

# Memoryless excursions

## Notes on joint distributions

Nicky van Foreest

January 13, 2021

## CONTENTS

1	A confusing problem	3
2	Discrete memoryless r.v.s	5
3	Continuous memoryless r.v.s	19
4	The solution	31
5	Working with densities	35

## 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 memorylessness, 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 noting that  $E[L]$  cancels we get  $2E[M] = 3E[X]$ , hence:

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

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

You might want to compare this to BH.7.2.2<sup>1</sup>, where they analyze the case with uniform  $X, Y$ .

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., BH.5.6.5. 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 some much more general goals. They are meant to

1. become familiar with all concepts BH 7.1.
2. provide a recap of concepts and results of BH chapters 1 to 6.
3. illustrate the change of variables formulas of BH 8.
4. 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 much simpler. All exercises below have solutions which focus on the computational steps. I will provide motivation and explanations in the online lectures.

Below I use the notation  $\partial_x = \partial/\partial x$ , and  $G_X(x) = 1 - F_X(x)$  for the survivor function of the r.v.  $X$ .

To solve the problems, I often use the following step-wise approach:

---

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

1. To compute probabilities of functions of r.v.s: I use the fundamental bridge, specifically, for discrete r.v.s

$$P\{g(X, Y) \in A\} = \sum_i \sum_j I_{(i,j) \in A} P\{X = i, Y = j\}, \quad (1.6)$$

and for continuous r.v.s,

$$\begin{aligned} P\{g(X, Y) \in A\} &= E\left[I_{g(X,Y) \in A}\right] \\ &= \iint I_{g(x,y) \in A} f_{X,Y}(x, y) \, dx \, dy \\ &= \iint_{g^{-1}(A)} f_{X,Y}(x, y) \, dx \, dy, \\ &= \iint_{\{(x,y): g(x,y) \in A\}} f_{X,Y}(x, y) \, dx \, dy. \end{aligned} \quad (1.7)$$

Conceptually this approach is easy, but simplifying the summation or integration is not always simple. However, the advantage is that I don't need special tricks or knowledge to solve the problem; it's 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  $F_{X,Y}(x, y) = P\{X \leq x, Y \leq y\} = E[I_{X \leq x, Y \leq y}]$ . In other words, I apply 2D LOTUS to the function  $g(x, y) = I_{X \leq x, Y \leq y}$ .
3. To get the joint density  $f_{X,Y}$ , I often first compute the joint distribution  $F_{X,Y}$  and then take partial derivatives because  $f_{X,Y} = \partial_x \partial_y F_{X,Y}$ . This approach is 'trick-free', but requires at times the hard work of solving the integrals to find  $F_{X,Y}$ .

Solving the exercises of Section 3 requires quite a lot of technical stuff with integrals. With the tools of BH ch 8 this can be done (considerably) easier. So in Section 5 we redo the work of Section 3, but now with change of variables techniques and densities. Read it once you study BH ch 8.

In the exercises below we first try to see which of (1.2) and (1.3) 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.

## 2 DISCRETE MEMORYLESS R.V.S

Let  $X, Y$  be iid.  $\sim \text{Geo}(p)$ . Write  $q = 1 - p$ .

**Ex 2.1.** As a recap of earlier chapters of BH, show that

$$\mathbb{P}\{X > j\} = q^{j+1}, \quad \mathbb{P}\{X \geq j\} = q^j, \quad \mathbb{E}[X] = \frac{q}{p}. \quad (2.1)$$

What is the domain of  $X$ ?

**s.2.1.** 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.

$$\mathbb{P}\{X > 0\} = \mathbb{P}\{\text{failure}\} = q \quad (2.2)$$

$$\mathbb{P}\{X > j\} = q \mathbb{P}\{X > j-1\} \implies \mathbb{P}\{X > j\} = q^j \mathbb{P}\{X > 0\} = q^{j+1}. \quad (2.3)$$

$$\mathbb{P}\{X = j\} = \mathbb{P}\{X > j-1\} - \mathbb{P}\{X > j\} = q^j - q^{j+1} = (1-q)q^j = pq^j. \quad (2.4)$$

$$\mathbb{E}[X] = p \cdot 0 + q(1 + \mathbb{E}[X]) \implies \mathbb{E}[X] = q/(1-q) = q/p. \quad (2.5)$$

Now the same answers, but with the regular method.

