richardwu.ca

STAT 333 Course Notes

Applied Probability

Steve Drekic • Winter 2018 • University of Waterloo

Last Revision: March 29, 2018

Table of Contents

Abstract

These notes are intended as a resource for myself; past, present, or future students of this course, and anyone interested in the material. The goal is to provide an end-to-end resource that covers all material discussed in the course displayed in an organized manner. These notes are my interpretation and transcription of the content covered in lectures. The instructor has not verified or confirmed the accuracy of these notes, and any discrepancies, misunderstandings, typos, etc. as these notes relate to course's content is not the responsibility of the instructor. If you spot any errors or would like to contribute, please contact me directly.

1 January 4, 2018

1.1 Example 1.1 solution

What is the probability that we roll a number less than 4 given that we know it's odd?

Solution. Let $A = \{1, 2, 3\}$ (less than 4) and $B = \{1, 3, 5\}$ (odd). We want to find $P(A \mid B)$. Note that $A \cap B = \{1, 3\}$ and there are six elements in the sample space S thus

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$$

1.2 Example 1.2 solution

Show that $BIN(n, p) \sim POI(\lambda)$ when $\lambda = np$ for n large and p small.

Solution. Let $\lambda = np$. Note that $p = \frac{\lambda}{n} \quad n > 0$. From the pmf for $X \sim BIN(n, p)$

$$\begin{aligned} p(x) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{n(n-1)...(n-x+1)}{x!} (\frac{\lambda}{n})^x (1 - \frac{\lambda}{n})^{n-x} \\ &= \frac{n(n-1)...(n-x+1)}{n^x} \cdot \frac{\lambda^x}{x!} \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^x} \end{aligned}$$

Recall $\lim_{n\to\infty}(1 - \frac{\lambda}{n})^n = e^{-\lambda}$ so

$$\lim_{n\to\infty} p(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$$

2 January 9, 2018

2.1 Example 1.3 solution

Find the mgf of $BIN(n, p)$ and use that to find $E[X]$ and $Var(X)$.

Solution. Recall the binomial series is

$$(a+b)^m = \sum_{x=0}^{m} \binom{m}{x} a^x b^{m-x} \quad a, b \in \mathbb{R}, m \in \mathbb{N}$$

Let $x \sim BIN(n, p)$ and so

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, \ldots, n$$

Taking the mgf $E[e^{tX}]$

$$\begin{aligned}\Phi_X(t) = E[e^{tX}] &= \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x(1-p)^{n-x} \\ &= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x(1-p)^{n-x}\end{aligned}$$

from the binomial series we have

$$\Phi_x(t) = (pe^t + 1 - p)^n \quad t \in \mathbb{R}$$

We can take the first and second derivatives for the first and second moment

$$\begin{aligned}&\Phi'_X(t) = n(pe^t + 1 - p)^{n-1}pe^t \\ &\Phi''_X(t) = np[(pe^t + 1 - p)^{n-1}e^t + e^t(n-1)(pe^t + 1 - p)^{n-2}pe^t]\end{aligned}$$

So $E[X] = \Phi_X(t) \mid_{t=0} = np$.
For the variance, we need the second moment

$$\begin{aligned}E[X^2] &= \Phi_X(t) \mid_{t=0} \\ &= np[1 + (n-1)p] \\ &= np + (np)^2 - np^2\end{aligned}$$

So

$$\begin{aligned}Var(X) &= E[X^2] - E[X]^2 \\ &= np + (np)^2 - np^2 - (np)^2 \\ &= np(1-p)\end{aligned}$$

2.2 Example 1.4 solution

Show that $Cov(X,Y) = 0 \not\Rightarrow$ independence.

Solution. We show this using a counter example

		y		
	$p(x,y)$	0	1	$p_X(x)$
	0	0.2	0	0.2
x	1	0	0.6	0.6
	2	0.2	0	0.2
	$p_Y(y)$	0.4	0.6	1

Note that

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$

where

$$E[XY] = \sum_{x=0}^{2}\sum_{y=0}^{1} xyp(x,y) = (1)(1)(0.6) = 0.6$$
$$E[X] = \sum_{x=0}^{2} xp_X(x) = (1)(0.6) + (2)(0.2) = 0.6 + 0.4 = 1$$
$$E[Y] = \sum_{y=0}^{1} yp_Y(y) = (1)(0.6) = 0.6$$

So $Cov(X,Y) = 0.6 - (1)(0.6) = 0$. However, $p(2,0) = 0.2 \neq p_X(2)p_Y(0) = (0.2)(0.4) = 0.08$, thus X and Y are not independent (they are dependent).

2.3 Example 1.5 solution

Given $X_1, \ldots, X_n$ are independent r.v's where $\Phi_X(t)$ is the mgf of X_i, show that $T = \sum_{i=1}^n X_i$ has mgf $\Phi_T(t) = \prod_{i=1}^n \Phi_{X_i}(t)$.

Solution. We take the definition of the mgf of T

$$\begin{aligned}
\Phi_T(t) &= E[e^{tT}] \\
&= E[e^{t(X_1+\ldots+X_n)}] \\
&= E[e^{tX_1} \cdot \ldots \cdot e^{tX_n}] \\
&= E[e^{tX_1}] \cdot \ldots \cdot E[e^{tX_n}] && \text{independence} \\
&= \prod_{i=1}^{n} \Phi_{X_i}(t)
\end{aligned}$$

2.4 Exercise 1.3

If $X_i \sim POI(\lambda_i)$ show that $T = \sum X_i \sim POI(\sum \lambda_i)$.

Solution. Recall that $POI(\lambda_i) \sim BIN(n_i, p)$ where $\lambda_i = n_i p$ and

$$\Phi_{X_i}(t) = (pe^t + 1 - p)^{n_i} \quad \forall t \in \mathbb{R}$$

where $X_i \sim BIN(n_i, p) \quad i = 1, \ldots, m$.
Therefore

$$\begin{aligned}
\Phi_T(t) &= \prod_{i=1}^{m} (pe^t + 1 - p)^{n_i} \\
&= (pe^t + 1 - p)^{n_1} \cdot \ldots \cdot (pe^t + 1 - p)^{n_m} \\
&= (pe^t + 1 - p)^{\sum n_i} \quad t \in \mathbb{R}
\end{aligned}$$

By the mgf uniqueness property, we have

$$T = \sum_{i=1}^{m} X_i \sim BIN(\sum_{i=1}^{m} n_i, p)$$

3 January 11, 2018

3.1 Theorem 2.1 (conditional variance)

Theorem 3.1.

$$Var(X_1 \mid X_2 = x_2) = E[X_1^2 \mid X_2 = x_2] - E[X_1 \mid X_2 = x_2]^2$$

Proof.

$$\begin{aligned} Var(X_1 \mid X_2 = x_2) &= E[(X_1 - E[X_1 \mid X_2 = x_2])^2 \mid X_2 = x_2] \\ &= E[(X_1^2 - 2E[X_1 \mid X_2 = x_2]X_1 + E[X_1 \mid X_2 = x_2]^2) \mid X_2 = x_2] \\ &= E[X_1^2 \mid X_2 = x_2] - 2E[X_1 \mid X_2 = x_2]E[X_1 \mid X_2 = x_2] + E[X_1 \mid X_2 = x_2]^2 \\ &= E[X_1^2 \mid X_2 = x_2] - E[X_1 \mid X_2 = x_2]^2 \end{aligned}$$

□

3.2 Example 2.1

Suppose that X and Y are discrete random variables having join pmf of the form

$$p(x, y) = \begin{cases} 1/5 & \text{, if x = 1 and y = 0,} \\ 2/15 & \text{, if x = 0 and y = 1,} \\ 1/15 & \text{, if x = 1 and y = 2,} \\ 1/5 & \text{, if x = 2 and y = 0,} \\ 2/5 & \text{, if x = 1 and y = 1,} \\ 0 & \text{, otherwise.} \end{cases}$$

Find the conditional probability of $X \mid (Y = 1)$. Also calculate $E[X \mid Y = 1]$ and $Var(X \mid Y = 1)$.

Solution. Note: for problems of this nature, construct a table.

		y			
	$p(x, y)$	0	1	2	$p_X(x)$
	0	0	2/15	0	2/15
x	1	1/5	2/5	1/15	2/3
	2	1/5	0	0	1/5
	$p_Y(y)$	2/5	8/15	1/15	1

Then we have

$$\begin{aligned} p(0 \mid 1) &= P(X = 0 \mid Y = 1) = \frac{2/15}{8/15} = \frac{1}{4} \\ p(1 \mid 1) &= P(X = 1 \mid Y = 1) = \frac{2/5}{8/15} = \frac{3}{4} \\ p(2 \mid 1) &= P(X = 2 \mid Y = 1) = \frac{0}{8/15} = 0 \end{aligned}$$

The conditional pmf of $X \mid (Y = 1)$ can be represented as follows

x	0	1
$p(x \mid 1)$	$1/4$	$3/4$

We observe $X \mid (Y = 1) \sim Bern(3/4)$. We can take the known $E[X] = p$ and $Var(X)p(1-p)$ for $X \sim Bern(p)$, thus

$$\begin{aligned}
&E[X \mid (Y = 1)] = 3/4 \\
&Var(X \mid (Y = 1)) = 3/4(1 - 3/4) = 3/16
\end{aligned}$$

3.3 Example 2.2

For $i = 1, 2$ suppose that $X_i \sim BIN(n_i, p)$ where X_1, X_2 are independent (but not identically distributed). Find conditional distribution of X_1 given $X_1 + X_2 = n$.

Solution. We want to find conditional pmf of $X \mid (X_1 + X_2 = n)$. Let this conditional pmf be denoted by

$$\begin{aligned}
p(x_1 \mid n) &= P(X_1 = x_1 \mid X_1 + X_2 = n) \\
&= \frac{P(X_1 = x_1, X_1 + X_2 = n)}{P(X_1 + X_2 = n)}
\end{aligned}$$

Recall: $X_1 + X_2 \sim BIN(n_1 + n_2, p)$ so

$$P(X_1 + X_2 = n) = \binom{n_1 + n_2}{n} p^n (1-p)^{n_1+n_2-n}$$

Next, consider

$$\begin{aligned}
P(X_1 = x_1, X_1 + X_2 = n) &= P(X_1 = x_1, x_1 + X_2 = n) \\
&= P(X_1 = x_1, X_2 = n - x_1) \\
&= P(X_1 = x_1)P(X_2 = n - x_1) && \text{independence} \\
&= \binom{n_1}{x_1} p^{x_1} (1-p)^{n_1 - x_1} \cdot \binom{n_2}{n - x_1} p^{n - x_1} (1-p)^{n_2 - (n - x_1)}
\end{aligned}$$

provided that $0 \leq x_1 \leq n_1$ **and**

$$\begin{aligned}
&0 \leq n - x_1 \leq n_2 \\
&-n_2 \leq x_1 - n \leq 0 \\
&n - n_2 \leq x_1 \leq n
\end{aligned}$$

(from the binomial coefficients). Therefore our domain for x_1 is

$$x_1 = \max\{0, n - n_2\}, \ldots, \min\{n_1, n\}$$

Thus we have

$$\begin{aligned}
p(x_1 \mid n) &= \frac{P(X_1 = x, X_1 + x_2 = n)}{P(X_1 + X_2 = n)} \\
&= \frac{\binom{n_1}{x_1} p^{x_1} (1-p)^{n_1 - x_1} \cdot \binom{n_2}{n - x_1} p^{n - x_1} (1-p)^{n_2 - (n - x_1)}}{\binom{n_1 + n_2}{n} p^n (1-p)^{n_1 + n_2 - n}} \\
&= \frac{\binom{n_1}{x_1}\binom{n_2}{n - x_1}}{\binom{n_1 + n_2}{n}}
\end{aligned}$$

for $x_1 = \max\{0, n - n_2\}, \ldots, \min\{n_1, n\}$.
Recall: A $HG(N, r, n)$ (hypergeometric) distribution has pmf

$$\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{x}} \quad x = \max\{0, n - N + r\}, \ldots, \min\{n, r\}$$

So this is precisely $HG(n_1 + n_2, x_1, n)$.
If you think about it: we are choosing x_1 successes from n_1 trials from the first set X_1 and choosing the remaining $n - x_1$ successes from n_2 trials from X_2.

