199)$$

Now observe that the second term, proportional to  $dx^2$  can be neglected.

All in all, we have checked and double checked all our expressions and limits for the geometric and exponential distribution. We had success too: the solution of the last exercise provides the key to understand why (1.2) and (1.3) are true for exponentially distributed r.v.s, but not for geometric random variables. In fact, in the solutions we can see the term corresponding to  $X = Y = i$  for  $X, Y \sim \text{Geo}(p)$  becomes negligibly small when  $n \rightarrow 0$ . In words, the probability that  $X$  and  $Y$  are the same is non-negligible when they are discrete, but it is when  $X$  and  $Y$  are continuous. So, to resolve our leading problem we should exploit that idea.

### 1.3 The solution

Let us now try to repair (1.3) for the case  $X, Y \sim \text{Geo}(p)$ . We should be careful about the non-negligible case that  $M = L$ , so we move on, carefully, step by step. The following must be true:

$$E[M] = E[L] + E[(M - L)I_{M > L}] = E[L] + 2E[(Y - X)I_{Y > X}], \quad (1.200)$$

because either  $M = L$  or  $M > L$ . Recall from earlier work that the factor 2 in the second equality follows from the fact that  $X, Y$  iid.

**Ex 1.25.** Show that

$$2E[(Y - X)I_{Y > X}] = \frac{2q}{1 - q^2}. \quad (1.201)$$

Combine this with the expression for  $E[L]$  of [1.4] to obtain (1.38) for  $E[M] = 2E[X] - E[L]$ , thereby verifying the correctness of (1.199).

**s.1.25.**

$$E[(Y - X)I_{Y > X}] = p^2 \sum_{ij} (j - i) I_{j > i} q^i q^j \quad (1.202)$$

$$= p^2 \sum_i q^i \sum_{j=i+1}^{\infty} (j - i) q^j \quad (1.203)$$

$$= p^2 \sum_i q^i q^i \sum_{k=1}^{\infty} k q^k, \quad k = j - i \quad (1.204)$$

$$= p \sum_i q^{2i} E[X] \quad (1.205)$$

$$= \frac{p}{1 - q^2} E[X] \quad (1.206)$$

$$= \frac{p}{1 - q^2} \frac{q}{p} \quad (1.207)$$

$$= \frac{q}{1 - q^2}. \quad (1.208)$$

With this,

$$E[L] + 2E[(Y - X)I_{Y > X}] = \frac{q^2}{1 - q^2} + \frac{2q}{1 - q^2} = \frac{q}{1 - q} \frac{q + 2}{1 + q}, \quad (1.209)$$

where I use that  $1 - q^2 = (1 - q)(1 + q)$ .

While (1.199) is correct, I am still not happy with the second part of (1.199) (I find this hard/unintuitive to interpret). Restarting again from scratch, here is another attempt to rewrite  $E[M]$ :

$$E[M] = E[L] + E[Z I_{M>L}], \quad (1.210)$$

where I take  $Z \sim \text{FS}(p)$ , i.e.,  $Z$  has the first success distribution with parameter  $p$ , in other words,  $Z \sim X + 1$  with  $X \sim \text{Geo}(p)$ . To see why this might be true, I reason like this. After ‘seeing’  $L$ , we want to restart. Let  $Z$  be the time from the restart to  $M$ . When  $Z \sim \text{Geo}(p)$ , it might happen that  $Z = 0$  (with positive probability  $p$ ). But if  $Z = 0$ , then  $M = L$ , and in that case, we should not restart. Hence, if  $Z \sim \text{Geo}(p)$  we are ‘double counting’ when  $Z = 0$ . By including the condition  $M > L$  and by taking  $Z \sim \text{FS}(p)$  (so that  $Z > 0$ ) I can prevent this.

**Ex 1.26.** Show that

$$E[Z I_{M>L}] = \frac{2q}{1 - q^2}, \quad (1.211)$$

i.e., the same as (1.200), hence (1.209) is correct.

**h.1.26.** Observe that  $Z$  is independent from  $X$  and  $Y$ , hence from  $M$  and  $L$

**s.1.26.** With the hint:

$$E[Z I_{M>L}] = E[Z] E[I_{M>L}] = E[Z] P\{M > L\} = \frac{1}{p} \frac{2pq}{1 - q^2} = \frac{2q}{1 - q^2}, \quad (1.212)$$

where, from the book,  $E[Z] = 1/p$ , and  $P\{M > L\}$  follows from [1.11].

I am nearly happy, but I want to see that (1.209), which is correct for discrete r.v.s, has also the correct limiting behavior, similar to [1.23].

**Ex 1.27.** Show that  $E[Z/n I_{M>L}] \rightarrow 1/\lambda$ , which is indeed the expectation of an  $\text{Exp}(\lambda)$  r.v. And thus, when  $X, Y \sim \text{Exp}(\lambda)$ ,  $E[M] = E[L] + E[X]$ .

**s.1.27.** By independence,

$$E[Z/n I_{M>L}] = E[Z/n] P\{M > L\}. \quad (1.213)$$

Then,

$$P\{M > L\} = \frac{2pq}{1-q^2} = \frac{2\lambda/n(1-\lambda/n)}{1-(1-\lambda/n)^2} = \frac{2\lambda/n(1-\lambda/n)}{2\lambda/n - \lambda^2/n^2} = \frac{2(1-\lambda/n)}{2-\lambda/n} \rightarrow 1, \quad (1.214)$$

and

$$E[Z/n] = \frac{1}{n} E[Z] = \frac{1}{n} \frac{p}{\lambda/n} = 1/\lambda. \quad (1.215)$$



It took me a long time, and a lot of work, to understand how to resolve the confusing problem. Perhaps how to tackle this problem is clear to any other person, but I had a nice time working on it, and I learned a lot. In particular, I find (1.209) a nice and revealing equation.