$$\mathbb{P}\{X > j\} = \sum_{i=j+1}^{\infty} \mathbb{P}\{X = i\} \quad (2.6)$$

$$= p \sum_{i=j+1}^{\infty} q^i \quad (2.7)$$

$$= p \sum_{i=0}^{\infty} q^{j+1+i} \quad (2.8)$$

$$= pq^{j+1} \sum_{i=0}^{\infty} q^i \quad (2.9)$$

$$= pq^{j+1} \frac{1}{1-q} = pq^{j+1} \frac{1}{p} = q^{j+1}. \quad (2.10)$$

From this,

$$\mathbb{P}\{X \geq j\} = \mathbb{P}\{X > j-1\} = q^j, \quad (2.11)$$

$$(2.12)$$

The use of indicator variables is particularly important:

$$\mathbb{E}[X] = \sum_{i=0}^{\infty} i \mathbb{P}\{X = i\} \quad (2.13)$$

$$= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} I_{j < i} \mathbb{P}\{X = i\} \quad (2.14)$$

$$= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} I_{j < i} \mathbb{P}\{X = i\} \quad (2.15)$$

$$= \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} \mathbb{P}\{X = i\} \quad (2.16)$$

$$= \sum_{j=0}^{\infty} \mathbb{P}\{X > j\} \quad (2.17)$$

$$= \sum_{j=0}^{\infty} q^{j+1} \quad (2.18)$$

$$= q \sum_{j=0}^{\infty} q^j \quad (2.19)$$

$$= q/(1 - q) = q/p. \quad (2.20)$$

**Ex 2.2.** Check that  $X$ , and hence  $Y$ , is memoryless.

**h.2.2.** What is the definition of memorylessness for discrete r.v.s?

**s.2.2.**

$$P\{X \geq n + m \mid X \geq m\} = \frac{P\{X \geq n + m, X \geq m\}}{P\{X \geq m\}} \quad (2.21)$$

$$= \frac{P\{X \geq n + m\}}{P\{X \geq m\}} \quad (2.22)$$

$$= \frac{q^{n+m}}{q^m} \quad (2.23)$$

$$= q^n = P\{X \geq n\}. \quad (2.24)$$

**Ex 2.3.** 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.2.3.**

$$P\{L \geq k\} = \sum_{ij} I_{\min\{i,j\} \geq k} P\{X = i, Y = j\} \quad (2.25)$$

$$= \sum_{i \geq k} \sum_{j \geq k} P\{X = i\} P\{Y = j\} \quad (2.26)$$

$$= P\{X \geq k\} P\{Y \geq k\} = q^k q^k = q^{2k}. \quad (2.27)$$

$P\{L > i\}$  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 2.4.** Show that  $E[L] = q^2 / (1 - q^2)$ .

**h.2.4.**  $X \sim \text{Geo}(1 - q) \implies E[X] = q / (1 - q)$ . Now use that  $L \sim \text{Geo}(1 - q^2)$ .

**s.2.4.** Immediate from the hint and [\[2.3\]](#).

**Ex 2.5.** Show that

$$\mathbb{E}[L] + \mathbb{E}[X] = \frac{q}{1-q} \frac{1+2q}{1+q}. \quad (2.28)$$

**h.2.5.** Use that  $1 - q^2 = (1 - q)(1 + q)$ .

**s.2.5.**

$$\mathbb{E}[L] + \mathbb{E}[X] = \frac{q^2}{1-q^2} + \frac{q}{1-q} \quad (2.29)$$

$$= \frac{q}{1-q} \left( \frac{q}{1+q} + 1 \right) \quad (2.30)$$

$$= \frac{q}{1-q} \frac{1+2q}{1+q} \quad (2.31)$$

**Ex 2.6.** Supposing that (1.2) is true, show that

$$E[M] = 2E[X] - E[L] = \frac{q}{1-q} \frac{2+q}{1+q}. \quad (2.32)$$

Show also that this is not the same as (2.28) (unless  $q = 0$ ). Conclude that this violates the simultaneous truth of (1.2) and (1.3).

**s.2.6.**

$$E[M] = 2E[X] - E[L] \quad (2.33)$$

$$= 2 \frac{q}{1-q} - \frac{q^2}{1-q^2} \quad (2.34)$$

$$= \frac{q}{1-q} \left( 2 - \frac{q}{1+q} \right) \quad (2.35)$$

$$= \frac{q}{1-q} \left( \frac{2+2q}{1+q} - \frac{q}{1+q} \right) \quad (2.36)$$

$$= \frac{q}{1-q} \frac{2+q}{1+q}. \quad (2.37)$$

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 (2.32), hence (1.2)

**Ex 2.7.** Idea 1: Show that for the PMF of  $M$ :

$$p_M(k) = P\{M = k\} = 2pq^k(1 - q^k) + p^2q^{2k}. \quad (2.38)$$

Then use this to show (2.32).