4 Tutorial 1

4.1 Exercise 1: MGF of Erlang

Find the mgf of $X \sim Erlang(\lambda)$ and use it to find $E[X], Var(X)$.
Note that the Erlang's pdf is for $n \in \mathbb{Z}^+$ and $\lambda > 0$

$$f(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \quad x > 0$$

Solution.

$$\begin{aligned}
\Phi_X(t) = E[e^{tX}] &= \int_0^\infty e^{tx} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} dx \\
&= \int_0^\infty \frac{\lambda^n x^{n-1} e^{-(\lambda - t)x}}{(n-1)!} dx
\end{aligned}$$

Note that the term in the integral is similar to the pdf of Erlang but for $\lambda = \lambda - t$. So we try to fix it so the integral is this pdf of Erlang

$$\begin{aligned}
\Phi_X(t) &= \int_0^\infty \frac{\lambda^n x^{n-1} e^{-(\lambda - t)x}}{(n-1)!} dx \\
&= (\frac{\lambda}{\lambda - t})^n \int_0^\infty \frac{(\lambda - t)^n x^{n-1} e^{-(\lambda - t)x}}{(n-1)!} dx \\
&= (\frac{\lambda}{\lambda - t})^n & t < \lambda
\end{aligned}$$

since the integral over the positive real line of the pdf of an $Erlang(n, \lambda - t)$ is 1 and $t < \lambda$ must hold so the rate parameter $\lambda - t$ is positive.

Differentiating,

$$
\begin{aligned}
\Phi_X^{(1)}(t) &= \frac{d}{dt}\left(\frac{\lambda}{(\lambda - t)^n}\right) \\
&= \frac{n\lambda^n}{(\lambda - t)^{n+1}} \\
\Phi_X^{(2)}(t) &= \frac{d}{dt}\left(\frac{n\lambda^n}{(\lambda - t)^{n+1}}\right) \\
&= \frac{n(n+1)\lambda^n}{(\lambda - t)^{n+2}}
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
E[X] = \Phi_X^{(1)}(0) &= \left.\frac{n\lambda^n}{(\lambda - t)^{n+1}}\right|_{t=0} = \frac{n}{\lambda} \\
E[X^2] = \Phi_X^{(2)}(0) &= \left.\frac{n(n+1)\lambda^n}{(\lambda - t)^{n+2}}\right|_{t=0} = \frac{n(n+1)}{\lambda^2} \\
Var(X) = E[X^2] - E[X]^2 &= \frac{n(n+1)}{\lambda^2} - \frac{n}{\lambda} = \frac{n}{\lambda^2}
\end{aligned}
$$

Remark 4.1. To solve any of these mgfs, it is useful to see if one can reduce the integral into a pdf of a known distribution (possibly itself).

4.2 Exercise 2: MGF of Uniform

Find the mgf of the uniform distribution on $(0, 1)$ and find $E[X]$ and $Var(X)$.

Solution. Let $X \sim U(0, 1)$ so that $f(x) = 1 \quad 0 \le x \le 1$. We have

$$
\begin{aligned}
\Phi_X(t) = E[e^{tX}] &= \int_0^1 e^{tx}(1)dx \\
&= \left.\frac{1}{t}e^{tx}\right|_{x=0}^{x=1} \\
&= t^{-1}(e^t - 1) \quad t \neq 0
\end{aligned}
$$

Differentiating

$$
\begin{aligned}
\Phi_X^{(1)}(t) &= \frac{d}{dt}(t^{-1}(e^t - 1)) \\
&= t^{-1}e^t - t^{-2}(e^t - 1) \\
&= \frac{te^t - e^t + 1}{t^2} \\
\Phi_X^{(2)}(t) &= \frac{d}{dt}\left(\frac{te^t - e^t + 1}{t^2}\right) \\
&= \frac{t^2(te^t + e^t - e^t) - 2t(te^t - e^t + 1)}{t^4} \\
&= \frac{t^2e^t - 2te^t + 2e^t - 2}{t^3}
\end{aligned}
$$

We may calculate the first two moments by applying **L'Hopital's rule** to calculate the limits

$$\begin{aligned}
E[X] = \Phi_X^{(1)}(t)\Big|_{t=0} &= \lim_{t\to\infty} \frac{te^t - e^t + 1}{t^2} \\
&= \lim_{t\to\infty} \frac{te^t + e^t - e^t}{2t} \\
&= \lim_{t\to\infty} \frac{e^t}{2} = \frac{1}{2}
\end{aligned}$$

Similarly

$$\begin{aligned}
E[X^2] = \Phi_X^{(2)}(t)\Big|_{t=0} &= \lim_{t\to\infty} \frac{t^2e^t - 2te^t + 2e^t - 2}{t^3} \\
&= \lim_{t\to\infty} \frac{t^2e^t + 2te^t - 2te^t - 2e^t + 2e^t}{3t^2} \\
&= \lim_{t\to\infty} \frac{e^t}{3} = \frac{1}{3}
\end{aligned}$$

So we have

$$Var(X) = E[X^2] - E[X]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

4.3 Exercise 3: Moments from PGF

Suppose X is a discrete r.v. on $\mathbb{N}$ with pmf $p(x)$. Show how to find the first two moments of X from its pgf.

Solution. By definition, the pgf of X is $\Psi_X(z) = E[z^X] = \sum_{x=0}^{\infty} z^x p(x)$.
If we let $z = 1$, then the sum equals 1. However, if we take its derivative with respect to z just once

$$\Psi_X^{(1)}(z) = \frac{d}{dz} \sum_{x=0}^{\infty} z^x p(x) = \sum_{x=1}^{\infty} xz^{x-1} p(x)$$

Letting $z = 1$ we can find the first moment

$$\begin{aligned}
\Psi_X^{(1)}(1) &= \lim_{z\to 1} \sum_{x=1}^{\infty} xz^{x-1}p(x) \\
&= \sum_{x=1}^{\infty} xp(x) \\
&= \sum_{x=0}^{\infty} xp(x) && \text{when } x = 0 \text{ the term is 0 anyways} \\
&= E[X]
\end{aligned}$$

For the second moment, we consider the second derivative

$$\begin{aligned}
\Psi_X^{(1)}(z) &= \frac{d^2}{dz^2} \sum_{x=0}^{\infty} z^x p(x) \\
&= \sum_{x=2}^{\infty} x(x-1) z^{x-2} p(x)
\end{aligned}$$

Letting $z = 1$

$$\begin{aligned}
\Psi_X^{(2)}(1) &= \lim_{z\to 1}\sum_{x=2}^{\infty} x(x-1)z^{x-2}p(x) \\
&= \sum_{x=2}^{\infty} x(x-1)p(x) \\
&= \sum_{x=0}^{\infty} x(x-1)p(x) \\
&= E[X(X-1)] \\
&= E[X^2] - E[X]
\end{aligned}$$

So we have $E[X^2] = \Psi_X^{(2)}(1) + \Psi_X^{(1)}(1)$. To find the variance

$$Var(X) = \Psi_X^{(2)}(1) + \Psi_X^{(1)}(1) - (\Psi_X^{(1)}(1))^2$$

4.4 Exercise 4: PGF of Poisson

Suppose $X \sim POI(\lambda)$. Find the pgf of X and use it to find $E[X]$ and $Var(X)$. The pmf of $POI(\lambda)$ for $\lambda > 0$

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!} \quad x = 0, 1, 2, \ldots$$

Solution.

$$\begin{aligned}
\Psi_X(z) = E[z^X] &= e^{-\lambda}\sum_{x=0}^{\infty}\frac{(z\lambda)^x}{x!} \\
&= e^{-\lambda}\cdot e^{z\lambda} \\
&= e^{\lambda(z-1)}
\end{aligned}$$

where the second equality holds since the summation is the Taylor expansion of $e^{z\lambda}$.
Differentiating

$$\begin{aligned}
\Psi_X^{(1)}(z) &= \frac{d}{dz}e^{\lambda(z-1)} \\
&= \lambda e^{\lambda(z-1)} \\
\Psi_X^{(2)}(z) &= \frac{d}{dz}\lambda e^{\lambda(z-1)} \\
&= \lambda^2 e^{\lambda(z-1)}
\end{aligned}$$

The moments are thus

$$\begin{aligned}
E[X] &= \Phi_X^{(1)}(1) = \lambda e^{\lambda(1-1)} = \lambda \\
E[X(X-1)] &= \Phi_X^{(2)}(1) = \lambda^2 e^{\lambda(1-1)} = \lambda^2 \\
E[X^2] &= E[X(X-1)] + E[X] = \lambda^2 + \lambda \\
Var(X) &= E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda
\end{aligned}$$

5 January 16, 2018

5.1 Example 2.3 solution

Let $X_1, \dots, X_m$ be independent r.v.'s where $X_i \sim POI(\lambda_i)$. Define $Y = \sum_{i=1}^m X_i$. Find the conditional distribution $X_j \mid (Y = n)$.

Solution. We set out to find

$$\begin{aligned}
p(x_j \mid n) = p(X_j = x_j \mid Y = n) &= \frac{P(X_j = x_j, Y = n)}{P(Y = n)} \\
&= \frac{P(X_j = x_j, \sum_{i=1}^m X_i = n)}{P(Y = n)} \\
&= \frac{P(X_j = x_j, X_j + \sum_{i=1, i\neq j}^m X_i = n)}{P(Y = n)} \\
&= \frac{P(X_j = x_j, \sum_{i=1, i\neq j}^m X_i = n - x_j)}{P(Y = n)} \\
&= \frac{P(X_j = x_j) P(\sum_{i=1, i\neq j}^m X_i = n - x_j)}{P(Y = n)} && \text{independence of } X_i
\end{aligned}$$

Remember that if $X_i \sim POI(\lambda_i)$, then

$$Y = \sum_{i=1}^m X_i \sim POI(\sum_{i=1}^m \lambda_i)$$

which can be derived from mgfs (Exercise 1.3). Therefore

$$\sum_{i=1, i\neq j}^m X_i \sim POI(\sum_{i=1, i\neq j}^m \lambda_i)$$

Expanding out $p(x_j \mid n)$ with the pdfs

$$p(x_j \mid n) = \frac{\frac{e^{-\lambda_j}\lambda_j^{x_j}}{x_j!} \cdot \frac{e^{-\sum_{i=1, i\neq j}\lambda_i}(\sum_{i=1, i\neq j}\lambda_i)^{n-x_j}}{(n-x_j)!}}{\frac{e^{-\sum_{i=1}^m \lambda_i} \cdot (\sum_{i=1}^m \lambda_i)^n}{n!}}$$

where $x_j \geq 0$ and $n - x_j \geq 0 \Rightarrow 0 \leq x_j \leq n$ (from the factorials).

Cancelling out the e^{λ} terms and let $\lambda_Y = \sum_{i=1}^{m} \lambda_i$

$$\begin{aligned} p(x_j \mid n) &= \frac{n!}{(n-x_j)!x_j!} \frac{\lambda_j^{x_j}}{\lambda_Y^{x_j}} \frac{(\lambda_Y - \lambda_j)^{n-x_j}}{\lambda_Y^{n-x_j}} \\ &= \binom{n}{x_j} (\frac{\lambda_j}{\lambda_Y})^{x_j} (1 - \frac{\lambda_j}{\lambda_Y})^{n-x_j} \end{aligned}$$

This is the binomial distribution, so we have

$$X_j \mid Y = n \sim BIN(n, \frac{\lambda_i}{\lambda_Y})$$

5.2 Example 2.4 solution

Suppose $X \sim POI(\lambda)$ and $Y \mid (X = x) \sim BIN(x, p)$. Find the conditional distribution $X \mid Y = y$.
(Note: range of y depends on x (that is $y \leq x$). Graphically, we have integral points on and below the $y = x$ line starting from 0 for both x and y).

Solution. We wish to find the conditional pmf given by $X \mid Y = y$ or

$$p(x \mid y) = P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

Note that also

$$\begin{aligned} P(Y = y \mid X = x) &= \frac{P(Y = y, X = x)}{P(X = x)} \\ \Rightarrow P(X = x, Y = y) &= P(X = x)P(Y = y \mid X = x) \\ &= \frac{e^{-\lambda}\lambda^x}{x!} \cdot \binom{x}{y} p^y (1-p)^{x-y} \end{aligned}$$

for $x = 0, 1, 2, \ldots$ **and** $y = 0, 1, 2, \ldots, x$ (range of y depends on x).
To find the marginal marginal pmf of Y, we use

$$p_Y(y) = \sum_x p(x, y)$$

To find the support for x, note that from the graphical region, we realize that $x = 0, 1, 2, \ldots$ **and** $y = 0, 1, 2, \ldots, x$ is equivalent to $y = 0, 1, 2, \ldots$ **and** $x = y, y+1, y+2, \ldots$.

