

MATH4406 HW1

Patrick Laub

August 2014

Exercise 1.

Part 1.

$$\begin{aligned}\emptyset \in \mathcal{F}, & \quad \text{by property 1 of a sigma algebra,} \\ \Rightarrow \emptyset^c \in \mathcal{F}, & \quad \text{by property 2 of a sigma algebra,} \\ \Rightarrow \Omega \in \mathcal{F}, & \quad \text{since } \emptyset^c = \Omega, \\ & \quad \therefore \Omega \in \mathcal{F}.\end{aligned}$$

Part 2.

Say $A_1, A_2, \dots \in \mathcal{F}$ then

$$\begin{aligned}A_1^c, A_2^c, \dots \in \mathcal{F}, & \quad \text{by property 2 of a sigma algebra,} \\ \Rightarrow \cup_i A_i^c \in \mathcal{F}, & \quad \text{by property 3 of a sigma algebra,} \\ \Rightarrow (\cap_i A_i)^c \in \mathcal{F}, & \quad \text{by De Morgan's rules,} \\ \Rightarrow \cap_i A_i \in \mathcal{F}, & \quad \text{by property 2 of a sigma algebra,} \\ & \quad \therefore A_1, A_2, \dots \in \mathcal{F} \Rightarrow \cap_i A_i \in \mathcal{F}.\end{aligned}$$

□

Exercise 2.

Part 1.

For any $A \in \mathcal{F}$ then

$$\mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1$$

by property 2 of a probability measure, and by property 3

$$\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) = 1$$

as A and A^c are disjoint. Rearranging gives

$$\therefore \mathbb{P}(A^c) = 1 - \mathbb{P}(A).$$

Part 2.

As $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ from previous proof and $\mathbb{P}(\Omega) = 1$ by property 2 of a probability measure then $\mathbb{P}(\Omega^c) = 1 - \mathbb{P}(\Omega) = 1 - 1 = 0$ and as $\Omega^c = \emptyset$ then

$$\therefore \mathbb{P}(\emptyset) = 0.$$

Part 3.

Say $A_1, A_2 \in \mathcal{F}$ but not necessarily disjoint. Note that $A_1 \cap A_2^c$ and A_2 are disjoint so

$$\mathbb{P}((A_1 \cap A_2^c) \cup A_2) = \mathbb{P}(A_1 \cap A_2^c) + \mathbb{P}(A_2).$$

Also see that $(A_1 \cap A_2^c) \cup A_2 = (A_1 \cup A_2) \cap \Omega = A_1 \cup A_2$ by the distributive law, so

$$\Rightarrow \mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 \cap A_2^c) + \mathbb{P}(A_2). \quad (1)$$

Now notice that $A_1 \cap A_2^c$ and $A_1 \cap A_2$ are disjoint so

$$\mathbb{P}((A_1 \cap A_2^c) \cup (A_1 \cap A_2)) = \mathbb{P}(A_1 \cap A_2^c) + \mathbb{P}(A_1 \cap A_2).$$

As $(A_1 \cap A_2^c) \cup (A_1 \cap A_2) = A_1$ then

$$\Rightarrow \mathbb{P}(A_1) = \mathbb{P}(A_1 \cap A_2^c) + \mathbb{P}(A_1 \cap A_2). \quad (2)$$

Rearranging (1) and (2) for $\mathbb{P}(A_1 \cap A_2^c)$ gives

$$\mathbb{P}(A_1 \cup A_2) - \mathbb{P}(A_2) = \mathbb{P}(A_1) - \mathbb{P}(A_1 \cap A_2)$$

$$\therefore \mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

□

Exercise 4.

Assume that $\mathbb{P}(A|B) = \mathbb{P}(A)$, and combine this with the definition of conditional probability

$$\Rightarrow \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

$$\Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

so this means that $\mathbb{P}(A|B) = \mathbb{P}(A)$ implies A and B are independent. The same holds in reverse by the above argument.

$$\therefore \mathbb{P}(A|B) = \mathbb{P}(A) \Leftrightarrow A, B \text{ independent.}$$

□

Exercise 5.

Say that R is the result of the die roll. The question asks for

$$\begin{aligned} \mathbb{P}(R = 6 | R \text{ is even}) &= \mathbb{P}(R = 6 | R \in \{2, 4, 6\}) \\ &= \frac{\mathbb{P}(R \in \{6\} \cap R \in \{2, 4, 6\})}{\mathbb{P}(R \in \{2, 4, 6\})} \\ &= \frac{\mathbb{P}(R = 6)}{\mathbb{P}(R \in \{2, 4, 6\})} \\ &= \frac{1/6}{3/6} \\ &= \frac{1}{3} \end{aligned}$$

using the definition of conditional probability, and the equilikely principle.

$$\therefore \mathbb{P}(R = 6 | R \text{ is even}) = \frac{1}{3}.$$

□

Exercise 6.

Part 1.

If $\mathbb{P}(B) > 0$ then $\mathbb{P}(A|B)$ is defined to be

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

and rearranging implies that

$$\therefore \mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B).$$

Part 2.

If $\{B_i\}$ partition Ω then note that

$$\bigcup_i (A \cap B_i) = A \cap \bigcup_i B_i = A \cap \Omega = A$$

by the distributive law. Also note that the events $\{A \cap B_i\}$ are mutually disjoint as the $\{B_i\}$ are mutually disjoint. Therefore

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_i (A \cap B_i)\right) = \sum_i \mathbb{P}(A \cap B_i)$$

and by the previous question we have

$$\begin{aligned} \sum_i \mathbb{P}(A \cap B_i) &= \sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i) \\ \therefore \mathbb{P}(A) &= \sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i) = \sum_i \mathbb{P}(A \cap B_i). \end{aligned}$$

Part 3.

$$\begin{aligned} \mathbb{P}(B_i|A) &= \frac{\mathbb{P}(B_i \cap A)}{\mathbb{P}(A)} && \text{by def. of conditional prob.} \\ &= \frac{\mathbb{P}(A \cap B_i)}{\sum_i \mathbb{P}(A \cap B_i)} && \text{by commutativity and L.O.T.P} \\ &= \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_i \mathbb{P}(A \cap B_i)} && \text{by multiplication rule.} \end{aligned}$$

$$\therefore \mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_i \mathbb{P}(A \cap B_i)}.$$

□

Exercise 7.