**s.2.7.**

$$P\{M = k\} = P\{\max\{X, Y\} = k\} \quad (2.39)$$

$$= p^2 \sum_{ij} I_{\max\{i,j\}=k} q^i q^j \quad (2.40)$$

$$= 2p^2 \sum_{ij} I_{i=k} I_{j < k} q^i q^j + p^2 \sum_{ij} I_{i=j=k} q^i q^j \quad (2.41)$$

$$= 2p^2 q^k \sum_{j < k} q^j + p^2 q^{2k} \quad (2.42)$$

$$= 2p^2 q^k \frac{1 - q^k}{1 - q} + p^2 q^{2k} \quad (2.43)$$

$$= 2pq^k(1 - q^k) + p^2 q^{2k} \quad (2.44)$$

$$= 2pq^k + (p^2 - 2p)q^{2k} \quad (2.45)$$

$$= 2P\{X = k\} - P\{L = k\}, \quad (2.46)$$

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

$$E[M] = \sum_k k P\{M = k\} \quad (2.47)$$

$$= \sum_k k(2P\{X = k\} - P\{L = k\}) \quad (2.48)$$

$$= 2E[X] - E[L]. \quad (2.49)$$

**Ex 2.8.** Idea 2: Use  $P\{M \leq k\} = (P\{X \leq k\})^2$  to check (2.38).

**h.2.8.**  $P\{M = k\} = P\{M \leq k\} - P\{M \leq k - 1\}$ .

**s.2.8.**

$$P\{M \leq k\} = P\{X \leq k, Y \leq k\} \quad (2.50)$$

$$= P\{X \leq k\} P\{Y \leq k\} \quad (2.51)$$

$$= (1 - P\{X > k\})(1 - P\{Y > k\}) \quad (2.52)$$

$$= (1 - q^{k+1})^2. \quad (2.53)$$

$$P\{M = k\} = P\{M \leq k\} - P\{M \leq k - 1\} \quad (2.54)$$

$$= 1 - 2q^{k+1} + q^{2k+2} - (1 - 2q^k + q^{2k}) \quad (2.55)$$

$$= 2q^k(1 - q) + q^{2k}(q^2 - 1) \quad (2.56)$$

$$= 2P\{X = k\} - q^{2k}(1 - q^2) \quad (2.57)$$

$$= 2P\{X = k\} - P\{L = k\}. \quad (2.58)$$

**Ex 2.9.** Idea 3a: show that the joint PMF

$$p_{L,M}(i, j) = \mathbf{P}\{L = i, M = j\} = 2p^2q^{i+j} I_{j>i} + p^2q^{2i} I_{i=j}. \quad (2.59)$$

**s.2.9.**

$$\mathbf{P}\{L = i, M = j\} = 2\mathbf{P}\{X = i, Y = j\} I_{j>i} + \mathbf{P}\{X = Y = i\} I_{i=j} \quad (2.60)$$

$$= 2p^2q^{i+j} I_{j>i} + p^2q^{2i} I_{i=j}. \quad (2.61)$$

**Ex 2.10.** Idea 3b: Use [2.9] to compute the marginal PMFs of  $M$  and of  $L$ .

**h.2.10.** For the first, marginalize out  $L$ , for the second, marginalize out  $M$ .

**s.2.10.**

$$P\{M = j\} = \sum_i P\{L = i, M = j\} \quad (2.62)$$

$$= \sum_i (2p^2 q^{i+j} I_{j>i} + p^2 q^{2i} I_{i=j}) \quad (2.63)$$

$$= 2p^2 q^j \sum_{i=0}^{j-1} q^i + p^2 q^{2j} \quad (2.64)$$

$$= 2p q^j (1 - q^j) + p^2 q^{2j} \quad (2.65)$$

$$= 2p q^j + (p^2 - 2p) q^{2j}, \quad (2.66)$$

which is equal to (2.45).

$$P\{L = i\} = \sum_j P\{L = i, M = j\} \quad (2.67)$$

$$= \sum_j (2p^2 q^{i+j} I_{j>i} + p^2 q^{2i} I_{i=j}) \quad (2.68)$$

$$= 2p^2 q^i \sum_{j=i+1}^{\infty} q^j + p^2 q^{2i} \quad (2.69)$$

$$= 2p^2 q^{2i+1} \sum_{j=0}^{\infty} q^j + p^2 q^{2i} \quad (2.70)$$

$$= 2p q^{2i+1} + p^2 q^{2i} \quad (2.71)$$

$$= p q^{2i} (2q + p) \quad (2.72)$$

$$= (1 - q) q^{2i} (q + 1), \quad p + q = 1, \quad (2.73)$$

$$= (1 - q^2) q^{2i}, \quad (2.74)$$

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

**h.2.11.** Show that  $P\{L = i \mid M > L\} = P\{L = i\}$ ; recall, to show independence we need the probability (measure)  $P\{\cdot\}$ .

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

**s.2.11.**

$$P\{L = i \mid M > L\} = \frac{P\{L = i, M > L\}}{P\{M > L\}}. \quad (2.75)$$

$$P\{L = i, M > L\} = 2P\{X = i, Y > X\} \quad (2.76)$$

$$= 2P\{X = i\} P\{Y > i\} \quad (2.77)$$

$$= 2pq^i \cdot q^{i+1} \quad (2.78)$$

$$= 2pq^{2i+1}. \quad (2.79)$$

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  $Y$ .