So

$$\begin{aligned}
p_Y(y) &= \sum_{x=y}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} \frac{x!}{(x-y)!y!} p^y (1-p)^{x-y} \\
&= \frac{\lambda^y e^{-\lambda} p^y}{y!} \sum_{x=y}^{\infty} \frac{\lambda^{x-y}(1-p)^{x-y}}{(x-y)!} \\
&= \frac{e^{-\lambda}(\lambda p)^y}{y!} \sum_{x=y}^{\infty} \frac{[\lambda(1-p)]^{x-y}}{(x-y)!} \\
&= \frac{e^{-\lambda}(\lambda p)^y}{y!} e^{\lambda(1-p)} \\
&= \frac{e^{-\lambda p}(\lambda p)^y}{y!} \qquad\qquad y = 0, 1, 2, \ldots
\end{aligned}$$

Note that $p_Y(y) \sim POI(\lambda p)$.
Thus

$$\begin{aligned}
p(x \mid y) &= \frac{P(X = x, Y = y)}{P(Y = y)} \\
&= \frac{\frac{e^{-\lambda}\lambda^x}{x!} \cdot \frac{x!}{(x-y)!y!} p^y (1-p)^{x-y}}{\frac{e^{-\lambda p}(\lambda p)^y}{y!}} \\
&= \frac{e^{-\lambda+\lambda p}[\lambda(1-p)]^{x-y}}{(x-y)!} \\
&= \frac{e^{-\lambda(1-p)}[\lambda(1-p)]^{x-y}}{(x-y)!} \qquad\qquad x = y, y+1, y+2, \ldots
\end{aligned}$$

This resembles the POIson distribution with $\lambda = \lambda(1-p)$ but with a slightly modified domain.
So we see that

$$W \mid (Y = y) \sim W + y$$

where $W \sim POI(\lambda(1-p))$. This is the **shifted Poisson pmf** y units to the right (note that W and y are random variables).
We can easily find the conditional expectations and variance e.g.

$$E[X \mid Y = y] = E[W + y] = E[W] + y$$

5.3 Example 2.5 solution

Suppose the joint pdf of X and Y is

$$f(x, y) = \begin{cases} \frac{12}{5} x(2 - x - y) & , 0 < x < 1, 0 < y < 1, \\ 0 & , \text{ elsewhere} \end{cases}$$

Determine the conditional distribution of X given $Y = y$ where $0 < y < 1$. Also calculate the mean of $X \mid (Y = y)$. (Note: the graphical region is a unit square box where the bottom left corner is at $0, 0$: the inside of the box is the support).

Solution. Using our theory, we wish to find the conditional pdf of $X \mid (Y = y)$ given by

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

For $0 < y < 1$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\ &= \int_0^1 \frac{12}{5}x(2-x-y)dx \\ &= \frac{12}{5}\int_0^1 (2x - x^2 - xy)dx \\ &= \frac{12}{5}(x^2 - \frac{x^3}{3} - \frac{x^2 y}{2})\Big|_0^1 \\ &= \frac{12}{5}(1 - \frac{1}{3} - \frac{y}{2}) \\ &= \frac{2}{5}(4 - 3y) \end{aligned}$$

So we have

$$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{\frac{12}{5}x(2-x-y)}{\frac{2}{5}(4-3y)} \\ &= \frac{6x(2-x-y)}{4-3y} \end{aligned}$$

Thus we have

$$\begin{aligned} E[X \mid Y] &= \int_0^1 x \cdot f_{X|Y}(x \mid y)dx \\ &= \frac{5-4y}{2(4-3y)} \end{aligned}$$

6 January 18, 2018

6.1 Example 2.6 solution

Suppose the joint pdf of X and Y is

$$f(x,y) = \begin{cases} 5e^{-3x-y} & , 0 < 2x < y < \infty, \\ 0 & , \text{ otherwise} \end{cases}$$

Find the conditional distribution of $Y \mid (X = x)$ where $0 < x < \infty$.
Note the region of support is a "flag" (upright triangle with downawrd point) where the slanted part is the line $y = 2x$.

Solution. We wish to find

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

For $0 < x < \infty$

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \int_{2x}^{\infty} 5e^{-3x-y}dy \\
&= 5e^{-3x}\int_{2x}^{\infty} 5e^{-y}dy \\
&= 5e^{-3x}(-e^{-y})\Big|_{2x}^{\infty} \\
&= 5e^{-3x}e^{-2x}) \\
&= 5e^{-5x}
\end{aligned}
$$

so we have $f_X(x) \sim Exp(5)$.

Remark 6.1. The bounds on the integral are in terms of y: it is dependent on x in our $f(x,y)$ definition.

Now

$$
\begin{aligned}
f_{Y|X}(y \mid x) &= \frac{5e^{-3x-y}}{5e^{-5x}} \\
&= e^{-y+2x} \quad y > 2x
\end{aligned}
$$

Note: recognize the conditional pdf of $Y \mid (X = x)$ as that of a shifted exponential distribution ($2x$ units to the right). Specifically, we have

$$Y \mid (X = x) \sim W + 2x$$

where $W \sim Exp(1)$. Thus $E[Y \mid (X = x)] = E(W) + 2x$ and $Var[Y \mid (X = x)] = Var(W)$.

6.2 Example 2.7 solution

Suppose $X \sim U(0,1)$ and $Y \mid (X = x) \sim Bern(x)$. Find the conditional distribution $X \mid (Y = y)$.
Note: X is continuous and $Y \mid (X = x)$ is discrete.

Solution. We wish to find

$$f_{X|Y}(x \mid y) = \frac{p(y \mid x)f_X(x)}{p_Y(y)}$$

From the given information, we have $f_X(x) = 1$ for $0 < x < 1$ Furthermore $p(y \mid x) = Bern(x) = x^y(1-x)^{1-y}$ for $y = 0, 1$.
For $y = 0, 1$ note that (from $\int f(x \mid y)dx = 1$)

$$
\begin{aligned}
p_Y(y) &= \int_{-\infty}^{\infty} p(y \mid x)f_X(x)dx \\
p_Y(y) &= \int_{0}^{1} x^y(1-x)^{1-y}dx
\end{aligned}
$$

To compute this integral, let's check $p_Y(0)$ and $p_Y(1)$

$$\begin{aligned}
p_Y(0) &= \int_0^1 x^0(1-x)^{1-0}dx \\
&= \int_0^1 1 - xdx \\
&= x - \frac{x^2}{2}\bigg|_0^1 \\
&= \frac{1}{2}
\end{aligned}$$

Similarly, take $y = 1$ where $p_Y(1) = \frac{1}{2}$.
In other words, we have that $p_Y(y) = \frac{1}{2}$ $y = 0, 1$ so

$$Y \sim Bern\left(\frac{1}{2}\right)$$

So

$$\begin{aligned}
f(x \mid y) &= \frac{p(y \mid x)f_X(x)}{p_Y(y)} \\
&= \frac{x^y(1-x)^{1-y} \cdot 1}{\frac{1}{2}} \\
&= 2x^y(1-x)^{1-y} \quad 0 < x < 1
\end{aligned}$$

6.3 Theorem 2.2 (law of total expectation)

Theorem 6.1. For random variables X and Y, $E[X] = E[E[X \mid Y]]$.

Proof. WLOG assume X, Y are jointly continuous random variables. We note

$$\begin{aligned}
E[E[X \mid Y]] &= \int_{-\infty}^{\infty} E[X \mid Y = y]f_Y(y)dy \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)dx\right] f_Y(y)dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\frac{f(x,y)}{f_Y(y)} \cdot f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f(x,y)dy\right] dx \\
&= \int_{-\infty}^{\infty} xf_X(x)dx \\
&= E[X]
\end{aligned}$$

□

6.4 Example 2.8 solution

Suppose $X \sim GEO(p)$ with pmf $p_X(x) = (1-p)^{x-1}p$ where $x = 1, 2, 3, \ldots$. Calculate $E[X]$ and $Var(X)$ using the law of total expectation.

Solution. Recall $E[X] = \frac{1}{p}$ and $Var(X) = \frac{1-p}{p^2}$ where X models the number of (independent) trials necessary to obtain the first success.
Remember: we could manually solve $E[X] = \sum_{x=1}^{\infty}(1-p)^{x-1}p$ and similarly $Var(X) = E[X^2] - E[X]$, or take the derivatives of the mgf $\Phi_X(t) = E[e^{tX}]$. This is tedious in general.

7 Tutorial 2

7.1 Sum of geometric distributions

Let X_i for $i = 1, 2, 3$ be independent geometric random variables having the same parameter p. Determine the value

$$P(X_j = x_j \mid \sum_{i=1}^{3} X_i = n)$$

Solution. Note that, by construction, the sum of k independent $GEO(p)$ random variables is distributed as $NB(k, p)$. Recall that

$$\begin{aligned}
&X_i \sim GEO(p) \Rightarrow P_{X_i}(x) = (1-p)^{x-1}p && x = 1, 2, 3, \ldots \\
&Y \sim NB(k, p) \Rightarrow P_Y(y) = \binom{y-1}{k-1} p^k (1-p)^{y-k} && y = k, k+1, k+2, \ldots
\end{aligned}$$

Breaking apart the summation

$$\begin{aligned}
P(X_j = x_j \mid \sum_{i=1}^{3} X_i = n) &= P(X_j = x_j \mid X_j + \sum_{i=1, i\neq j}^{3} X_i = n) \\
&= \frac{P(X_j = x_j, X_j + \sum_{i=1,i\neq j}^{3} X_i = n)}{P(\sum_{i=1}^{3} X_i = n)} \\
&= \frac{P(X_j = x_j, \sum_{i=1,i\neq j}^{3} X_i = n - x_j)}{P(\sum_{i=1}^{3} X_i = n)} \\
&= \frac{P(X_j = x_j) \cdot P(\sum_{i=1,i\neq j}^{3} X_i = n - x_j)}{P(\sum_{i=1}^{3} X_i = n)} && X_i\text{'s are independent} \\
&= \frac{(1-p)^{x_j-1}p \cdot \binom{n-x_j-1}{1} p^2 (1-p)^{n-x_j-2}}{\binom{n-1}{2} p^3 (1-p)^{n-3}} && \text{provided that } x_j \geq 1 \text{ and } n - x_j \geq 2 \\
&= \frac{(1-p)^{x_j-1}p \cdot \binom{n-x_j-1}{1} p^2 (1-p)^{n-x_j-2}}{\binom{n-1}{2} p^3 (1-p)^{n-3}} \\
&= \frac{(n-x_j-1)!}{1!(n-x_j-2)!} \cdot \frac{2!(n-3)!}{(n-1)!} \\
&= \frac{2(n-x_j-1)}{(n-1)(n-2)} \quad x_j = 1, 2, \ldots, n-2
\end{aligned}$$

Note this is a pmf so we can check

$$\begin{aligned}
\sum_{x_1}^{n-2} \frac{2(n-x_1)}{(n-1)(n-2)} &= \sum_{x_1}^{n-2} \frac{2(n-1)}{(n-1)(n-2)} - \sum_{x_1}^{n-2} \frac{2x}{(n-1)(n-2)} \\
&= \frac{2(n-1)(n-2)}{(n-1)(n-2)} - \frac{2}{(n-1)(n-2)} \sum_{x=1}^{n-2} x \\
&= 2 - \frac{2}{(n-1)(n-2)} \cdot \frac{(n-2)(n-1)}{2} \\
&= 2 - 1 \\
&= 1
\end{aligned}$$

which satisfies the cdf axiom.

7.2 Conditional card drawing

Given $N \in \mathbb{Z}^+$ cards labelled $1, 2, \ldots, N$, let X represent the number that is picked. Suppose a second card Y is picked from $1, 2, \ldots, X$.
Assuming $N = 10$, calculate the expected value of X given $Y = 8$.

Solution. Clearly we have that $P_X(x) = \frac{1}{N}$ where $x = 1, 2, \ldots, N$ and $P_{Y|X}(y \mid x) = \frac{1}{x}$ for $y = 1, 2, \ldots, x$.
To find the conditional distribution of $X \mid (Y = y)$ we must identify the joint distribution of X, Y. It immediately follows that

$$p(x, y) = P(X = x, Y = y) = P_{Y|X}(y \mid x) P_X(x) = \frac{1}{xN}$$

for $x = 1, 2, \ldots, N$ and $y = 1, 2, \ldots, x$. or equivalently the range can be re-expressed as

$$y = 1, 2, \ldots, N \text{ and } x = y, y+1, \ldots, N$$

Remark 7.1. Whenever we want to find the marginal pmf/pdf for a given rv Y, we generally need to re-map the support such that the support of Y is independent of the other rv X.

Note that

$$\begin{aligned}
P_Y(y) = \sum_{x=y}^{N} p(x, y) &= \sum_{x=y}^{N} \frac{1}{xN} \\
&= \frac{1}{N} \sum_{x=y}^{N} \frac{1}{x} \quad y = 1, 2, \ldots, N
\end{aligned}$$

Letting $N = 10$, we can calculate

$$\begin{aligned}
E[X \mid Y = 8] &= \sum_{x=8}^{10} x P_{X|Y}(x \mid 8) \\
&= \sum_{x=8}^{10} x \frac{P(x,8)}{P_Y(8)} \\
&= \sum_{x=8}^{10} x \frac{\frac{1}{10x}}{\frac{1}{10}\sum_{z=8}^{10}\frac{1}{z}} \\
&= \sum_{x=8}^{10} x (\sum_{z=8}^{10} \frac{1}{z})^{-1} \\
&= 3(\frac{1}{8} + \frac{1}{9} + \frac{1}{10})^{-1} \\
&= 3(\frac{242}{720})^{-1} \\
&= \frac{1080}{121} \approx 8.9256
\end{aligned}$$

7.3 Conditional points from interval

Let us choose a random point from interal $(0,1)$ denoted as rv X_1. We then choose a random point X_2 on the interval $(0, x_1)$ hwere x_1 is the realized value of X_1.