When originally guessing you have no information about the location of the prize, and guess correctly with probability $1/3$. When one of the doors is opened then now you have some information (i.e. the prize isn't behind that door). If you stick with your current guess then you are not *using* the information you were just given, and win with probability $1/3$. However, if you swap to the third door then you're effectively updating your guess to the reduced situation with only two doors.

□

Exercise 8.

Assume that $X \in \mathbb{R}$ or in some other set that doesn't include $\pm\infty$ (like, for example, the extended reals). So by assumption $-\infty < X < \infty$.

Part 1.

$$\begin{aligned} \lim_{x \rightarrow -\infty} F_X(x) &= \lim_{x \rightarrow -\infty} \mathbb{P}(X \leq x) \\ &= \mathbb{P}(X \leq -\infty) \\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq -\infty\}) \\ &= \mathbb{P}(\emptyset) \\ &= 0. \end{aligned}$$

Part 2.

$$\begin{aligned} \lim_{x \rightarrow \infty} F_X(x) &= \lim_{x \rightarrow \infty} \mathbb{P}(X \leq x) \\ &= \mathbb{P}(X \leq \infty) \\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq \infty\}) \\ &= \mathbb{P}(\Omega) \\ &= 1. \end{aligned}$$

Part 3.

Consider $\forall x, y$ s.t. $-\infty < x_1 < x_2 < \infty$ the values

$$F_X(x_1) = \mathbb{P}(\{X \in (-\infty, x_1]\})$$

and

$$\begin{aligned} F_X(x_2) &= \mathbb{P}(\{X \in (-\infty, x_2]\}) \\ &= \mathbb{P}(\{X \in (-\infty, x_1]\} \cup \{X \in (x_1, x_2]\}) \\ &= \mathbb{P}(\{X \in (-\infty, x_1]\}) + \mathbb{P}(\{X \in (x_1, x_2]\}) && \text{as events disjoint} \\ &= F_X(x_1) + \mathbb{P}(\{X \in (x_1, x_2]\}) \\ &\geq F_X(x_1) && \text{as } \mathbb{P}(\cdot) \geq 0 \end{aligned}$$

$\therefore F_X(\cdot)$ is non-decreasing.

□

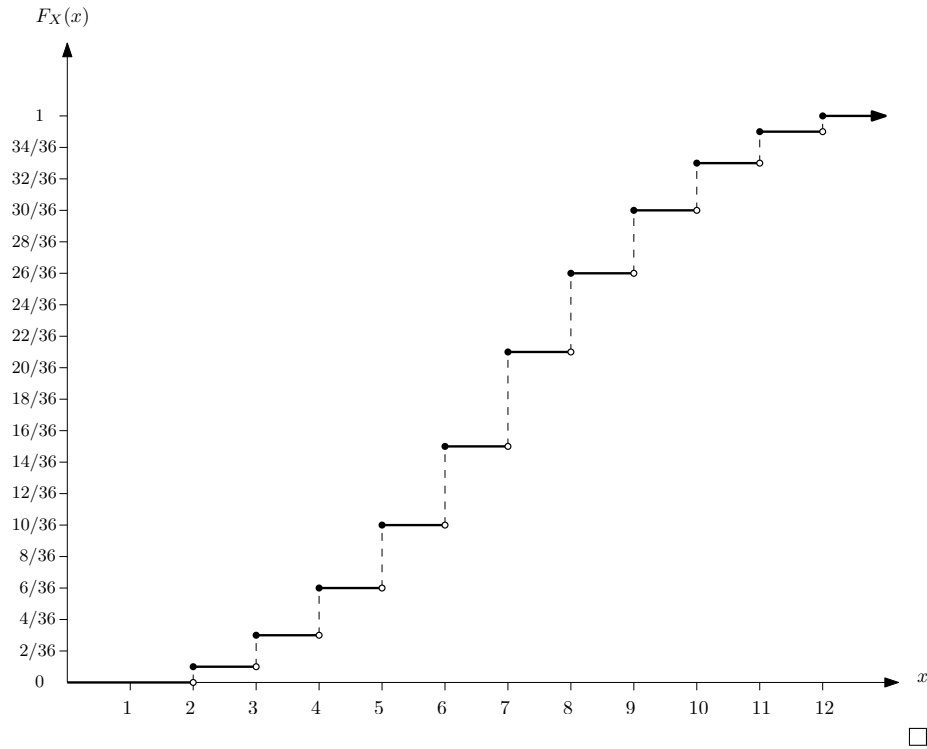
Exercise 9.

Say X is the sum of the two fair die. The pmf for X is

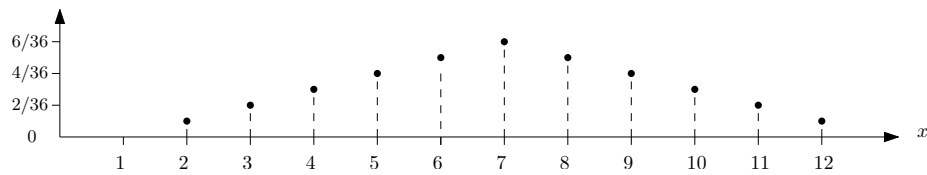
$$p_X(x) = \begin{cases} 1/36, & x = 2 \\ 2/36, & x = 3 \\ 3/36, & x = 4 \\ 4/36, & x = 5 \\ 5/36, & x = 6 \\ 6/36, & x = 7 \\ 5/36, & x = 8 \\ 4/36, & x = 9 \\ 3/36, & x = 10 \\ 2/36, & x = 11 \\ 1/36, & x = 12 \\ 0, & \text{otherwise} \end{cases}$$

so therefore the cdf is

$$F_X(x) = \begin{cases} 0, & x < 2 \\ 1/36, & 2 \leq x < 3 \\ 3/36, & 3 \leq x < 4 \\ 6/36, & 4 \leq x < 5 \\ 10/36, & 5 \leq x < 6 \\ 15/36, & 6 \leq x < 7 \\ 21/36, & 7 \leq x < 8 \\ 26/36, & 8 \leq x < 9 \\ 30/36, & 9 \leq x < 10 \\ 33/36, & 10 \leq x < 11 \\ 35/36, & 11 \leq x < 12 \\ 1, & 12 \leq x \end{cases}$$