$$P\{M > L\} = P\{X > Y\} + P\{Y > X\} \quad (2.80)$$

$$= 2P\{X > Y\} \quad (2.81)$$

$$= 2 \sum_j P\{X > j, Y = j\} \quad (2.82)$$

$$= 2 \sum_j P\{X > j\} P\{Y = j\} \quad (2.83)$$

$$= 2p \sum_j q^{j+1} q^j \quad (2.84)$$

$$= 2p \sum_j q^{2j+1} \quad (2.85)$$

$$= 2pq \sum_j q^{2j} \quad (2.86)$$

$$= 2pq / (1 - q^2). \quad (2.87)$$



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

$$P\{M > L\} = P\{\max\{X, Y\} > \min\{X, Y\}\} \quad (2.88)$$

$$= p^2 \sum_{ij} I_{\max\{i,j\} > \min\{i,j\}} p_{XY}(i, j) \quad (2.89)$$

$$= p^2 \sum_{ij} I_{\max\{i,j\} > \min\{i,j\}} p_X(i) p_Y(j) \quad (2.90)$$

$$= p^2 \sum_{ij} I_{\max\{i,j\} > \min\{i,j\}} q^i q^j \quad (2.91)$$

$$= 2p^2 \sum_i \sum_{j>i} q^i q^j \quad (2.92)$$

$$= 2p^2 \sum_i q^i \sum_{j>i} q^j \quad (2.93)$$

$$= 2p^2 \sum_i q^i \sum_{j=0}^{\infty} q^{i+1+j} \quad (2.94)$$

$$= 2p^2 \sum_i q^{2i+1} \sum_{j=0}^{\infty} q^j \quad (2.95)$$

$$= 2p \sum_i q^{2i+1} \quad (2.96)$$

$$= 2pq \sum_i q^{2i} \quad (2.97)$$

$$= 2pq / (1 - q^2). \quad (2.98)$$

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

$$P\{M > L\} = \sum_i P\{L = i, M > L\} = \sum_i 2pq^{2i+1} = 2pq \sum_i q^{2i} = \frac{2pq}{1 - q^2}. \quad (2.99)$$

Hence,

$$P\{L = i | M > L\} = \frac{P\{L = i, M > L\}}{P\{M > L\}} \quad (2.100)$$

$$= \frac{2pq^{2i+1}}{2pq / (1 - q^2)} \quad (2.101)$$

$$= q^{2i} (1 - q^2) \quad (2.102)$$

$$= P\{L = i\}, \quad (2.103)$$

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.

### 3 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 3.1.** Show that  $X$ , and hence  $Y$ , is memoryless.

**s.3.1.** See BH.

**Ex 3.2.** Show that  $E[X] = 1/\lambda$ .

**h.3.2.** Use partial integration.

**s.3.2.** It is essential that you know both methods to solve this integral.

$$E[X] = \int_0^{\infty} x f_X(x) dx \quad (3.1)$$

$$= \int_0^{\infty} x \lambda e^{-\lambda x} dx \quad (3.2)$$

$$= -xe^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \quad (3.3)$$

$$= -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}. \quad (3.4)$$

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

$$E[X] = \int_0^{\infty} x f_X(x) dx \quad (3.5)$$

$$= \int_0^{\infty} x \lambda e^{-\lambda x} dx \quad (3.6)$$

$$= \frac{1}{\lambda} \int_0^{\infty} u e^{-u} du, \quad (3.7)$$

by the substitution  $u = \lambda x \implies du = d(\lambda x) \implies du = \lambda dx \implies dx = du/\lambda$ . 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 3.3.** Show that  $F_L(x) = 1 - e^{-2\lambda x}$ , hence  $L \sim \text{Exp}(2\lambda)$  and therefore  $E[L] = 1/(2\lambda)$ . Realize that with this we have an independent check of (1.5).

**s.3.3.** Using independence and the specific property of the r.v.  $L$  that  $\{L > x\} \iff \{X > x, Y > x\}$ :

$$G_L(x) = P\{L > x\} = P\{X > x, Y > x\} = G_X(x)^2 = e^{-2\lambda x}. \quad (3.8)$$

The result follows since  $F_L(x) = 1 - G_L(x)$ .

**Ex 3.4.** Compute  $F_M(v) = P\{M \leq v\}$ . With this, show that  $f_M(v) = 2(1 - e^{-\lambda v})\lambda e^{-\lambda v}$ , and then use this to compute  $E[M]$  to show that (1.4) holds.

**s.3.4.** Here are two related ways. This is one:

$$P\{M \leq v\} = P\{X \leq v, Y \leq v\} = (F_X(v))^2, \quad (3.9)$$

since  $X, Y$  iid.

The second is based on the fundamental bridge:

$$P\{M \leq v\} = E[I_{M \leq v}] \quad (3.10)$$

$$= \int_0^\infty \int_0^\infty I_{x \leq v, y \leq v} f_{XY}(x, y) dx dy \quad (3.11)$$

$$= \int_0^\infty \int_0^\infty I_{x \leq v, y \leq v} f_X(x) f_Y(y) dx dy \quad (3.12)$$

$$= \int_0^v f_X(x) dx \int_0^v f_Y(y) dy \quad (3.13)$$

$$= F_X(v)F_Y(v) = (F_X(v))^2. \quad (3.14)$$

We get the same result.