1. Make assumptions about the marginal pdf $f_1(x_1)$ and conditional pdf $f_{2|1}(x_2 \mid x_1)$.
2. Find the conditional mean $E[X_1 \mid X_2 = x_2]$.
3. Compute $P(X_1 + X_2 \geq 1)$.

Solution.

1. It makes sense that $X_1 \sim U(0,1)$ and $X_2 \mid (X_1 = x_1) \sim U(0, x_1)$ so that $f_1(x_1) = 1$, $0 < x_1 < 1$ and $f_{2|1}(x_2 \mid x_1) = \frac{1}{x_1}$ for $0 < x_2 < x_1 < 1$.
2. Note that $f_{1|2}(x_1 \mid x_2) = \frac{f(x_1,x_2)}{f_2(x_2)}$ and so we need to identify the joint distribution of x_1 and x_2 as well as the marginal distribution of X_2. We have

$$\begin{aligned}
f(x_1, x_2) &= f_{2|1}(x_2 \mid x_1) \cdot f_1(x_1) \\
&= \frac{1}{x_1} \qquad\qquad 0 < x_2 < x_1 < 1 \quad 0 < x_1 < 1
\end{aligned}$$

or equivalently, the region of support can be re-expressed as

$$\begin{aligned}
&0 < x_2 < 1 \\
&x_2 < x_1 < 1
\end{aligned}$$

so the marginal pdf of $f_2(x_2)$ is

$$\begin{aligned}
f_2(x_2) &= \int_{x_1=x_2}^{1} p(x_1,x_2)dx_1 \\
&= \int_{x_1=x_2}^{1} \frac{1}{x_1}dx_1 \\
&= \ln(x_1)\Big|_{x_1=x_2}^{x_1=1} \\
&= -\ln(x_2) \quad 0<x_2<1
\end{aligned}$$

so the conditional pdf is

$$\begin{aligned}
f_{1|2}(x_1 \mid x_2) &= \frac{f(x_1,x_2)}{f_2(x_2)} \\
&= \frac{1}{-x_1\ln(x_2)} \quad 0<x_2<x_1<1
\end{aligned}$$

Taking the expectation

$$\begin{aligned}
E[X_1 \mid X_2=x_2] &= \int_{x_1=x_2}^{1} x_1 p_{1|2}(x_1,x_2)dx_1 \\
&= \int_{x_1=x_2}^{1} x_1 \cdot \frac{1}{-x_1\ln(x_2)}dx_1 \\
&= \int_{x_1=x_2}^{1} \frac{1}{-\ln(x_2)}dx_1 \\
&= \frac{1-x_2}{-\ln(x_2)} \quad 0<x_2<1
\end{aligned}$$

Exercise: solve for $lim_{x_2\to 1}E[X_1 \mid X_2=x_2]$ (use LHR).

3. The probability that $X_1+X_2 \geq 1$ may be calculated by taking the double integral over the region R of their support where $X_1+X_2\geq 1$ holds. This region may be found as follows:

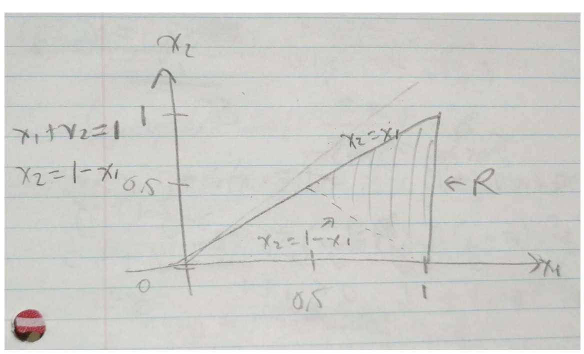


Figure 7.1: The region R is the support where $X_1+X_2\geq 1$.

The region R is equivalent to the bounds $\frac{1}{2}<x_1<1$ and $1-x_1<x_2<x_1$.

Integrating $f(x_1, x_2)$ over R we obtain

$$
\begin{aligned}
P(X_1 + X_2 \geq 1) &= \int_R \int f(x_1, x_2) dx_2 dx_1 \\
&= \int_{\frac{1}{2}}^{1} \int_{1-x_1}^{x_1} \frac{1}{x_1} dx_2 dx_1 \\
&= \int_{\frac{1}{2}}^{1} \frac{x_2}{x_1} \Big|_{x_2=1-x_1}^{x_2=x_1} dx_1 \\
&= \int_{\frac{1}{2}}^{1} (2 - \frac{1}{x_1}) dx_1 \\
&= (2x_1 - \ln(x_1)) \Big|_{x_1=\frac{1}{2}}^{x_1=1} \\
&= 1 + \ln(\frac{1}{2}) \\
&= 1 - \ln(2) \\
&\approx 0.3068528
\end{aligned}
$$

8 January 23, 2018

8.1 Example 2.8 solution

Suppose $X \sim GEO(p)$ with pmf $p_X(x) = (1-p)^{x-1}p$ for $x = 1, 2, 3 \ldots$. Calculate $E[X], Var(X)$ using the law of total expectation.

Solution. Recall X is modelling the number of trials needed to obtain the **1st success**. We want to calculate $E[X]$ and $Var(X)$ using the total law of expectation.
Define

$$
Y = \begin{cases} 0 & \text{if the 1st trial is a failure} \\ 1 & \text{if the 1st trial is a success} \end{cases}
$$

Note that $Y \sim Bern(p)$ so that $P_Y(0) = P(Y = 0) = 1 - p$ and similarly $P_Y(1) = P(Y = 1) = p$.
Thus by the law of total expectation

$$
\begin{aligned}
E[X] &= E[E[X \mid Y]] \\
&= \sum_{y=0}^{1} E[X \mid Y = y] p_Y(y) \\
&= (1-p)E[X \mid Y = 0] + pE[X \mid Y = 1]
\end{aligned}
$$

Note that

$$
X \mid (Y = 1) = 1
$$

with probability 1 (one success is equivalent to $X = 1$ for $GEO(p)$), and

$$
X \mid (Y = 0) \sim 1 + X
$$

(the first one failed, we expect to take X more trials; same initial problem - recurse. See course notes for formal proof).

Thus we have

$$\begin{aligned}
E[X] &= (1-p)E[1+X] + p(1) \\
&= (1-p)(1+E[X]) + p \\
&= 1 + (1-p)E[X] \\
\Rightarrow & E[X](1-(1-p)) = 1 \\
\Rightarrow & E[X] = \frac{1}{p}
\end{aligned}$$

as expected.
For $Var(X)$, notice that

$$\begin{aligned}
E[X^2] &= E[E[X^2 \mid Y]] \\
&= \sum_{y=0}^{1} E[X^2 \mid Y=y]p_Y(y) \\
&= (1-p)E[X^2 \mid Y=0] + pE[X^2 \mid Y=1] \\
&= (1-p)E[(1+X)^2] + p(1)^2 && \text{from above} \\
&= (1-p)E[1+2X+X^2] + p \\
&= (1-p)(1+2E[X]+E[X^2]) + p \\
&= 1 + 2(1-p)E[X] + (1-p)E[X^2] \\
\Rightarrow & E[X^2](1-(1-p)) = 1 + \frac{2(1-p)}{p} \\
\Rightarrow & E[X^2] = \frac{1}{p} + \frac{2(1-p)}{p^2}
\end{aligned}$$

So we have

$$\begin{aligned}
Var(X) &= E[X^2] - E[X]^2 \\
&= \frac{1}{p} + \frac{2(1-p)}{p^2} - \frac{1}{p^2} \\
&= \frac{p+2-2p-1}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}$$

Remark 8.1. For law of total expectations, a large part of it is choosing the right random variable to condition on (i.e. $Y = Bern(p)$ in this example).

8.2 Theorem 2.3 (variance as expectation of conditionals)

Theorem 8.1. For random variables X and Y

$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$

Proof. Recall that

$$Var(X \mid Y=y) = E[X^2 \mid Y=y] + E[X \mid Y=y]^2$$

so more generally we have

$$Var(X \mid Y) = E[X^2 \mid Y] + E[X \mid Y]^2$$

Taing the expectation of this

$$\begin{aligned} E[Var(X \mid Y)] &= E[E[X^2 \mid Y] - E[X \mid Y]^2] \\ &= E[E[X^2 \mid Y]] - E[E[X \mid Y]^2] \\ &= E[X^2] - E[E[X \mid Y]^2] && E[A] = E[E[A \mid B]] \text{ (law of total expectation)} \end{aligned}$$

Note that

$$Var(E[X \mid Y]) = Var(v(Y))$$

where $v(Y) = E[X \mid Y]$ is a function of Y (not X!).

$$\begin{aligned} Var(v(Y)) &= E[v(Y)^2] - E[v(Y)]^2 \\ &= E[E[X \mid Y]^2] - E[X]^2 && \text{law of total expectation} \end{aligned}$$

Therefore we have

$$\begin{aligned} E[Var(X \mid Y)] + Var(E[X \mid Y]) &= E[X^2] - E[E[X \mid Y]^2] + E[E[X \mid Y]^2] - E[X]^2 \\ &= E[X^2] - E[X]^2 \\ &= Var(X) \end{aligned}$$

as desired. □

8.3 Example 2.9 solution

Suppose $\{X_i\}_{i=1}^{\infty}$ is an iid sequence of random variables with common mean μ and variance σ^2. Let N be a discrete, non-negative integer-valued rv that is independent of each X_i.
Find the mean and variance of $T = \sum_{i=1}^{N} X_i$ (referred to as a **random sum**).

Solution. To find the mean:
We condition on N since the value of our T depends on how many X_i's there are which depends on N. By the law of total expectations

$$E[T] = E[E[T \mid N]]$$

Note that

$$\begin{aligned} E[T \mid N = n] &= E[\sum_{i=1}^{N} X_i \mid N = n] \\ &= E[\sum_{i=1}^{n} X_i \mid N = n] \\ &= \sum_{i=1}^{n} E[X_i \mid N = n] && \text{due to independence of } X_i \text{ and } N \\ &= \sum_{i=1}^{n} E[X_i] \\ &= n\mu \end{aligned}$$

So we have $E[T \mid N] = N\mu$.

Remark 8.2. We needed to first condition on a concrete $N = n$ in order to unwrap the summation, then revert back to the random variable N.

Thus we have

$$E[T] = E[E[T \mid N]] = E[N\mu] = \mu E[N]$$

which intuitively makes sense.
To find the variance:
We use our previous theorem on variance as expectation of conditionals

$$Var(T) = E[Var(T \mid N)] + Var(E[T \mid N])$$

We know from before that

$$Var(E[T \mid N]) = Var(N\mu) = \mu^2 Var(N)$$

We can break apart the variance as

$$\begin{aligned}
Var(T \mid N = n) &= Var(\sum_{i=1}^{N} X_i \mid N = n) \\
&= Var(\sum_{i=1}^{n} X_i \mid N = n) \\
&= Var(\sum_{i=1}^{n} X_i \\
&= \sum i = 1^n Var(X_i) && \text{independence of } X_i \\
&= \sigma^2 n
\end{aligned}$$

Therefore $Var(T \mid N) Var(T \mid N = n)\Big|_{n=N} = \sigma^2 N$.
So

$$E[Var(T \mid N)] = E[\sigma^2 N] = \sigma^2 E[N]$$

and thus

$$Var(T) = \sigma^2 E[N] + \mu^2 Var(N)$$

9 January 25, 2018

9.1 Example 2.10 solution ($P(X < Y)$)

Suppose X and Y are independent continuous random variables. Find an expression for $P(X < Y)$.