Exercise 10.
 $p_X(x)$



The jumps in the cdf specify the heights of the pmf. It is as if the cdf is pushed down towards the x axis. □

Exercise 11.

Say X is the sum of the two fair die. Using the pmf specified earlier then

$$\begin{aligned}
 \mathbb{E}[X] &= \sum_{x=2}^{12} xp_X(x) \\
 &= 2 \times 1/36 + 3 \times 2/36 + 4 \times 3/36 + 5 \times 4/36 + \\
 &\quad 6 \times 5/36 + 7 \times 6/36 + 8 \times 5/36 + 9 \times 4/36 + \\
 &\quad 10 \times 3/36 + 11 \times 2/36 + 12 \times 1/36 \\
 &= 7.
 \end{aligned}$$

$$\therefore \mathbb{E}[X] = 7.$$

□

Exercise 12.If $X \in \mathbb{N}_0$ then

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{\infty} k\mathbb{P}(X = k) \\
&= 0\mathbb{P}(X = 0) + 1\mathbb{P}(X = 1) + 2\mathbb{P}(X = 2) + \dots \\
&= \mathbb{P}(X = 1) + \\
&\quad \mathbb{P}(X = 2) + \mathbb{P}(X = 2) + \\
&\quad \mathbb{P}(X = 3) + \mathbb{P}(X = 3) + \mathbb{P}(X = 3) + \\
&\quad \mathbb{P}(X = 4) + \mathbb{P}(X = 4) + \mathbb{P}(X = 4) + \mathbb{P}(X = 4) + \dots \\
&= \mathbb{P}(X > 0) + \mathbb{P}(X > 1) + \mathbb{P}(X > 2) + \mathbb{P}(X > 3) + \dots \\
&= \sum_{k=0}^{\infty} \mathbb{P}(X > k) \\
&= \sum_{k=0}^{\infty} \bar{F}_X(k).
\end{aligned}$$

$$\therefore X \in \mathbb{N}_0 \Rightarrow \mathbb{E}[X] = \sum_{k=0}^{\infty} \bar{F}_X(k).$$

□

Exercise 13.*Part 1.*

$$\begin{aligned}
\mathbb{E}[cX] &= \int_{\omega \in \Omega} cX(\omega) \, d\omega \\
&= c \int_{\omega \in \Omega} X(\omega) \, d\omega \\
&= c\mathbb{E}[X].
\end{aligned}$$

$$\therefore \text{For deterministic } c \Rightarrow \mathbb{E}[cX] = c\mathbb{E}[X].$$

Part 2.

$$\begin{aligned}
\mathbb{E}[X + Y] &= \int_{\omega \in \Omega} X(\omega) + Y(\omega) \, d\omega \\
&= \int_{\omega \in \Omega} X(\omega) \, d\omega + \int_{\omega \in \Omega} Y(\omega) \, d\omega \\
&= \mathbb{E}[X] + \mathbb{E}[Y].
\end{aligned}$$

$$\therefore \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

□

Exercise 14.For some constants c_1, c_2 then

$$\begin{aligned}
\text{Var}[c_1X + c_2] &= \mathbb{E}[(c_1X + c_2)^2] - \mathbb{E}[c_1X + c_2]^2 \\
&= \mathbb{E}[c_1^2X^2 + 2c_1c_2X + c_2^2] - (c_1\mathbb{E}[X] + c_2)^2 \\
&= c_1^2\mathbb{E}[X^2] + 2c_1c_2\mathbb{E}[X] + c_2^2 - c_1^2\mathbb{E}[X]^2 - 2c_1c_2\mathbb{E}[X] - c_2^2 \\
&= c_1^2\mathbb{E}[X^2] - c_1^2\mathbb{E}[X]^2 \\
&= c_1^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) \\
&= c_1^2\text{Var}[X].
\end{aligned}$$

$$\therefore \text{Var}[c_1X + c_2] = c_1^2\text{Var}[X].$$

□

Exercise 15.Assume that $\text{Var}[X] = 0$, so

$$\text{Var}[X] = \sum_{k=-\infty}^{\infty} (k - \mathbb{E}[X])^2 p_X(k) = 0.$$

Say that $p_X(k) > 0$ for some set $k \in S$, the support of X . So

$$\text{Var}[X] = \sum_{k \in S} (k - \mathbb{E}[X])^2 p_X(k) = 0.$$

As each summand is non-negative then the summands must each be zero.

$$\Rightarrow \forall k \in S : (k - \mathbb{E}[X])^2 p_X(k) = 0$$

Then, since $p_X(k) \neq 0$ for $k \in S$ then

$$\Rightarrow \forall k \in S : (k - \mathbb{E}[X])^2 = 0.$$

The only solution to this set of equalities for $|S| = 1$ and $k = \mathbb{E}[X]$, i.e. $S = \{\mathbb{E}[X]\}$. As $p_X(k) > 0$ for $k \in S$, and as $\sum_k p_X(k) = 1$ then therefore $p_X(\mathbb{E}[X]) = 1$.

$$\therefore \text{Var}[X] = 0 \Rightarrow \text{the support of } X \text{ only contains a single value.}$$

□

Exercise 17.