Now the density:

$$f_M(v) = F'_M(v) = 2F_X(v)f_X(v) = 2(1 - e^{-\lambda v})\lambda e^{-\lambda v}. \quad (3.15)$$

With this,

$$E[M] = \int_0^\infty v f_M(v) dv = \quad (3.16)$$

$$= 2\lambda \int_0^\infty v(1 - e^{-\lambda v})e^{-\lambda v} dv = \quad (3.17)$$

$$= 2\lambda \int_0^\infty v e^{-\lambda v} dv - 2\lambda \int_0^\infty v e^{-2\lambda v} dv \quad (3.18)$$

$$= 2E[X] - E[L], \quad (3.19)$$

where the last equality follows from the previous exercises.

We can also compute  $f_M(y)$  (and  $f_L(x)$ ) by marginalization of the joint density  $f_{L,M}(x, y)$ . However, to get  $f_{L,M}$ , we first need  $F_{L,M}$ .

**Ex 3.5.** Use the fundamental bridge (1.7) to show that

$$F_{L,M}(u, v) = P\{L \leq u, M \leq v\} = 2 \int_0^u (F_Y(v) - F_Y(x)) f_X(x) dx,$$

with  $u \leq v$  (of course). Then take partial derivatives to show that

$$f_{L,M}(u, v) = 2f_X(u)f_Y(v) I_{u \leq v}. \quad (3.20)$$

**s.3.5.** First the joint distribution. With  $u \leq v$ ,

$$F_{L,M}(u, v) = P\{L \leq u, M \leq v\} \quad (3.21)$$

$$= 2 \iint I_{x \leq u, y \leq v, x \leq y} f_{X,Y}(x, y) dx dy \quad (3.22)$$

$$= 2 \int_0^u \int_x^v f_Y(y) dy f_X(x) dx \quad \text{independence} \quad (3.23)$$

$$= 2 \int_0^u (F_Y(v) - F_Y(x)) f_X(x) dx. \quad (3.24)$$

Taking partial derivatives,

$$f_{L,M}(u, v) = \partial_v \partial_u F_{L,M}(u, v) \quad (3.25)$$

$$= 2 \partial_v \partial_u \int_0^u (F_Y(v) - F_Y(x)) f_X(x) dx \quad (3.26)$$

$$= 2 \partial_v \{(F_Y(v) - F_Y(u)) f_X(u)\} \quad (3.27)$$

$$= 2 f_X(u) \partial_v F_Y(v) \quad (3.28)$$

$$= 2 f_X(u) f_Y(v). \quad (3.29)$$

**Ex 3.6.** In (3.20) marginalize out  $L$  to find  $f_M$ , and marginalize out  $M$  to find  $f_L$ .

**s.3.6.**

$$f_M(v) = \int_0^\infty f_{L,M}(u, v) \, du \quad (3.30)$$

$$= 2 \int_0^v f_X(u) f_Y(v) \, du \quad (3.31)$$

$$= 2f_Y(v) \int_0^v f_X(u) \, du \quad (3.32)$$

$$= 2f_Y(v)F_X(v), \quad (3.33)$$

$$f_L(u) = \int_0^\infty f_{L,M}(u, v) \, dv \quad (3.34)$$

$$= 2f_X(u) \int_u^\infty f_Y(v) \, dv \quad (3.35)$$

$$= 2f_X(u)G_Y(u). \quad (3.36)$$



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  $E[M - L]$ .

**Ex 3.7.** Show that

$$F_{L,M-L}(x,y) = (1 - e^{-\lambda y})(1 - e^{-2\lambda x}) = F_Y(y)F_L(x). \quad (3.37)$$

Conclude that  $M - L$  and  $L$  are independent, and  $M - L \sim Y$ , hence  $E[M - L] = E[Y] = E[X]$ , as  $X$  and  $Y$  are iid.

**s.3.7.** I'll use the standard method, i.e., the fundamental bridge. In itself the idea is not difficult, but the technical details require attention, in particular the limits in the integrations.

$$P\{L \leq x, M - L \leq y\} = 2P\{X \leq x, Y - X \leq y, X \leq Y\} \quad (3.38)$$

$$= 2 \int_0^\infty \int_0^\infty I_{u \leq x, v-u \leq y, u \leq v} f_{X,Y}(u,v) du dv \quad (3.39)$$

$$= 2 \int_0^\infty \int_0^\infty I_{u \leq x, v-u \leq y, u \leq v} \lambda^2 e^{-\lambda u} e^{-\lambda v} du dv \quad (3.40)$$

$$= 2 \int_0^x \int_0^\infty I_{u \leq v \leq u+y} \lambda^2 e^{-\lambda u} e^{-\lambda v} dv du \quad (3.41)$$

$$= 2 \int_0^x \lambda e^{-\lambda u} \int_u^{u+y} \lambda e^{-\lambda v} dv du \quad (3.42)$$