Solution. Define our event of interest as

$$A = \{X < Y\}$$

Thus we have

$$\begin{aligned}
P(X<Y)=P(A) &= \int_{-\infty}^{\infty} P(A \mid Y=y) f_Y(y)dy && \text{law of total probability}\\
&= \int_{-\infty}^{\infty} P(X<Y \mid Y=y) f_Y(y)dy\\
&= \int_{-\infty}^{\infty} P(X<y \mid Y=y) f_Y(y)dy\\
&= \int_{-\infty}^{\infty} P(X<y) f_Y(y)dy && X<y \text{ only depends on X; } Y=y \text{ only depends on Y}\\
&= \int_{-\infty}^{\infty} P(X\le y) f_Y(y)dy && X \text{ is a continuous rv}\\
&= \int_{-\infty}^{\infty} F_X(y) f_Y(y)dy
\end{aligned}$$

Suppose that X and Y have the same distribution. We expect $P(X<Y)=\frac{1}{2}$. Let's verify it with our expression

$$\begin{aligned}
P(X<Y) &= \int_{-\infty}^{\infty} F_X(y) f_Y(y)dy\\
&= \int_{-\infty}^{\infty} F_Y(y) f_Y(y)dy && X \sim Y
\end{aligned}$$

Let $u=F_Y(y)$, thus $\frac{du}{dy}=f_Y(y) \iff du=f_Y(y)dy$. So we have

$$\begin{aligned}
P(X<Y) &= \int_0^1 udu && \text{domain for a CDF is } [0,1]\\
&= \frac{u^2}{2}\bigg|_0^1\\
&= \frac{1}{2}
\end{aligned}$$

9.2 Example 2.11 solution ($P(X<Y)$ where $X \sim EXP(\lambda_1)$ and $Y \sim EXP(\lambda_2)$

Suppose $X \sim Exp(\lambda_1)$ and $Y \sim Exp(\lambda_2)$ are independent exponential rvs. Show that

$$P(X<Y)=\frac{\lambda_1}{\lambda_1+\lambda_2}$$

Solution. Since $Y \sim Exp(\lambda_2)$, then we have $f_Y(y)=\lambda_2 e^{-\lambda y}$ for $y>0$.
Since $X \sim Exp(\lambda_1)$, we have

$$\begin{aligned}
F_X(x)=P(X\le x) &= \int_0^x \lambda_1 e^{-\lambda_1 x} dx\\
&= -e^{-\lambda_1 x}\bigg|_0^x\\
&= 1-e^{-\lambda_1 x} \quad x\ge 0
\end{aligned}$$

From the expression in Example 2.10, we have

$$\begin{aligned}
P(X < Y) &= \int_0^\infty F_X(y) f_Y(y) dy \\
&= \int_0^\infty (1 - e^{-\lambda_1 y})(\lambda_2 e^{-\lambda_2 y}) dy \\
&= \int_0^\infty \lambda_2 e^{-\lambda y} - \lambda_2 e^{-(\lambda_1+\lambda_2)y} dy \\
&= \int_0^\infty \lambda_2 e^{-\lambda y} + \frac{\lambda_2}{\lambda_1+\lambda_2} e^{-(\lambda_1+\lambda_2)y} \Big|_0^\infty \qquad = 1 - \frac{\lambda_2}{\lambda_1+\lambda_2} \\
&= \frac{\lambda_1}{\lambda_1+\lambda_2}
\end{aligned}$$

9.3 Example 2.12 solution

Consider an experiment in which independent trials each having probability $p \in (0,1)$ are performed until $k \in \mathbb{Z}^+$ consecutive successes are achieved. Determined the expected number of trails for k consecutive successes.

Solution. Let N_k be the rv which counts the number of trials needed to obtain k consecutive successes.
Current goal: we want to find $E[N_k]$.
Note: when $n = 1$, then we have $N_1 \sim GEO(p)$, and so $E[N_1] = \frac{1}{p}$.
For arbitrary $k \geq 2$, we will try to find $E[N_k]$ using the law of total expectations, namely

$$E[N_k] = E[E[N_k \mid W]]$$

for some W rv we *choose carefully.*
Suppose we choose W where (we will later see why this won't work)

$$W = \begin{cases} 0 & \text{if first trial is a failure} \\ 1 & \text{if first trial is a success} \end{cases}$$

So we have

$$\begin{aligned}
E[N_k] &= \sum_w E[N_k \mid W = w] P(W = w) \\
&= P(W = 0) E[N_k \mid W = 0] + P(W = 1) E[N_k \mid W = 1] \\
&= (1-p) E[N_k \mid W = 0] + p E[N_k \mid W = 1]
\end{aligned}$$

Note that

$$\begin{aligned}
&N_k \mid (W = 0) \sim 1 + N_k \\
&N_k \mid (W = 1) \sim ?
\end{aligned}$$

We can't simply have $N_k \mid (W = 1) \sim 1 + N_{k-1}$ since N_{k-1} does not guarantee that the $k-1$ consecutive successes are followed immediately after our first $W = 1$.
Perhaps we need another W, $W = N_{k-1}$ so we attempt to find

$$E[N_k] = E[E[N_k \mid N_{k-1}]]$$

Consider

$$E[N_k \mid N_{k-1} = n]$$

conditional on $N_{k-1} = n$, defin

$$Y = \begin{cases} 0 & \text{if the } (n+1)\text{th trial is a failure} \\ 1 & \text{if the } (n+1)\text{th trial is a success} \end{cases}$$

Now we have

$$\begin{aligned} E[N_k \mid N_{k-1} = n] &= \sum_y E[N_k \mid N_{k-1} = n, Y = y] P(Y = y \mid N_{k-1} = n) \\ &= P(Y = 0 \mid N_{k-1} = n) E[N_k \mid N_{k-1} = n, Y = 0] \\ &+ P(Y = 1 \mid N_{k-1} = n) E[N_k \mid N_{k-1} = n, Y = 1] \\ &= (1-p) E[N_k \mid N_{k-1} = n, Y = 0] + pE[N_k \mid N_{k-1} = n, Y = 1] \quad Y \text{ is independent from } N_{k-1} \end{aligned}$$

Note that

$$\begin{aligned} N_k \mid (N_{k-1} = n \mid Y = 0) &\sim n + 1 + N_k \text{we need to start over again} \\ N_k \mid (N_{k-1} = n \mid Y = 1) &\sim n + 1 \text{ with probability 1} \end{aligned}$$

Therefore

$$\begin{aligned} E[N_k \mid N_{k-1} = n] &= (1-p)(n + 1 + E[N_k]) + p(n+1) \\ &= n + 1 + (1-p)E[N_k] \end{aligned}$$

which in terms of the rv N_{k-1}

$$E[N_k \mid N_{k=1}] = E[N_k \mid N_{k-1} = n]\Big|_{n = N_{k-1}} = N_{k-1} + 1 + (1-p)E[N_k]$$

Thus from the law of total expectations

$$\begin{aligned} E[N_k] &= E[E[N_k \mid N_{k-1}]] \\ &= E[N_{k-1} + 1 + (1-p)E[N_k]] \\ &= E[N_{k-1}] + 1 + (1-p)E[N_k] \\ \Rightarrow E[N_k] &= \frac{1}{p} + \frac{E[N_{k-1}]}{p} \end{aligned}$$

This is a recurrence relation for $k = 2, 3, 4, \ldots$. To solve, we check for some k values to gain some intuition

$$\begin{aligned} &k = 2 \Rightarrow E[N_2] = \frac{1}{p} + \frac{E[N_1]}{p} = \frac{1}{p} + \frac{1}{p^2} \\ &k = 3 \Rightarrow E[N_3] = \frac{1}{p} + \frac{E[N_2]}{p} = \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} \\ &\vdots \\ &E[N_k] = \sum_{i=1}^{k} \frac{1}{p^i} \quad k = 1, 2, 3, \ldots \qquad \text{by induction} \end{aligned}$$

This is the finite geometric series for $r = \frac{1}{p}$, thus we have

$$E[N_k] = \frac{\frac{1}{p} - \frac{1}{p^{k+1}}}{1 - \frac{1}{p}}$$

10 Tutorial 3

10.1 Mixed conditional distribution

Suppose X is $Erlang(n, \lambda)$ with pdf

$$f_X(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \quad x > 0$$

Suppose $Y \mid (X = x)$ is $POI(x)$ with pmf

$$p_{Y|X}(y \mid x) = \frac{e^{-x} x^y}{y!} \quad y = 0, 1, 2, \ldots$$

Find the condition distribution $X \mid (Y = y)$.

Solution. The marginal distribution of Y is characterized by its pmf

$$\begin{aligned}
p_Y(y) &= \int_{-\infty}^{\infty} p_{Y|X}(y \mid x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} p_{Y|X}(y \mid x) f_X(x) dx \\
&= \int_0^{\infty} \frac{e^{-x} x^y}{y!} \cdot \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} dx \\
&= \frac{\lambda^n}{y!(n-1)!} \int_0^{\infty} x^{n+y-1} e^{-(\lambda+1)x} dx \\
&= \frac{\lambda^n (n+y-1)!}{(\lambda+1)^{n+y} y!(n-1)!} \int_0^{\infty} \frac{(\lambda+1)^{n+y} x^{n+y-1} e^{-(\lambda+1)x}}{(n+y-1)!} dx \\
&= \frac{\lambda^n (n+y-1)!}{(\lambda+1)^{n+y} y!(n-1)!} && \text{integral of pdf } Erlang(n+y, \lambda+1) \\
&= \binom{n+y-1}{n-1} \left(\frac{\lambda}{\lambda+1}\right)^n \left(\frac{1}{\lambda+1}\right)^y \quad y = 0, 1, 2, \ldots
\end{aligned}$$

Note that $p_Y(y)$ is the Negative Binomial distribution shifted to the left n units. In other words, it counts the number of "failures" before n successes, where the probability of success if $\lambda/(\lambda+1)$.
The distribution of $X \mid (Y = y)$ is thus

$$\begin{aligned}
f_{X|Y}(x \mid y) &= \frac{p_{Y|X}(y \mid x) f_X(x)}{p_Y(y)} \\
&= \frac{\frac{e^{-x} x^y}{y!} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}}{\frac{(n+y-1)!}{y!(n-1)!} \frac{\lambda^n}{(\lambda+1)^{n+y}}} \\
&= \frac{(\lambda+1)^{n+y} x^{n+y-1} e^{-(\lambda+1)x}}{(n+y-1)!} \quad x > 0
\end{aligned}$$

Note that $f_{X|Y}(x \mid y)$ is exactly the Erlang distribution $Erlang(n+y, \lambda+1)$.

10.2 Law of total expectations

1. Let $\{X_i\}_{i=1}^{\infty}$ an iid sequence of $EXP(\lambda)$ random variables and let $N \sim GEO(p)$ be independent of each X_i. Find $E[\prod_{i=1}^{N} X_i]$.

2. Let $\{X_i\}_{i=0}^{\infty}$ an iid sequence where $X_i \sim BIN(10, 1/2^i)$, $i = 0, 1, 2, \ldots$. Also let $N \sim POI(\lambda)$ be independent of each X_i. Find $E[X_N]$.

Solution. 1. We want to first find $E[\prod_{i=1}^{N} X_i \mid N = n]$ (conditioning on $N = n$)

$$\begin{aligned}
E[\prod_{i=1}^{N} X_i \mid N = n] &= E[\prod_{i=1}^{n} X_i \mid N = n] \\
&= E[\prod_{i=1}^{n} X_i] && \text{independence of } X_i's \text{ and } N \\
&= \prod_{i=1}^{n} E[X_i] && \text{independence of } X_i's \\
&= \prod_{i=1}^{n} \frac{1}{\lambda} \\
&= \frac{1}{\lambda^n}
\end{aligned}$$

Thus by the law of total expectations

$$\begin{aligned}
E[\prod_{i=1}^{N} X_i] &= E[E[\prod_{i=1}^{N} X_i \mid N = n]] \\
&= \sum_{n=1}^{\infty} \frac{1}{\lambda^n}(1-p)^{n-1}p \\
&= \frac{p}{\lambda^n} \sum_{n=1}^{\infty} (1-p)^{n-1} \\
&= \frac{p}{\lambda} \sum_{n=1}^{\infty} \left(\frac{1-p}{\lambda}\right)^{n-1} \\
&= \frac{p}{\lambda(1-\frac{1-p}{\lambda})} \sum_{n=1}^{\infty} \left(\frac{1-p}{\lambda}\right)^{n-1} \left(1 - \frac{1-p}{\lambda}\right) \\
&= \frac{p}{\lambda(1-\frac{1-p}{\lambda})} && \text{summation of pmf of } GEO(\frac{1-p}{\lambda}) \\
&= \frac{p}{\lambda - 1 + p}
\end{aligned}$$

provided that $\frac{1-p}{\lambda} < 1$ or $1 - p < \lambda$.

2. Condition on $N = n$ we have

$$E[X_N \mid N = n] = E[X_n \mid N = n] = E[X_n] = 10 \cdot \frac{1}{2^n} = \frac{10}{2^n}$$

From the law of total expectations

$$\begin{aligned} E[X_N] &= E[E[X_N \mid N = n]] \\ &= \sum_{n=0}^{\infty} \frac{10}{2^n} \frac{e^{-\lambda}\lambda^n}{n!} \\ &= 10e^{-\lambda/2} \sum_{n=0}^{\infty} \frac{e^{-\lambda/2}\left(\frac{\lambda}{2}\right)^n}{n!} \\ &= 10e^{-\lambda/2} && \text{summation of pmf of } POI(\lambda/2) \end{aligned}$$

10.3 Conditioning on wins and losses

A, B, C are evenly matched tennis players. Initially A and B play a set, and the winner plays C. The winner of each set continues playing the waiting player until one player wins two sets *in a row*. What is the probability that A is the overall winner?