By applying the binomial theorem one can see that

$$\begin{aligned}
\sum_{i=1}^n \mathbb{P}(X = i) &= \sum_{i=1}^n \binom{n}{i} p^i (1-p)^{n-i} \\
&= (p + 1 - p)^n \\
&= 1.
\end{aligned}$$

$$\therefore \sum_{i=1}^n \mathbb{P}(X = i) = 1.$$

□

Exercise 18.

Part 1.

Given $X \sim \text{Bin}(n, p)$ then one could split this up to be

$$X = \sum_{i=1}^n X_i$$

where $X_i \sim \text{Ber}(p)$ with $\mathbb{E}[X_i] = p$. Next calculate

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{i=1}^n X_i\right] \\ &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n] \\ &= n\mathbb{E}[X_1] \quad \text{as identically distributed} \\ &= np. \end{aligned}$$

$$\therefore X \sim \text{Bin}(n, p) \Rightarrow \mathbb{E}[X] = np.$$

Part 2.

If $X \sim \text{Bin}(n, p)$ then one can see this as the number of “successes” out of n independent trials each with probability p of success. Equivalently, one can look at $n - X$ the number of failures out of n independent trials each with probability $1 - p$ of failing, and conclude that

$$\therefore n - X \sim \text{Bin}(n, 1 - p).$$

□

Exercise 19.

Say that $X \sim \text{Bin}(20, 0.25)$ is the number of answers guessed correctly out of the multiple choice exam.

$$\begin{aligned} \mathbb{P}(X \geq 10) &= \sum_{i=10}^{20} \mathbb{P}(X = i) \\ &= \sum_{i=10}^{20} \binom{20}{i} 0.25^i 0.75^{20-i} \\ &= 0.00992228 + 0.00300675 + 0.000751688 + 0.000154192 + \\ &\quad 0.0000256987 + 3.4265\text{e-}6 + 3.56927\text{e-}7 + 2.79942\text{e-}8 + \\ &\quad 1.55524\text{e-}9 + 5.45697\text{e-}11 + 9.09495\text{e-}13 \\ &= 0.01386442. \end{aligned}$$

$$\therefore \mathbb{P}(\text{Guessing 10 or more correctly}) = 0.01386442.$$

□

Exercise 20.

Say that $X \sim \text{Geo}(p)$, then

$$\begin{aligned} \mathbb{P}(X = k) &= \mathbb{P}(\inf\{i \in \{1, 2, 3, \dots\} | b_i^\omega = 1\} = k) \\ &= \mathbb{P}(b_1^\omega = 0, b_2^\omega = 0, \dots, b_{k-1}^\omega = 0, b_k^\omega = 1) \\ &= \mathbb{P}(b_1^\omega = 0)\mathbb{P}(b_2^\omega = 0)\dots\mathbb{P}(b_{k-1}^\omega = 0)\mathbb{P}(b_k^\omega = 1) \\ &= (1-p)(1-p)\dots(1-p)p \\ &= (1-p)^{k-1}p. \\ &\Rightarrow \mathbb{P}(X = k) = (1-p)^{k-1}p. \end{aligned}$$

Also,

$$\begin{aligned} \sum_{k=1}^{\infty} \mathbb{P}(X = k) &= \sum_{k=1}^{\infty} (1-p)^{k-1}p \\ &= p \sum_{k=1}^{\infty} (1-p)^{k-1} \\ &= p \sum_{k=0}^{\infty} (1-p)^k \\ &= p \frac{1}{1-(1-p)} \quad \text{sum of geometric series} \\ &= \frac{p}{p} \\ &= 1. \end{aligned}$$

$$\therefore \mathbb{P}(X = k) = (1-p)^{k-1}p \quad \text{and} \quad \sum_{k=1}^{\infty} \mathbb{P}(X = k) = 1.$$

□

Exercise 21.

Assuming the event of getting a flat tyre is iid Bernoulli over the consecutive bicycle rides with $p = 0.01$. So the chance of 20 of these trials failing consecutively is $(1-p)^{20} = 0.99^{20} \approx 0.8179$.

$$\therefore \mathbb{P}(\text{20 consecutive rides without a flat}) = 0.99^{20} \approx 0.8179.$$

□

Exercise 22.

This random variable is simply one less than the equivalent geometric distribution originally defined, i.e. $X \sim \text{Geo}(p)$ and $X^* = X - 1 \sim \text{Geo}_0(p)$.

The original distribution has support $S = \{1, 2, \dots\}$ and for $k \in S$ then $p_X(k) = (1 - p)^{k-1}p$.

Therefore the new shifted distribution has support $S^* = \{0, 1, \dots\}$ and for $k \in S^*$ then $p_{X^*}(k) = (1 - p)^k p$.

$\therefore X^* \sim \text{Geo}_0(p)$ has support \mathbb{N}_0 and for $k \in \mathbb{N}_0$ then $p_{X^*}(k) = (1 - p)^k p$.

□

Exercise 26.

Note the Taylor series expansion

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}.$$

Say that $X \sim \text{Poi}(\lambda)$, then

$$\begin{aligned} \mathbb{E}[X] &= \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= \lambda e^{-\lambda} e^\lambda \\ &= \lambda. \end{aligned}$$

To calculate variance start with

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} \\ &= e^{-\lambda} \left[\sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \right] \\ &= e^{-\lambda} \left[\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \right] \\ &= e^{-\lambda} \left[\lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right] \\ &= e^{-\lambda} \left[\lambda^2 e^{\lambda} + \lambda e^{\lambda} \right] \\ &= \lambda^2 + \lambda.\end{aligned}$$

then combine these results to get

$$\begin{aligned}\text{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \lambda^2 + \lambda - (\lambda)^2 \\ &= \lambda.\end{aligned}$$

$$\therefore \mathbb{E}[X] = \lambda \text{ and } \text{Var}[X] = \lambda.$$

□

Exercise 28.