$$= 2 \int_0^x \lambda e^{-\lambda u} (-e^{-\lambda v}) \Big|_u^{u+y} du \quad (3.43)$$

$$= 2 \int_0^x \lambda e^{-\lambda u} (e^{-\lambda u} - e^{-\lambda(u+y)}) du \quad (3.44)$$

$$= 2\lambda \int_0^x e^{-2\lambda u} du - 2\lambda \int_0^x e^{-\lambda(2u+y)} du \quad (3.45)$$

$$= 2\lambda \int_0^x e^{-2\lambda u} du - 2\lambda e^{-\lambda y} \int_0^x e^{-2\lambda u} du \quad (3.46)$$

$$= (1 - e^{-\lambda y}) 2\lambda \int_0^x e^{-2\lambda u} du \quad (3.47)$$

$$= (1 - e^{-\lambda y}) (-e^{-2\lambda u}) \Big|_0^x \quad (3.48)$$

$$= (1 - e^{-\lambda y}) (1 - e^{-2\lambda x}). \quad (3.49)$$

We did a number of checks for the case  $X, Y, \text{iid.}, \sim \text{Exp}(\lambda)$ . But I have still 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 that by taking proper limits, I can obtain the same results.

More specically, I want to see intuitively<sup>2</sup> how  $X \sim \text{Geo}(\lambda/n)$  approaches  $Y \sim \text{Exp}(\lambda)$  as  $n \rightarrow \infty$ .

**Ex 3.8.** 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  $n \gg 0$ . Then take some  $x \geq 0$  and let  $i$  be such that  $x \in [i/n, (i+1)/n)$ . Show that

$$\mathbb{P}\{X/n \approx x\} \approx \frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right)^{xn}. \quad (3.50)$$

**s.3.8.** First,

$$\mathbb{P}\{X/n \approx x\} = \mathbb{P}\{X/n \in [i/n, (i+1)/n]\} = \mathbb{P}\{X \in [i, i+1]\} = pq^i \quad (3.51)$$

$$\approx pq^{nx} = \frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right)^{xn}, \quad (3.52)$$

since  $p = \lambda/n$ ,  $q = 1 - p = 1 - \lambda/n$ , and  $i = nx$ .

---

<sup>2</sup> A formal proof is too hard for this course.

From BH.A.2.5,  $\lim_{n \rightarrow \infty} (1 - \lambda/n)^n = e^{-\lambda}$ .<sup>3</sup> Next, by introducing the highly intuitive definition  $dx = \lim_{n \rightarrow \infty} 1/n$ ,<sup>4</sup> it follows from (3.50) that

$$P\{X/n \approx x\} \rightarrow \lambda e^{-\lambda x} dx = f_X(x) dx, \quad \text{as } n \rightarrow \infty. \quad (3.53)$$

If you don't like this trick with  $dx$ , it is also possible to obtain this limit with moment-generating functions.

**Ex 3.9.** 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.3.9.** If you recall the Poisson distribution, you know that  $e^\lambda = \sum_{i=0}^{\infty} \lambda^i / i!$ . In fact, this is precisely Taylor's expansion of  $e^\lambda$ .

**s.3.9.**

$$M_{X/n}(s) = E[e^{sX/n}] = \sum_i e^{si/n} p q^i = p \sum_i (q e^{s/n})^i = \frac{p}{1 - q e^{s/n}}. \quad (3.54)$$

With  $p = \lambda/n$  this becomes

$$M_{X/n}(s) = \frac{\lambda/n}{1 - (1 - \lambda/n)(1 + s/n + 1/n^2 \times (\dots))} \quad (3.55)$$

$$= \frac{\lambda/n}{\lambda/n - s/n + 1/n^2 \times (\dots)} \quad (3.56)$$

$$= \frac{\lambda}{\lambda - s + 1/n \times (\dots)} \quad (3.57)$$

$$\rightarrow \frac{\lambda}{\lambda - s}, \quad \text{as } n \rightarrow \infty, \quad (3.58)$$

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>3</sup> This is a standard limit but not entirely trivial to prove. If you like mathematics, check the neat proof in Rudin's Principles of mathematical analysis.

<sup>4</sup> In your math classes you learned that  $\lim_{n \rightarrow \infty} 1/n = 0$ . Doesn't this definition therefore imply that  $dx = 0$ ? Well, no, because  $dx$  is not a real number but an infinitesimal. Infinitesimals allow us to consider a quantity that is so small that it cannot be distinguished from 0 within the real numbers.