Solution. Key idea: we condition on wins and losses each time until we can find some sort of recurrent relationship, eliminating trivial cases along the way.
Let A denote the event that A is the overall winner, and W_i, L_i denote that A wins or loses game i, respectively. Then we have

$$P(A) = P(W_1)P(A \mid W_1) + P(L_1)P(A \mid L_1) = \frac{1}{2}P(A \mid W_1) + \frac{1}{2}P(A \mid L_1)$$

We can then continue conditioning on subsequent games and their possible outcomes

$$\begin{aligned} P(A \mid W_1) &= \frac{1}{2}P(A \mid W_1, W_2) + \frac{1}{2}P(A \mid W_1, L_2) \\ &= \frac{1}{2}(1) + \frac{1}{2}[P(A \mid W_1, L_2, C \text{ wins}) \\ &+ \frac{1}{2}P(A \mid W_1, L_2, C \text{ loses})] \\ &= \frac{1}{2} + \frac{1}{4}P(A \mid W_1, L_2, C \text{ loses}) && P(A \mid W_1, L_2, C \text{ wins}) = 0 \\ & && \text{since C wins twice in a row} \\ &= \frac{1}{2} + \frac{1}{4}[\frac{1}{2}P(A \mid W_1, L_2, C \text{ loses}, W_4) \\ &+ \frac{1}{2}P(A \mid W_1, L_2, C \text{ loses}, L_4)] \\ &= \frac{1}{2} + \frac{1}{8}P(A \mid W_1, L_2, C \text{ loses}, W_4) && P(A \mid W_1, L_2, C \text{ loses}, L_4) = 0 \\ & && \text{since B wins twice in a row} \\ &= \frac{1}{2} + \frac{1}{8}P(A \mid W_1) && \text{since the probability is the same as} \\ & && \text{A winning its second game after a win} \end{aligned}$$

Solving this recurrence we get $P(A \mid W_1) = \frac{8}{14}$. Similarly

$$\begin{aligned}
P(A \mid L_1) &= \frac{1}{2}P(A \mid L_1, B \text{ wins}) + \frac{1}{2}P(A \mid L_1, B \text{ loses}) \\
&= \frac{1}{2}[\frac{1}{2}P(A \mid L_1, B \text{ loses}, W_3) + \frac{1}{2}P(A \mid L_1, B \text{ loses}, L_3)] && P(A \mid L_1, B \text{ wins}) = 0 \\
& && \text{since B wins twice in a row} \\
&= \frac{1}{4}P(A \mid L_1, B \text{ loses}, W_3) && P(A \mid L_1, B \text{ loses}, L_3) = 0 \\
& && \text{since C wins twice in a row} \\
&= \frac{1}{4}P(A \mid W_1)
\end{aligned}$$

So $P(A \mid L_1) = \frac{2}{14}$.
Plugging this into our initial equation we get $P(A) = \frac{5}{14}$.

11 February 1, 2018

11.1 Example 3.1 solution

A particle moves along the state $[0, 1, 2]$ according to a DTMC whose TPM is given by

$$P = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

where P_{ij} is the transition probability $P(X_n = j \mid X_{n-1} = i)$.
Let X_n denote the position of the particle after the nth move. Suppose the particle is likely to start in any of the three states.

1. Calculate $P(X_3 = 1 \mid X_0 = 0)$.
2. Calculate $P(X_4 = 2)$.
3. Calculate $P(X_6 = 0, X_4 = 2)$.

Solution. 1. We wish to determine $P^{(3)}_{0,1}$. To get this, we proceed to calculate $P^{(3)} = P^3$. So we have

$$P^3 = (P^2)P = \begin{bmatrix} 0.54 & 0.26 & 0.2 \\ 0.2 & 0.36 & 0.44 \\ 0.6 & 0.1 & 0.3 \end{bmatrix} \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{bmatrix} \tag{11.1}$$

$$= \begin{bmatrix} 0.478 & 0.264 & 0.258 \\ 0.36 & 0.256 & 0.384 \\ 0.57 & 0.18 & 0.25 \end{bmatrix} \tag{11.2}$$

So $P(X_3 = 1 \mid X_0 = 0) = P^{(3)}_{0,1} = 0.264$.

2. We wish to find $\alpha_{4,2} = P(X_4 = 2)$. So

$$\begin{aligned}
\alpha_4 &= (\alpha_{4,0}, \alpha_{4,1}, \alpha_{4,2}) \\
&= \alpha_0 P^{(4)} \\
&= (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) P^3 P \\
&= (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \begin{bmatrix} 0.478 & 0.264 & 0.258 \\ 0.36 & 0.256 & 0.384 \\ 0.57 & 0.18 & 0.25 \end{bmatrix} \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{bmatrix} \\
&= (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \begin{bmatrix} 0.4636 & 0.254 & 0.2824 \\ 0.444 & 0.2256 & 0.3304 \\ 0.524 & 0.222 & 0.254 \end{bmatrix} \\
&= (0.4772, 0.233867, 0.288933)
\end{aligned}$$

So we have $P(X_4 = 2) = 0.288933$.

3. We wish to calculate $P(X_6 = 0, X_4 = 2)$, which is

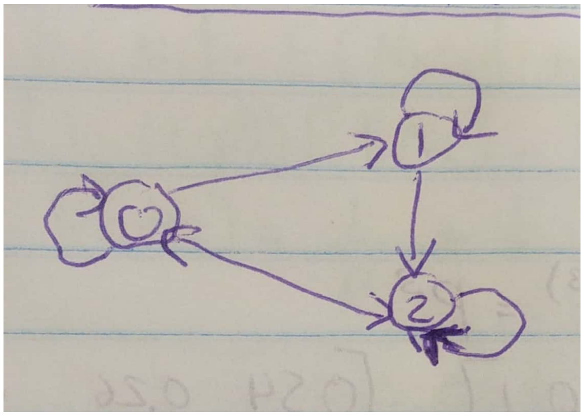
$$\begin{aligned}
P(X_6 = 0, X_4 = 2) &= P(X_4 = 2) P(X_6 = 0 \mid X_4 = 2) \\
&= (0.288433) P(X_2 = 0 \mid X_0 = 2) && \text{by stationary assumption} \\
&= (0.288433) P^{(2)}_{2,0} \\
&= (0.288433)(0.6) \\
&= 0.1733598
\end{aligned}$$

Continued: what are the equivalence classes of the DTMC?

Solution. Remember we have the TPM

$$P = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

To answer questions of this nature, it is useful to draw a **statement transition diagram**



We see that all states communicate with each other (there is some path from state i to j and vice versa). Thre is only one equivalence class, namely $\{0, 1, 2\}$. This is an **irreduicble DTMC**.

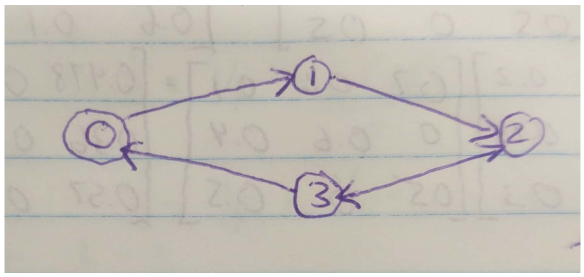
11.2 Example 3.2 solution

Consider a DTMC with TPM

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}$$

What are the equivalence classes of this DTMC?

Solution. Using a state diagram we have



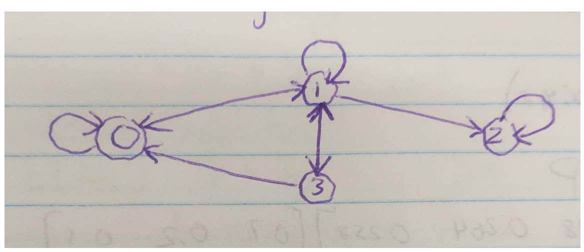
From the diagram, there is only one equivalence class $\{0, 1, 2, 3\}$. This DTMC is irreducible.

11.3 Example 3.3 solution

Consider a DTMC with TPM

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & 1 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \end{bmatrix}$$

What are the equivalence classes of this DTMC?



Solution. From the state diagram there are two equivalence classes: $\{2\}$ and $\{0, 1, 3\}$. Thus this DTMC is not irreducible.

12 Tutorial 4

12.1 Law of total expectations with indicator variables

Suppose the number of people who get on the ground floor of an elevator follows $POI(\lambda)$. If there are m floors above the ground floor and if each person is equally likely to get off at each of the m floors, independent of where the others get off, calculate the expected number of stops the elevator will make before discharging all passengers.

Solution. Let X_i denote the whether or not someone gets off at floor i, that is

$$X_i = \begin{cases} 0 & \text{no one gets off at floor } i \\ 1 & \text{someone gets off at floor } i \end{cases}$$

It is easier to think of the case where no one gets off at a floor. That is for N people

$$P(X_i) = (1 - \frac{1}{m})^N$$

Since X_i is bernoulli, we have

$$E[X_i \mid N = n] = 1 - (1 - \frac{1}{m})^n$$

Let $X = X_1 + \ldots + X_m$ denote the total number of stops for the elevator. Thus we have

$$\begin{aligned} E[X] = E[E[X \mid N = n]] &= \sum_{n=0}^{\infty} E[\sum_{i=1}^{m} X_i \mid N = n] p_N(n) \\ &= \sum_{n=0}^{\infty} m(1 - (1 - \frac{1}{m})^n) \cdot \frac{e^{-\lambda}\lambda^n}{n!} \\ &= m[\sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} - \sum_{n=0}^{\infty} \frac{e^{-\lambda}((1-\frac{1}{m})\lambda)^n}{n!}] \\ &= m[1 - e^{\frac{-\lambda}{m}} \sum_{n=0}^{\infty} \frac{e^{-(1-\frac{1}{m})\lambda}((1-\frac{1}{m})\lambda)^n}{n!}] \\ &= m(1 - e^{\frac{-\lambda}{m}}) \end{aligned}$$

where the second last and third last equality follows from the summation of Poisson distributions.

12.2 Discrete time Markov chain urn example

Three white and three black balls are distributed in two urns in such a way that each contains three balls. We say that the system is in state i, $i = 0, 1, 2, 3$, if the first urn contains i white balls. At each step, we draw one ball from each urn and place the ball drawn from the first urn into the second, and conversely with the ball from the second urn. Let X_n denote the state of the system after the nth step. Explain why $\{X_n, n = 0, 1, 2, \ldots\}$ is a Markov chain and calculate its transition probability matrix.

Solution. Since the determination of the number of white balls in the first urn in the nth step is only dependent on the same information in the $(n-1)$th step, the process $\{X_n, n = 0, 1, 2, \ldots\}$ satisfies the Markov property assumption, and hence is a Markov chain.
Let Urn 1 contain i white balls and $3 - i$ black balls (and Urn 2 has $3 - i$ white balls and i black balls). There are 9 possible ways to choose a ball from each so we do a case analysis on what color ball is chosen from each

Urn 1	Urn 2	X_n	X_{n+1}	
W	B	i	$i-1$	with probability $i^2/9$
W	W	i	i	with probability $i(3-i)/9$
B	B	i	i	with probability $(3-i)i/9$
B	W	i	$i+1$	with probability $(3-i)^2/9$

So for $i = 0, 1, 2, 3$ we have

$$P_{i,i-1} = \frac{i^2}{9} \quad P_{i,i} = \frac{2i(3-i)}{9} \quad P_{i,i+1}\frac{(3-i)^2}{9}$$

thus we have the TPM

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} & 0 \\ 0 & \frac{4}{9} & \frac{4}{9} & \frac{1}{9} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

12.3 Discrete time Markov chain weather example

On a given day, the weather is either clear (C), overcast (O), or raining (R). If the weather is clear today, then it will be C, O, or R tomorrow with respective probabilities $0.6, 0.3, 0.1$. If the weather is overcast today, then it will be C, O, or R tomorrow with probabilities $0.2, 0.5, 0.3$. If the weather is raining today, then it will be C, O, or R tomorrow with probabilities $0.4, 0.2, 0.4$. Construct the one-step transition probability matrix and use it to find α_1, $P^{(2)}$, and the probability that it rains on both the first and third days. Assume that the initial probability row vector is given by $a_0 = (0.5, 0.3, 0.2)$.

Solution. Let state $0, 1, 2$ be the states for clear, overcast, and raining. Then we can easily construct the TPM from the given probabilities of transitioning to the other states given an initial state

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.2 & 0.4 \end{bmatrix}$$

To find α_1, we simply take $\alpha_0 P$ or

$$\alpha_1 = (0.5, 0.3, 0.2) \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.2 & 0.4 \end{bmatrix} = (0.44, 0.34, 0.22)$$

To find $P^{(2)}$, we take PP or

$$\begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.2 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.46 & 0.35 & 0.19 \\ 0.34 & 0.37 & 0.29 \\ 0.44 & 0.30 & 0.26 \end{bmatrix}$$

We know the probability it rains on the first day $\alpha_{1,2}$ or 0.22.