Noting that $\sum_n \frac{1}{n} \rightarrow \infty$ by the p -series test

$$\begin{aligned}
 \mathbb{E}[X] &= \sum_{k=-\infty}^{\infty} k\mathbb{P}(X = k) \\
 &= \sum_{k=1}^{\infty} \frac{k}{k(k+1)} \\
 &= \sum_{k=1}^{\infty} \frac{1}{k+1} \\
 &= \sum_{k=2}^{\infty} \frac{1}{k} \\
 &= \infty. \\
 \therefore \mathbb{E}[X] &= \infty.
 \end{aligned}$$

□

Exercise 29.

Since $\forall \ell \in \mathbb{Z}$ then the events $\{Y = \ell\}$ are disjoint and partition Ω . Hence we can use the law of total probability conditioning on these events:

$$\begin{aligned}
 p_X(k) &= \mathbb{P}(X = k) \\
 &= \sum_{\ell=-\infty}^{\infty} \mathbb{P}(X = k|Y = \ell)\mathbb{P}(Y = \ell) \\
 &= \sum_{\ell=-\infty}^{\infty} \mathbb{P}(X = k, Y = \ell) \\
 &= \sum_{\ell=-\infty}^{\infty} \mathbb{P}(X = k, Y = \ell) \\
 &= \sum_{\ell=-\infty}^{\infty} p_{X,Y}(k, \ell). \\
 \therefore p_X(k) &= \sum_{\ell=-\infty}^{\infty} p_{X,Y}(k, \ell).
 \end{aligned}$$

□

Exercise 30.

Using the facts that

$$\sum_k p_{X,Y}(k, l) = p_Y(l), \quad \sum_l p_{X,Y}(k, l) = p_X(k), \quad \sum_k \sum_l p_{X,Y}(k, l) = 1$$

then

$$\begin{aligned}
\text{Cov}[X, Y] &= \mathbb{E}[(x - \mathbb{E}[X])(y - \mathbb{E}[Y])] \\
&= \sum_k \sum_l (k - \mathbb{E}[X])(l - \mathbb{E}[Y])p_{X,Y}(k, l) \\
&= \sum_k \sum_l (kl - k\mathbb{E}[Y] - l\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y])p_{X,Y}(k, l) \\
&= \sum_k \sum_l kl p_{X,Y}(k, l) - \sum_k \sum_l k\mathbb{E}[Y]p_{X,Y}(k, l) - \\
&\quad \sum_k \sum_l l\mathbb{E}[X]p_{X,Y}(k, l) + \sum_k \sum_l \mathbb{E}[X]\mathbb{E}[Y]p_{X,Y}(k, l) \\
&= \mathbb{E}[XY] - \mathbb{E}[Y] \sum_k k \sum_l p_{X,Y}(k, l) - \\
&\quad \mathbb{E}[X] \sum_l l \sum_k p_{X,Y}(k, l) + \mathbb{E}[X]\mathbb{E}[Y] \sum_k \sum_l p_{X,Y}(k, l) \\
&= \mathbb{E}[XY] - \mathbb{E}[Y] \sum_k kp_X(k) - \mathbb{E}[X] \sum_l lp_Y(l) + \mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

$$\therefore \text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

□

Exercise 31.

If X, Y are independent then $p_{X,Y}(k, l) = p_X(k)p_Y(l)$ so

$$\begin{aligned}
\mathbb{E}[XY] &= \sum_k \sum_l kl p_{X,Y}(k, l) \\
&= \sum_k \sum_l kl p_X(k)p_Y(l) \\
&= \left(\sum_k kp_X(k) \right) \left(\sum_l lp_Y(l) \right) \\
&= \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

And note that

$$\text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] = 0$$

$$\therefore X, Y \text{ independent} \Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \text{ and } \text{Cov}[X, Y] = 0.$$

□

Exercise 32.*Part 1.*

Take

$$p_{X,Y}(x,y) = \begin{cases} 0.1, & x = 1, y = 1 \\ 0.1, & x = 1, y = 2 \\ 0.1, & x = 2, y = 1 \\ 0.1, & x = 2, y = 2 \\ 0.1, & x = 3, y = 1 \\ 0.5, & x = 3, y = 2 \\ 0, & \text{otherwise} \end{cases}.$$

So therefore

$$p_X(x) = \begin{cases} 0.2, & x = 1 \\ 0.2, & x = 2 \\ 0.6, & x = 3 \\ 0, & \text{otherwise} \end{cases}, \quad p_Y(y) = \begin{cases} 0.3, & y = 1 \\ 0.7, & y = 2 \\ 0, & \text{otherwise} \end{cases}.$$

So to calculate the expectations

$$\Rightarrow \mathbb{E}[X] = 1 \times 0.2 + 2 \times 0.2 + 3 \times 0.6 = 2.4,$$

$$\Rightarrow \mathbb{E}[Y] = 1 \times 0.3 + 2 \times 0.7 = 1.7,$$

$$\begin{aligned} \mathbb{E}[XY] &= 1 \times 1 \times 0.1 + 1 \times 2 \times 0.1 + \\ &\quad 2 \times 1 \times 0.1 + 2 \times 2 \times 0.1 + \\ &\quad 3 \times 1 \times 0.1 + 3 \times 2 \times 0.5 \\ &= 4.2 \end{aligned}$$

$$\Rightarrow \text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 4.2 - 2.4 \times 1.7 = 0.12$$

$$\therefore \text{Cov}[X, Y] \neq 0$$

Part 2.