**Ex 3.10.** 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.3.10.** Use [2.4].

**s.3.10.**

$$E[L/n] = \frac{1}{n} E[L] = \frac{1}{n} \frac{q^2}{1-q^2} \quad (3.59)$$

$$= \frac{1}{n} \frac{(1-\lambda/n)^2}{1-(1-\lambda/n)^2} \quad (3.60)$$

$$= \frac{1}{n} \frac{1-2\lambda/n+(\lambda/n)^2}{2\lambda/n+(\lambda/n)^2} \quad (3.61)$$

$$= \frac{1-2\lambda/n+(\lambda/n)^2}{2\lambda+\lambda^2/n} \quad (3.62)$$

$$\rightarrow \frac{1}{2\lambda}. \quad (3.63)$$

**Ex 3.11.** Here is yet another check on the correctness of  $f_M(x)$ . Show that the PMF  $P\{M = k\}$  of (2.38) converges to  $f_M(x)$  of [5.2] when  $n \rightarrow \infty$ . Take  $k$  suitable.

**s.3.11.** Take  $p = \lambda/n$ ,  $q = 1 - \lambda/n$ , and  $k \approx xn$ , hence  $k/n \approx x$ . Then,

$$P\{M/n = k/n\} = 2pq^{k/n}(1 - q^{k/n}) + p^2q^{2k/n} \quad (3.64)$$

$$= 2\frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right)^{k/n} \left(1 - \left(1 - \frac{\lambda}{n}\right)^{k/n}\right) + \frac{\lambda^2}{n^2} \left(1 - \frac{\lambda}{n}\right)^{2k/n} \quad (3.65)$$

$$\rightarrow 2\lambda \, dx e^{-\lambda x} (1 - e^{-\lambda x}) + \lambda^2 \, dx^2 e^{-2\lambda x}. \quad (3.66)$$

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 other words,  $P\{X = Y\} > 0$  when  $X$  and  $Y$  are discrete, but  $P\{X = Y\} = 0$  when  $X$  and  $Y$  are continuous. So, to resolve our leading problem we should exploit that idea.

## 4 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:

$$\mathbb{E}[M] = \mathbb{E}[L] + \mathbb{E}[(M - L) I_{M > L}] = \mathbb{E}[L] + 2 \mathbb{E}[(Y - X) I_{Y > X}], \quad (4.1)$$

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 4.1.** Show that

$$2 \mathbb{E}[(Y - X) I_{Y > X}] = \frac{2q}{1 - q^2}. \quad (4.2)$$

Combine this with the expression for  $\mathbb{E}[L]$  of [2.4] to obtain (2.32) for  $\mathbb{E}[M] = 2 \mathbb{E}[X] - \mathbb{E}[L]$ , thereby verifying the correctness of (4.1).

**s.4.1.**

$$\mathbb{E}[(Y - X) I_{Y > X}] = p^2 \sum_{ij} (j - i) I_{j > i} q^i q^j \quad (4.3)$$

$$= p^2 \sum_i q^i \sum_{j=i+1}^{\infty} (j - i) q^j \quad (4.4)$$

$$= p^2 \sum_i q^i q^i \sum_{k=1}^{\infty} k q^k, \quad k = j - i \quad (4.5)$$

$$= p \sum_i q^{2i} \mathbb{E}[X] \quad (4.6)$$

$$= \frac{p}{1 - q^2} \mathbb{E}[X] \quad (4.7)$$

$$= \frac{p}{1 - q^2} \frac{q}{p} \quad (4.8)$$

$$= \frac{q}{1 - q^2}. \quad (4.9)$$

With this,

$$\mathbb{E}[L] + 2 \mathbb{E}[(Y - X) I_{Y > X}] = \frac{q^2}{1 - q^2} + \frac{2q}{1 - q^2} = \frac{q}{1 - q} \frac{q + 2}{1 + q}, \quad (4.10)$$

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

While (4.1) is correct, I am still not happy with the second part of (4.1) (I find this hard/unintuitive to interpret). Restarting again from scratch, here is another attempt to rewrite  $E[M]$ :

$$E[M] = E[L] + E[Z I_{M>L}], \quad (4.11)$$

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 4.2.** Show that

$$E[Z I_{M>L}] = \frac{2q}{1-q^2}, \quad (4.12)$$

i.e., the same as (4.2), hence (4.11) is correct.

**h.4.2.** Observe that  $Z$  is independent from  $X$  and  $Y$ , hence from  $M$  and  $L$

**s.4.2.** With the hint:

$$E[Z I_{M>L}] = E[Z] E[I_{M>L}] = E[Z] P\{M > L\} = \frac{1}{p} \frac{2pq}{1-q^2} = \frac{2q}{1-q^2}, \quad (4.13)$$

where, from the book,  $E[Z] = 1/p$ , and  $P\{M > L\}$  follows from [2.11].



I am nearly happy, but I want to see that (4.11), which is correct for discrete r.v.s, has also the correct limiting behavior, similar to [3.10].

**Ex 4.3.** 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.4.3.** By independence,

$$E[Z/n I_{M>L}] = E[Z/n] P\{M > L\}. \quad (4.14)$$

Then,

$$P\{M > L\} = \frac{2pq}{1-q^2} = \frac{2\lambda/n(1-\lambda/n)}{1-(1-\lambda/n)^2} = \frac{2\lambda/n(1-\lambda/n)}{2\lambda/n - \lambda^2/n^2} = \frac{2(1-\lambda/n)}{2-\lambda/n} \rightarrow 1, \quad (4.15)$$

and

$$E[Z/n] = \frac{1}{n} E[Z] = \frac{1}{n} \frac{1}{p} = \frac{1}{n\lambda/n} = 1/\lambda. \quad (4.16)$$

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 (4.11) a nice and revealing equation.