Remark 12.1. We can't just take $\alpha_{3,2}$ as the probability it rains on the third day then apply the multiplication rule since we have some prior that the first day also rains. So we must use conditional probability here.

Thus we have

$$\begin{aligned} P(X_3 = 2, X_1 = 2) &= P(X_3 = 2 \mid X_1 = 2)P(X_1 = 2) \\ &= P(X_2 = 2 \mid X_0 = 2)P(X_1 = 2) && \text{stationary assumption} \\ &= P^{(2)}_{2,2}\alpha_{1,2} = (0.26)(0.22) = 0.0572 \end{aligned}$$

Where $P(X_2 = 2 \mid X_0 = 2) = P^{(2)}_{2,2}$ since we already know day 0 has state 2.

13 February 6, 2018

13.1 Example 3.4 solution

Consider the DTMC with TPM

$$P = \begin{bmatrix} \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & 1 & 0 \\ \frac{3}{4} & 0 & 0 & \frac{1}{4} \end{bmatrix}$$

which has equivalence classes $\{0,3\}$, $\{1\}$, and $\{2\}$. Determine the period of each state.

Solution. Consider the state 0. There are many paths which we can go from state 0 to 0 in n steps, but one obvious one is simply going from 0 to 0 n times which has probability $(P_{0,0})^n$. Therefore

$$P_{0,0}^{(n)} \geq (P_{0,0})^n = (1/3)^n > 0 \quad \forall n \in \mathbb{Z}^+$$

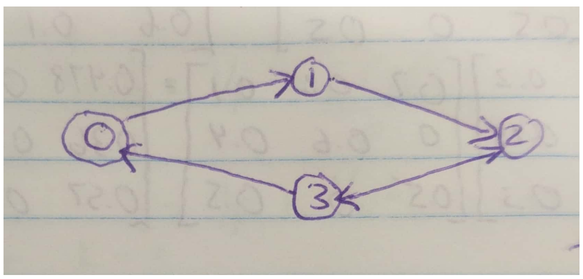
Thus $d(0) = \gcd\{n \in \mathbb{Z}^+ \mid P_{0,0}^{(n)} > 0\} = \gcd\{1,2,3,\ldots\} = 1$ (there is a way to get from 0 to 0 in any $n \in \mathbb{Z}^+$ steps, so we take the gcd of $\mathbb{Z}^+$ which is 1).
In fact, since every term on the main diagonal of P is positive, the same argument holds for every state. Thus $d(1) = d(2) = d(3) = 1$.

13.2 Example 3.2 (continued) solution

Recall for the DTMC with TPM

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}$$

there is 1 equivalence class $\{0,1,2,3\}$ with state diagram



(note the bi-direction between 2 and 3). Determine the period for each state.

Solution. We see that in one of the loops $0 \to 1 \to 2 \to 3 \to 0$ we have

$$P_{0,0}^{(n)} > 0 \text{ for } n = 4, 8, 12, 16, \ldots$$

Also we have (for the cycle $0 \to 1 \to 2 \to 3 \to 2 \to 3 \to 0$)

$$P_{0,0}^{(n)} > 0 \text{ for } n = 6, 10, 14, 18, \ldots$$

Thus $d(0) = \gcd\{4, 6, 8, 10, 12\} = 2$.

Following a similar line of logic, we find

$$\begin{aligned} d(1) &= \gcd\{4,6,8,10,12,\ldots\} = 2 \\ d(2) &= \gcd\{2,4,6,8,10,12,\ldots\} = 2 \\ d(3) &= \gcd\{2,4,6,8,10,12,\ldots\} = 2 \end{aligned}$$

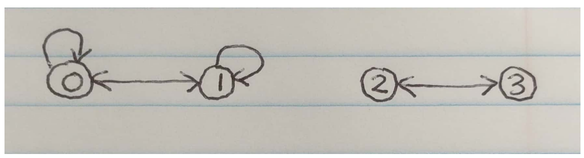
13.3 Example 3.5 solution

Consider the DTMC with TPM

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 2/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Find the equivalence class of this DTMC and determine the period of each state.

Solution. To determine the equivalence class we draw the state diagram



Clearly the equivalence classes are $\{0,1\}$ and $\{2,3\}$.
As in Example 3.4, the main diagonal terms for rows 0 and 1 are positive (i.e. $P_{0,0}, P_{1,1} > 0$) and so $d(0) = d(1) = 1$.
For states 2 and 3 the DTMC will continually alternate (with probability 1) between each other at every step (i.e. $2 \to 3 \to 2 \to 3 \to \ldots$). Thus it is clear that

$$\begin{aligned} d(2) &= \gcd\{n \in \mathbb{Z}^+ \mid P_{2,2}^{(n)} > 0\} = \gcd\{2,4,6,8,\ldots\} = 2 \\ d(3) &= \gcd\{n \in \mathbb{Z}^+ \mid P_{3,3}^{(n)} > 0\} = \gcd\{2,4,6,8,\ldots\} = 2 \end{aligned}$$

13.4 Theorem 3.1 (equivalent states have equivalent periods)

Theorem 13.1. If $i \leftrightarrow j$ (they communicate), then $d(i) = d(j)$.

Proof. Assume $i \neq j$. Since $i \leftrightarrow j$, we know by definition that $P_{i,j}^{(n)} > 0$ for some $n \in \mathbb{Z}^+$ and $P_{j,i}^{(m)} > 0$ for some $m \in \mathbb{Z}^+$. Moreover since state i is accessible from state j and state j is accessible from state i, $\exists s \in \mathbb{Z}^+$ such that $P_{j,j}^{(s)} > 0$.
Clearly we have that

$$P_{i,i}^{(n+m)} \geq P_{i,j}^{(n)} \cdot P_{j,i}^{(m)} > 0$$

(paths that take n steps to i to j then m steps to j to i is one such possible path from i to i in $n+m$ steps. There could be more $n+m$ paths hence the $\geq$. This also follows from the Chapman-Kolmogorov equations.)
In addition,

$$P_{i,i}^{(n+s+m)} \geq P_{i,j}^{(n)} \cdot P_{j,j}^{(s)} \cdot P_{j,i}^{(m)} > 0$$

So we have paths with $n+m$ and $n+s+m$ steps, thus $d(i)$ divides both $n+m$ and $n+s+m$. Therefore it follows that $d(i)$ divides their differencely, namely $(n+s+m)-(n+m) = s$. Since this holds true for *any* s which satisfies $P_{j,j}^{(s)} > 0$, then it must be the case that $d(i)$ divides $d(j)$.
Using the same line of logic, it is straightforward to show $d(j)$ divides $d(i)$.

Putting these two arguments together, we deduce that $d(i) = d(j)$. □

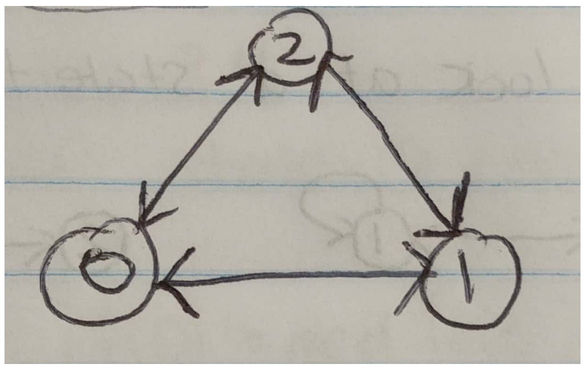
13.5 Example 3.6 solution

Consider the DTMC with TPM

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

Find the equivalence classes and determine the period of the states.

Solution. The state transition diagram looks like



Clearly the DTMC is irreducible (i.e. there is just one class $\{0, 1, 2\}$).
Note that

$$P_{0,0}^{(1)} = 0$$
$$P_{0,0}^{(2)} \geq P_{0,1}P_{1,0} = \left(\frac{1}{2}\right)^2 = \frac{1}{4} > 0$$
$$P_{0,0}^{(3)} \geq P_{0,1}P_{1,2}P_{2,0} = \left(\frac{1}{2}\right)^3 = \frac{1}{8} > 0$$

Clearly $d(0) = \gcd\{2, 3, \ldots\} = 1$.
By the above theorem, we have $d(0) = d(1) = d(2) = 1$.

14 February 8, 2018

14.1 Theorem 3.2 (communication and recurrent state i implies recurrent state j)

Theorem 14.1. If $i \leftrightarrow j$ (communicate) and state i is recurrent, then state j is recurrent.

Proof. Since $j \leftrightarrow j$, $\exists m, n \in \mathbb{N}$ such that

$$P_{i,j}^{(m)} > 0$$
$$P_{j,i}^{(n)} > 0$$

Also since state i is recurrent, then we have

$$\sum_{l=1}^{\infty} P_{i,i}^{(l)} = \infty$$

Suppose that $s \in \mathbb{Z}^+$. Note that

$$P_{j,j}^{(n+s+m)} \geq P_{j,i}^{(m)} \cdot P_{i,i}^{(s)} \cdot P_{i,j}^{(n)}$$

Now to show that j is recurrent, we show that the following series diverges

$$\begin{aligned}
\sum_{k=1}^{\infty} P_{j,j}^{(k)} &\geq \sum_{k=n+m+1}^{\infty} P_{j,j}^{(k)} \\
&= \sum_{s=1}^{\infty} P_{j,j}^{(n+s+m)} \\
&\geq \sum_{s=1}^{\infty} P_{j,i}^{(m)} \cdot P_{i,i}^{(s)} \cdot P_{i,j}^{(n)} \\
&= P_{j,i}^{(m)} \cdot P_{i,j}^{(n)} \sum_{s=1}^{\infty} P_{i,i}^{(s)} \\
&= \infty
\end{aligned}$$

since $P_{j,i}^{(m)}, P_{i,j}^{(n)} > 0$ and the series diverges by our premise. Therefore j is recurrent. $\square$

Remark 14.1. A by-product of the above theorem is that if $i \leftrightarrow j$ and state i is transient, then state j is transient.

14.2 Theorem 3.3 (communication and recurrent state i implies mutual recurrence among all states)

Theorem 14.2. If $i \leftrightarrow j$ and state i is recurrent, then

$$f_{i,j} = P(\text{DTMC ever makes a future visit to state } j \mid X_0 = i) = 1$$

Proof. Clearly the result is true if $i = j$. Therefore suppose that $i \neq j$. Since $i \leftrightarrow j$, the fact that state i is recurrent implies that state j is recurrent by the previous theorem and $f_{j.j} = 1$.
To prove $f_{i,j} = 1$, suppose that $f_{i,j} < 1$ and try to get a contradiction.
Since $i \leftarrow j$, $\exists n \in \mathbb{Z}^+$ such that $P_{j,i}^{(n)} > 0$ i.e. each time the DTMC visits state j, there is the possibility of being in state i n time units later with probability $P_{j,i}^{(n)} > 0$.
If we are assuming that $f_{i,j} < 1$, then this implies that the probability of returning to state j after visiting i in the future is not guaranteed (as $1 - f_{i,j} > 0$). Therefore

$$\begin{aligned}
1 - f_{j,j} = P(\text{DTMC never makes a future visit to state } j \mid X_0 = j) & \\
&= P_{j,i}^{(n)} \cdot (1 - f_{i,j}) \\
&> 0 \qquad \text{both } > 0
\end{aligned}$$

This implies that $1 - f_{j,j} > 0$ or $f_{j,j} < 1$, which is a contradiction. Therefore $f_{i,j} = 1$. $\square$

14.3 Theorem 3.4 (finite-state DTMCs have at least one recurrent state)

Theorem 14.3. A finite-state DTMC has at least one recurrent state.

Proof. Equivalently, we want to show that not all states can be transient.
Suppose that $\{0, 1, 2, \ldots, N\}$ represents the states of the DTMC where $N < \infty$ (finite).

To prove that not all states can be transient, we suppose they are all transient and try to get a contradiction.
Now for each $i = 0, 1, \ldots, N$, if state i is assumed to be transient, we know that after a *finite* amount of time (denoted by T_i), state i will never be visited again. As a result, after a finite amount of time $T = \{T_0, T_1, \ldots, T_n\}$ has gone by *none of the states will be visited again.*
However, the DTMC **must be in some state** after time T but we have exhausted all states fro the DTMC to be in. This is a contradiction thus not all states can be transient in a finite state DTMC. □

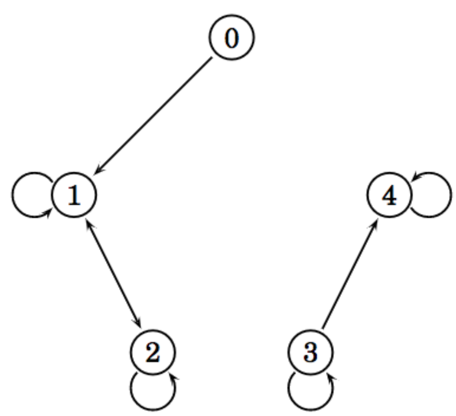
15 Tutorial 5

15.1 Determining diagram, equivalence classes, period, and transience/recurrence of DTMC

For the following Markov chain, draw its state transition diagram, determine its equivalence classes, and the periods of states within each class, and determine whether they are transient or recurrent

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{5} & \frac{4}{5} & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution. The state diagram for the DTMC is



From the diagram we see the only closed loop path containing more than one state is $1 \to 2 \to 1$, thus we have the equivalence classes $\{0\}, \{1, 2\}, \{3\}, \{4\}$.
Clearly $0 \to 1$ with a probability of 1 so 0 is transient and its period is ∞.
Notice for states $1, 2, 3, 4$ the main diagonal entries $P_{i,i} > 0$ so their periods are 1.