Take

$$p_{X,Y}(x,y) = \begin{cases} 0.1, & x = 1, y = 1 \\ 0.1, & x = 1, y = 2 \\ 0.5, & x = 2, y = 1 \\ 0.1, & x = 2, y = 2 \\ 0.1, & x = 3, y = 1 \\ 0.1, & x = 3, y = 2 \\ 0, & \text{otherwise} \end{cases}.$$

So therefore

$$p_X(x) = \begin{cases} 0.2, & x = 1 \\ 0.6, & x = 2 \\ 0.2, & x = 3 \\ 0, & \text{otherwise} \end{cases}, \quad p_Y(y) = \begin{cases} 0.7, & y = 1 \\ 0.3, & y = 2 \\ 0, & \text{otherwise} \end{cases}.$$

The variables X and Y are clearly not independent, e.g.

$$p_{X,Y}(1, 1) = 0.1 \neq 0.14 = 0.2 \times 0.7 = p_X(1)p_Y(1).$$

So to calculate the expectations

$$\Rightarrow \mathbb{E}[X] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2,$$

$$\Rightarrow \mathbb{E}[Y] = 1 \times 0.7 + 2 \times 0.3 = 1.3,$$

$$\begin{aligned} \mathbb{E}[XY] &= 1 \times 1 \times 0.1 + 1 \times 2 \times 0.1 + \\ &\quad 2 \times 1 \times 0.5 + 2 \times 2 \times 0.1 + \\ &\quad 3 \times 1 \times 0.1 + 3 \times 2 \times 0.1 \\ &= 2.6 \end{aligned}$$

$$\Rightarrow \text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 2.6 - 2 \times 1.3 = 0$$

$\therefore X$ and Y dependent and $\text{Cov}[X, Y] = 0$.

□

Exercise 33.

Given that $\mathbb{P}(Y = \ell) > 0$ then the conditional probability $p_{X|Y=\ell}(\cdot, \ell)$ is well defined. As it equals

$$p_{X|Y=\ell}(\cdot, \ell) = \frac{\mathbb{P}(X = \cdot, Y = \ell)}{\mathbb{P}(Y = \ell)}$$

then it is always non-negative ($\mathbb{P}(\cdot) > 0$ and the ratio of non-negative numbers is also non-negative). Also,

$$\begin{aligned} \sum_k p_{X|Y=\ell}(k, \ell) &= \sum_k \frac{\mathbb{P}(X = k, Y = \ell)}{\mathbb{P}(Y = \ell)} \\ &= \frac{\sum_k \mathbb{P}(X = k, Y = \ell)}{\mathbb{P}(Y = \ell)} \\ &= \frac{\mathbb{P}(Y = \ell)}{\mathbb{P}(Y = \ell)} \\ &= 1. \end{aligned}$$

$\therefore p_{X|Y=\ell}(\cdot, \ell) \geq 0$ and $\sum_k p_{X|Y=\ell}(k, \ell) = 1$ so it is a probability mass function.

□

Exercise 34.

Using the fact that independence implies $\mathbb{P}(X = \cdot, Y = \ell) = \mathbb{P}(X = \cdot)\mathbb{P}(Y = \ell)$ then

$$\begin{aligned} p_{X|Y=\ell}(\cdot, \ell) &= \frac{\mathbb{P}(X = \cdot, Y = \ell)}{\mathbb{P}(Y = \ell)} \\ &= \frac{\mathbb{P}(X = \cdot)\mathbb{P}(Y = \ell)}{\mathbb{P}(Y = \ell)} \\ &= \mathbb{P}(X = \cdot) \\ &= p_X(\cdot). \end{aligned}$$

\therefore If X, Y are independent $\Rightarrow p_{X|Y=\ell}(\cdot, \ell) = p_X(\cdot)$.

□

Exercise 38.

Firstly note that

$$p_{X|Y}(k, \ell) = \frac{p_{X,Y}(k, \ell)}{p_Y(\ell)} \Rightarrow p_{X,Y}(k, \ell) = p_{X|Y}(k, \ell)p_Y(\ell).$$

Also remember that

$$\mathbb{E}[g(Y)] = \sum_{\ell} g(\ell)p_Y(\ell)$$

for some function g of Y . Then if we consider

$$g(Y) = \mathbb{E}[h(X)|Y] = \sum_k h(k)p_{X|Y}(k, Y)$$

then

$$\begin{aligned}\mathbb{E}[g(Y)] &= \mathbb{E}[\mathbb{E}[h(X)|Y]] \\ &= \sum_{\ell} g(\ell) p_Y(\ell) \\ &= \sum_{\ell} \sum_k h(k) p_{X|Y=\ell}(k, \ell) p_Y(\ell) \\ &= \sum_{\ell} \sum_k h(k) p_{X,Y}(k, \ell) \\ &= \sum_{\ell} \sum_k h(k) p_{X,Y}(k, \ell) \\ &= \sum_k h(k) \sum_{\ell} p_{X,Y}(k, \ell) \\ &= \sum_k h(k) p_X(k) \\ &= \mathbb{E}[h(X)]. \\ \therefore \mathbb{E}[\mathbb{E}[h(X)|Y]] &= \mathbb{E}[h(X)].\end{aligned}$$

□

Exercise 41.

Say that $X \sim U[a, b]$ then

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_a^b x \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \int_a^b x dx \\ &= \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b \\ &= \frac{1}{b-a} \frac{b^2 - a^2}{2} \\ &= \frac{(b-a)(b+a)}{2(b-a)} \\ \Rightarrow \mathbb{E}[X] &= \frac{a+b}{2}.\end{aligned}$$

This makes sense, if numbers are scattered uniformly between a and b then on

average they would be near the middle of the two points, i.e. $(a + b)/2$.

$$\begin{aligned}
 \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x) \, dx \\
 &= \int_a^b x^2 \frac{1}{b-a} \, dx \\
 &= \frac{1}{b-a} \int_{x=a}^b x^2 \, dx \\
 &= \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b \\
 &= \frac{1}{b-a} \frac{b^3 - a^3}{3} \\
 &= \frac{(b-a)(b^2 + ab + a^2)}{3(b-a)} \\
 \Rightarrow \mathbb{E}[X^2] &= \frac{a^2 + ab + b^2}{3}.
 \end{aligned}$$

$$\begin{aligned}
 \text{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
 &= \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2} \right)^2 \\
 &= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} \\
 &= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} \\
 &= \frac{a^2 - 2ab + b^2}{12} \\
 &= \frac{(a-b)^2}{12}.
 \end{aligned}$$

$$\therefore \mathbb{E}[X] = \frac{a+b}{2} \text{ and } \text{Var}[X] = \frac{(a-b)^2}{12}.$$

□

Exercise 42.