## 5 WORKING WITH DENSITIES

You should notice that working with densities is often easier than working with the fundamental bridge because you can work with equalities rather than inequalities. However, don't forget to properly scale lengths and areas; use the Jacobian!

**Ex 5.1.** Derive the density  $f_L(u)$  from the joint density  $f_{X,Y}$ .

**s.5.1.** Hence, with  $u = x$ ,

$$\mathbf{P}\{L \approx u\} = 2 \mathbf{P}\{X \approx x, Y > x\}. \quad (5.1)$$

Hence, as  $u = x \implies du/dx = 1$ ,

$$f_L(u) du = 2f_X(x)G_Y(x) dx \implies f_L(u) = 2f_X(u)G_Y(u).$$

**Ex 5.2.** Derive the density  $f_M(v)$  from the joint density  $f_{X,Y}$ .

**s.5.2.** With  $v = y$

$$\mathbf{P}\{M \approx v\} = 2 \mathbf{P}\{X \leq v, Y \approx v\}. \quad (5.2)$$

As  $v = y \implies dv/dy = 1$ ,

$$f_M(v) dv = 2F_X(y)f_Y(y) dy \implies f_M(v) = 2F_X(v)f_Y(v). \quad (5.3)$$

**Ex 5.3.** Find a function  $g$  that maps  $X, Y$  to  $L, M$ . Then compute  $f_{L,M}$  by using change of variables.

**s.5.3.** We need  $g$  first, and then we compute the Jacobian.

$$(u, v) = g(x, y) = (\min\{x, y\}, \max\{x, y\}). \quad (5.4)$$

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

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} I_{x \leq y} & I_{y < x} \\ I_{x > y} & I_{y \geq x} \end{vmatrix} = |I_{x \leq y} - I_{x > y}| = 1. \quad (5.5)$$

Now with the change of variable formula in which I drop the Jacobian right away as it is 1 and substitute  $u = x$  and  $y = v$ ,

$$f_{L,M}(u, v) I_{u \leq v} = 2f_{X,Y}(u, v) I_{u \leq v}. \quad (5.6)$$

**Ex 5.4.** 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.5.4.** 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

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} I_{x \leq y} & 0 \\ -I_{x \leq y} & I_{x \leq y} \end{vmatrix} = I_{x \leq y}. \quad (5.7)$$

And with this,

$$f_{L,M-L}(u, v) = f_{L,M}(u, u + v) I_{u \leq u+v} \quad (5.8)$$

$$= f_{L,M}(u, u + v) \quad (5.9)$$

$$= 2f_X(u)f_Y(u + v) \quad (5.10)$$

$$= 2\lambda^2 e^{-2\lambda u} e^{-\lambda v} \quad (5.11)$$

$$= f_L(u)f_Y(v). \quad (5.12)$$

Now two exercises just to practice with the method.

**Ex 5.5.** Compute  $f_{L^2}(l)$ , where  $L = \min\{X, Y\}$  and  $X, Y$  iid  $\sim \text{Exp}(\lambda)$ .

**s.5.5.** With  $l = x^2, x \leq y$

$$f_{L^2}(l) \, dl = 2f_X(x)G_Y(x) \, dx.$$

As  $dl/dx = 2x = 2\sqrt{l}$ ,

$$f_{L^2}(l) = 2f_X(\sqrt{l})G_Y(\sqrt{l}) \, dx/dl = f_X(\sqrt{l})G_Y(\sqrt{l})/\sqrt{l}.$$

**Ex 5.6.** We have three r.v.s,  $X, Y, Z$ , iid. Let  $L = \min\{X, Y, Z\}$  and  $M = \max\{X, Y, Z\}$ . Find  $f_{L,M}(u, v)$ .

**s.5.6.** There are  $3!$  ways to order  $X, Y, Z$ . Therefore, with  $u = x \leq y \leq z = v$ ,

$$f_{L,M}(u, v) \, du \, dv = 6f_X(x)(F_Y(z) - F_Y(x))f_Z(z) \, dx \, dz$$

Since  $\partial(u, v)/\partial(x, z) = 1$ ,

$$f_{L,M}(u, v) = 6f_X(u)(F_Y(v) - F_Y(u))f_Z(v) I_{u \leq v}.$$

Here is a neat method with CDFs.

$$\begin{aligned} F_{L,M}(u, v) &= \mathbf{P}\{L \leq u, M \leq v\} \\ &= \mathbf{P}\{M \leq v\} - \mathbf{P}\{L \geq u, M \leq v\} \\ &= (F_X(v))^3 - (F_X(v) - F_X(u))^3. \end{aligned}$$

Taking  $\partial_v \partial_u$  gives the same result as above.