Say that $X \sim U[a, b]$. Since the support of X is $[a, b]$ then for $x < a$ then

$F_X(x) = 0$ and for $x > b$ then $F_X(x) = 1$. For $x \in [a, b]$ then

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x f_X(z) \, dz \\ &= \int_a^x \frac{1}{b-a} \, dz \\ &= \frac{1}{b-a} \int_a^x \, dz \\ &= \frac{x-a}{b-a}. \end{aligned}$$

In summary,

$$\therefore F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & x \in [a, b] \\ 1, & x > b \end{cases}.$$

□

Exercise 45.

Say that $X \sim \text{Exp}(\lambda)$.

Part 1.

$$\begin{aligned} \int_0^{\infty} f_X(x) \, dx &= \int_0^{\infty} \lambda e^{-\lambda x} \, dx \\ &= \lambda \int_0^{\infty} e^{-\lambda x} \, dx \\ &= \lambda \left[\frac{-e^{-\lambda x}}{\lambda} \right]_0^{\infty} \\ &= [1 - 0] \\ &= 1. \end{aligned}$$

$$\therefore \int_0^{\infty} f_X(x) \, dx = 1.$$

Part 2.

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^{\infty} x f_X(x) \, dx \\
&= \int_0^{\infty} \lambda x e^{-\lambda x} \, dx \\
&= \left[-x e^{-\lambda x} \right]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} \, dx \\
&= [0 - 0] + \left[-\frac{e^{-\lambda x}}{\lambda} \right]_0^{\infty} \\
&= \frac{1 - 0}{\lambda} \\
&= \frac{1}{\lambda}.
\end{aligned}$$

$$\therefore \mathbb{E}[X] = \frac{1}{\lambda}.$$

Part 3.

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^{\infty} x^2 f_X(x) \, dx \\
&= \int_0^{\infty} \lambda x^2 e^{-\lambda x} \, dx \\
&= \left[-x^2 e^{-\lambda x} \right]_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} \, dx \\
&= [0 - 0] + \frac{2}{\lambda} \int_0^{\infty} \lambda x e^{-\lambda x} \, dx \\
&= \frac{2}{\lambda} \mathbb{E}[X] \\
&= \frac{2}{\lambda^2}.
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \text{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2}.
\end{aligned}$$

$$\therefore \text{Var}[X] = \frac{1}{\lambda^2}.$$

□

Exercise 47.

Say that $X \sim \text{Exp}(\lambda)$. The distribution is memoryless if for all $s, t > 0$ then

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s).$$

Remember that for all $u > 0$ then

$$\begin{aligned} \mathbb{P}(X > u) &= \int_u^\infty \lambda e^{-\lambda x} dx \\ &= \lambda \int_u^\infty e^{-\lambda v} dx \\ &= \lambda \left[\frac{-e^{-\lambda x}}{\lambda} \right]_u^\infty \\ &= e^{-\lambda u}. \end{aligned}$$

So

$$\begin{aligned} \mathbb{P}(X > s + t | X > t) &= \frac{\mathbb{P}(X > s + t, X > t)}{\mathbb{P}(X > t)} \\ &= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > t)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} \\ &= \mathbb{P}(X > s). \end{aligned}$$

For any $s, t > 0 \Rightarrow \mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s)$.

□

Exercise 48.

Say that $Y_i = \mathbb{1}(X_i > \alpha)$ so

$$\#_n\{X_i > \alpha\} = \sum_{i=1}^n Y_i.$$

It is clear that $Y_i \stackrel{iid}{\sim} \text{Ber}(q)$, and as such have finite mean ($\mathbb{E}[Y_i] = q < \infty$). Therefore the SLLN implies that with probability 1

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_i = q &\Rightarrow \lim_{n \rightarrow \infty} \frac{\#_n\{X_i > \alpha\}}{n} = q \\ \therefore \lim_{n \rightarrow \infty} \frac{\#_n\{X_i > \alpha\}}{n} &= q \quad \text{w.p.1.} \end{aligned}$$

□

Exercise 49.

$$\begin{aligned}
a_n &= \mathbb{E} \left[\frac{\sum_{i=1}^n X_i}{n} \right] \\
&= \frac{1}{n} \mathbb{E} \left[\sum_{i=1}^n X_i \right] \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] \\
&= \frac{1}{n} n \mathbb{E}[X_1] \quad \text{as } X_i \text{ identically distributed} \\
&= \frac{n\mu}{n} \\
&= \mu.
\end{aligned}$$

$$\begin{aligned}
b_n^2 &= \mathbb{V}\text{ar} \left[\frac{\sum_{i=1}^n X_i}{n} \right] \\
&= \frac{1}{n^2} \mathbb{V}\text{ar} \left[\sum_{i=1}^n X_i \right] \\
&= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}\text{ar}[X_i] \quad \text{as } X_i \text{ independent} \\
&= \frac{1}{n^2} n \mathbb{V}\text{ar}[X_1] \quad \text{as } X_i \text{ identically distributed} \\
&= \frac{\sigma^2}{n} \\
\Rightarrow b_n &= \frac{\sigma}{\sqrt{n}}.
\end{aligned}$$

$$\therefore a_n = \mu, b_n = \frac{\sigma}{\sqrt{n}}.$$

□

Exercise 51.

Below is the output for a MATLAB session.

```

1 >> p = 0.25;
2 >> B = (rand(1,1e4) < p);
3 >> mean(B)
4 ans =

```

```

5      0.2504
6  >> var(B)
7  ans =
8      0.1877

```

If $B \text{ Ber}(p)$ then $\mathbb{E}[B] = p$ and $\text{Var}[B] = p(1 - p)$ which for $p = 0.25$ means

$$\mathbb{E}[B] = 0.25, \quad \text{Var}[B] = \frac{3}{16} = 0.1875.$$

So

$$\widehat{\mathbb{E}[B]} = 0.2504 \approx 0.25 = \mathbb{E}[B]$$

and

$$\widehat{\text{Var}[B]} = 0.1877 \approx 0.1875 = \text{Var}[B].$$

□

Exercise 52.

For $E \sim \text{Exp}(1/2)$ then $F_E(x) = 1 - e^{-x/2}$ so set

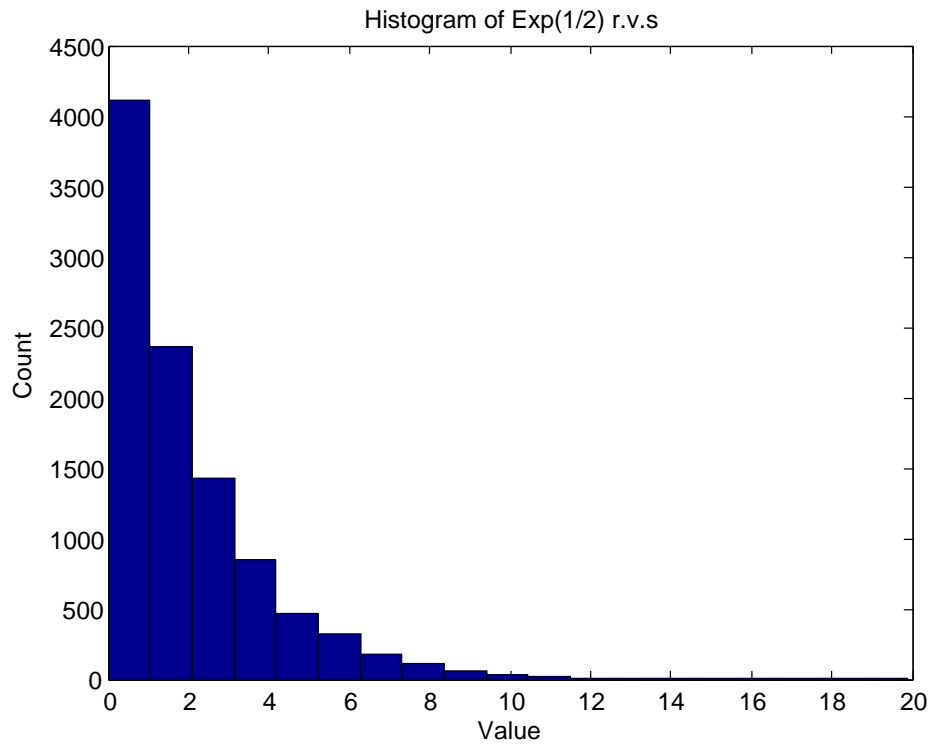
$$\begin{aligned} 1 - e^{-x/2} = u &\Rightarrow \log(1 - u) = -x/2 \Rightarrow x = -2 \log(1 - u) \\ &\Rightarrow F_E^{-1}(u) = -2 \log(1 - u) \end{aligned}$$

Note that if $U_i \stackrel{iid}{\sim} U(0, 1)$ then also $(1 - U_i) \stackrel{iid}{\sim} U(0, 1)$, so the exponential random variables can be generated by $E_i = -2 \log(U_i)$.

```

1  >> E = -log(rand(1,1e4))/0.5;
2  >> mean(E)
3  ans =
4      2.0076
5  >> var(E)
6  ans =
7      4.1118
8  >> hist(E, max(E));
9  >> title('Histogram of Exp(1/2) r.v.s');
10 >> xlabel('Value'); ylabel('Count');

```

If $E \sim \text{Exp}(\lambda)$ then $\mathbb{E}[E] = \lambda^{-1}$ and $\text{Var}[E] = \lambda^{-2}$ which for $\lambda = 1/2$ means

$$\mathbb{E}[E] = 2, \quad \text{Var}[E] = 4.$$

So

$$\widehat{\mathbb{E}[B]} = 2.0076 \approx 2 = \mathbb{E}[B]$$

and

$$\widehat{\text{Var}[B]} = 4.1118 \approx 4 = \text{Var}[B].$$

□

Exercise 53.

```

1 function X = samplePDF(p)
2     X = find(rand() < cumsum(p), 1);
3 end
4 samples = zeros(1, 1e4);
5 for s=1:1e4
6     samples(s) = samplePDF([0.35, 0.25, 0.1, 0.3]);
7 end
8 freqs=histc(samples,1:4)/1e4
9     0.3456    0.2510    0.1012    0.3022

```

So

$$\therefore \hat{\mathbf{p}} = [0.3456, 0.2510, 0.1012, 0.3022].$$

□

Exercise 56.

The P to select would be the one with every row being $p_X(\cdot)$. This means $\mathbb{P}(X(t+1) = x | X(t) = \cdot) = p_X(x)$, so it will have the correct distribution, and it will be independent as the transitions are independent of the current state.

□

Exercise 73.

Consider constructing the sets in $2^{\mathcal{A}}$ by looping through each element in \mathcal{A} many times. At each $a \in \mathcal{A}$ then one might choose to include a or not include a in a particular set in the power set (i.e. 2 options for each element in \mathcal{A}). This means the number of different combinations is

$$2 \cdot 2 \cdot \dots \cdot 2 = 2^{|\mathcal{A}|}.$$

$$\therefore |2^{\mathcal{A}}| = 2^{|\mathcal{A}|}.$$

□