For transience/recurrence, note that

$$
\begin{aligned}
\sum_{n=1}^{\infty} P_{0,0}^{(n)} &= \sum_{n=1}^{\infty} 0 && < \infty \Rightarrow \{0\} \text{ is transient} \\
\sum_{n=1}^{\infty} P_{3,3}^{(n)} &= \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n = \frac{\frac{3}{4}}{1-\frac{3}{4}} = 3 && < \infty \Rightarrow \{3\} \text{ is transient} \\
\sum_{n=1}^{\infty} P_{4,4}^{(n)} &= \sum_{n=1}^{\infty} 1 && = \infty \Rightarrow \{4\} \text{ is recurrent}
\end{aligned}
$$

For class $\{1,2\}$, consider the values of $f_{1,1}^{(n)}$

$$
\begin{aligned}
f_{1,1}^{(1)} &= P(X_1 = 1 \mid X_0 = 1) && = P_{1,1} = \frac{2}{3} \\
f_{1,1}^{(2)} &= P(X_2 = 1, X_1 \neq 1 \mid X_0 = 1) = P(X_2 = 1, X_1 = 2 \mid X_0 = 1) && = P_{1,2}P_{2,1} = \frac{1}{3} \cdot \frac{1}{5} \\
f_{1,1}^{(3)} &= P(X_3 = 1, X_2 \neq 1, X_1 \neq 1 \mid X_0 = 1) \\
&= P(X_3 = 1, X_2 = 2, X_1 = 2 \mid X_0 = 1) && = P_{1,2}P_{2,2}P_{2,1} = \frac{1}{3} \cdot \frac{4}{5} \cdot \frac{1}{5} \\
f_{1,1}^{(4)} &= P(X_4 = 1, X_3 = 2, X_2 = 2, X_1 = 2 \mid X_0 = 1) && = P_{1,2}(P_{2,2})^2 P_{2,1} = \frac{1}{3} \cdot \left(\frac{4}{5}\right)^2 \cdot \frac{1}{5} \\
f_{1,1}^{(n)} &= P(X_n = 1, X_{n-1} = 2, \dots, X_2 = 2, X_1 = 2 \mid X_0 = 1) && = P_{1,2}(P_{2,2})^{n-2} P_{2,1} = \frac{1}{3} \cdot \left(\frac{4}{5}\right)^{n-2} \cdot \frac{1}{5}
\end{aligned}
$$

So we have

$$
\begin{aligned}
f_{1,1} &= \sum_{n=1}^{\infty} f_{1,1}^{(n)} \\
&= \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{5} \sum_{n=2}^{\infty} \left(\frac{4}{5}\right)^{n-2} \\
&= \frac{2}{3} + \frac{1}{15} \sum_{n=0}^{\infty} \left(\frac{4}{5}\right)^{n} \\
&= \frac{2}{3} + \frac{1}{15} \frac{1}{1-\frac{4}{5}} \\
&= \frac{2}{3} + \frac{1}{3} \\
&= 1
\end{aligned}
$$

which by definition implies class $\{1,2\}$ is recurrent.

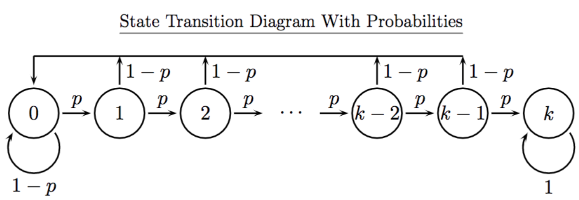
15.2 Discrete time Markov chain consecutive successes example

Let $k \in \mathbb{Z}^+$ and consider an experiment in which independent trials, each having success probability $p \in (0,1)$, are conducted indefinitely. We say that the system is in state i at time n if after the nth trial we have observed i consecutive successes occurring at times $n-i+1, n-i+2, \dots, n-1, n$ where $i = 1, 2, \dots, k-1$. In addition, we say that the system is in state k at time n if we have ever observed at least k consecutive successes occur at least once by time n. Finally, if we have not observed k consecutive successes by time n, and the nth trial was a failure, then we say that the system is in state 0. Letting X_n denote the state of the system after the nth trial, calculate its transition probability matrix.

Solution. Note at any state $k = 1, 2, \dots, k-1$ we always have a probability of $1-p$ of failing and going back to state 0.
To transition from state k to state $k+1$ successes, we must get a success on the next trial with probability p, unless it's already in state k.
Finally, once we reach state k (we have seen k consecutive successes), any successes or failures will not change the system's state at k, so $P_{k,k} = 1$.
Thus we have

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ k-2 \\ k-1 \\ k \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & \dots & k-2 & k-1 & k \end{array} \\ \begin{bmatrix} 1-p & p & 0 & \dots & 0 & 0 & 0 \\ 1-p & 0 & p & \dots & 0 & 0 & 0 \\ 1-p & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & & & & & & \\ 1-p & 0 & 0 & \dots & 0 & p & 0 \\ 1-p & 0 & 0 & \dots & 0 & 0 & p \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} \end{array}$$

The corresponding state diagram of this DTMC is



15.3 Limiting behaviour of discrete Markov chains

For a DTMC $\{X_n, n \in \mathbb{N}\}$ suppose we are given initial probability row vector $\alpha_0 = (1, 0)$ and TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1-p & p \\ 0 & 1 \end{bmatrix} \end{array}$$

where $0 < p < 1$. Find $P^{(n)}$ and α_n.

Solution. We know the n-step TPM is simply the one-step TPM multipled by itself n times or

$$P^{(n)} = \prod_{i=1}^{n} \begin{bmatrix} 1-p & p \\ 0 & 1 \end{bmatrix}$$

Note that

$$
\begin{aligned}
P_{0,0}^{(n)} &= \sum_{k=0}^{1} P_{0,k}^{(n-1)} P_{k,0} = P_{0,0}^{(n-1)} P_{0,0} + P_{0,1}^{(n-1)} P_{1,0} = P_{0,0}^{(n-1)}(1-p) && P_{1,0} = 0 \\
&= P_{0,0}^{(n-2)}(1-p)^2 = \ldots = P_{0,0}^{(1)}(1-p)^{n-1} = (1-p)^n \\
P_{1,0}^{(n)} &= \sum_{k=0}^{1} P_{1,k}^{(n-1)} P_{k,0} = P_{1,0}^{(n-1)} P_{0,0} + P_{1,1}^{(n-1)} P_{1,0} = P_{1,0}^{(n-1)}(1-p) && P_{1,0} = 0 \\
&= P_{1,0}^{(n-2)}(1-p)^2 = \ldots = P_{1,0}^{(1)}(1-p)^{n-1} = 0(1-p)^{n-1} = 0 \\
P_{1,1}^{(n)} &= \sum_{k=0}^{1} P_{1,k}^{(n-1)} P_{k,1} = P_{1,0}^{(n-1)} P_{0,1} + P_{1,1}^{(n-1)} P_{1,1} = P_{1,1}^{(n-1)} 1 = \ldots = P_{1,1}^{(1)} = 1
\end{aligned}
$$

for all $n \in \mathbb{N}$. Using the fact that each row sum of a TPM must equal 1 we have

$$P_{0,1}^{(n)} = 1 - P_{0,0}^{(n)} = 1 - (1-p)^n \quad n \in \mathbb{N}$$

So we have

$$P^{(n)} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \left[\begin{array}{cc} (1-p)^n & 1-(1-p)^n \\ 0 & 1 \end{array}\right] \end{array} \quad n \in \mathbb{N}$$

and by definition

$$\alpha_n = \alpha_0 P^{(n)} = (1,0) \begin{bmatrix} (1-p)^n & 1-(1-p)^n \\ 0 & 1 \end{bmatrix} = ((1-p)^n, 1-(1-p)^n)$$

16 February 13, 2018

16.1 Theorem 3.5 (recurrent i and not communicate with j implies $P_{i,j} = 0$)

Theorem 16.1. If state i is recurrent and state i does not communicate with state j, then $P_{i,j} = 0$.

Proof. Let us assume that $P_{i,j} > 0$ (one-step transition probability) and try to get a contradiction.
Then $P_{j,i}^{(n)} = 0 \ \forall n \in \mathbb{Z}^+$ otherwise states i, j communicate.
However, the DTMC, starting in state i, would have a positive probability of at least $P_{i,j}$ of never returning to state i. This contradicts the recurrence of state i, therefore we must have $P_{i,j} = 0$. ☐

16.2 Example 3.3 (continued) solution

Recall our earlier DTMC with TPM

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & 1 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \end{bmatrix}$$

Determine whether each state is transient or recurrent.

Solution. Note that $\sum_{n=1}^{\infty} P_{2,2}^{(n)} = \sum_{n=1}^{\infty} 1 = \infty$ which implies state 2 is recurrent.

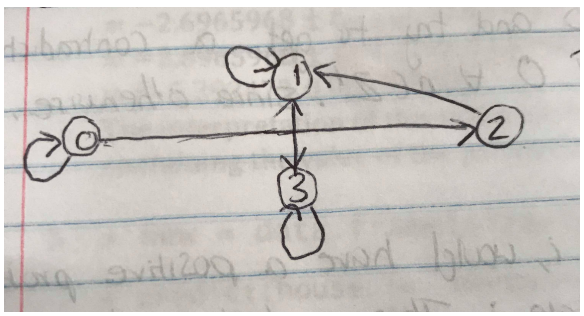
Looking at the possible transitions that can take place among states 0, 1, and 3 we strongly suspect state 1 to be transient (since there is a positive probability of never returning to state 1 if a transition to state 2 occurs). To show this formally assume state 1 is recurrent and try to get a contradiction.
So if state 1 is recurrent, note that state 1 does not communicate with state 2. By theorem 3.5, we have that $P_{1,2} = 0$, but in fact $P_{1,2} = \frac{1}{8} > 0$. This is a contradiction, so state 1 must be transient and so $\{0, 1, 3\}$ is transient.

16.3 Example 3.7 solution

Consider a DTMC with TPM

$$P = \begin{bmatrix} \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 & 0 \\ 0 & \frac{2}{5} & 0 & \frac{3}{5} \end{bmatrix}$$

Determine whether each state is transient or recurrent.



Solution. The equivalence classes are $\{0\}, \{2\}, \{1, 3\}$.
Note that (since 0 is in its own equivalence classes)

$$\sum_{n=1}^{\infty} P_{0,0}^{(n)} = \sum_{n=1}^{\infty} (\frac{1}{4})^n = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3} < \infty$$

so state 0 is transient.
Furthermore (since state 2 is in its own equivalence class)

$$\sum_{n=1}^{\infty} P_{2,2}^{(n)} = \sum_{n=1}^{\infty} 0 = 0 < \infty$$

so state 2 is also transient.
On the other hand, concerning the class $\{1, 3\}$ we observe

$$f_{1,1}^{(1)} = P_{1,1} = \frac{1}{3}$$

AND

$$f_{1,1}^{(n)} = \sum_{n=2} (2/3)(3/5)^{n-2}(2/5) \quad n \geq 2$$

Now

$$\begin{aligned} f_{1,1} = \sum_{n=1}^{\infty} f_{1,1}^{(n)} &= \frac{1}{3} \sum_{n=2}^{\infty} (\frac{2}{3})(\frac{3}{5})^{n-2}(\frac{2}{5}) \\ &= \frac{1}{3} + (\frac{2}{3})(\frac{2}{5}) \sum_{n=2} (\frac{3}{5})^{n-2} \\ &= \frac{1}{3} + (\frac{2}{3})(\frac{2}{5}) \cdot \frac{1}{1-\frac{3}{5}} \\ &= 1 \end{aligned}$$

By definition since $f_{1,1} = 1$ state 1 is recurrent and so class $\{1,3\}$ is recurrent.
(Or we could have concluded from Theorem 3.4 that because $\{0\}$ and $\{2\}$ are both transient, $\{1,3\}$ must be recurrent).

16.4 Example 3.8 solution

Consider a DTMC $\{X_n, n \in \mathbb{N}\}$ whose state space is all integers i.e. $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$. Suppose the TPM for the DTMC satisfies

$$P_{i,i-1} = 1 - p \text{ and } P_{i,i+1} = p \quad \forall i \in \mathbb{Z} \text{ and } 0 < p < 1$$

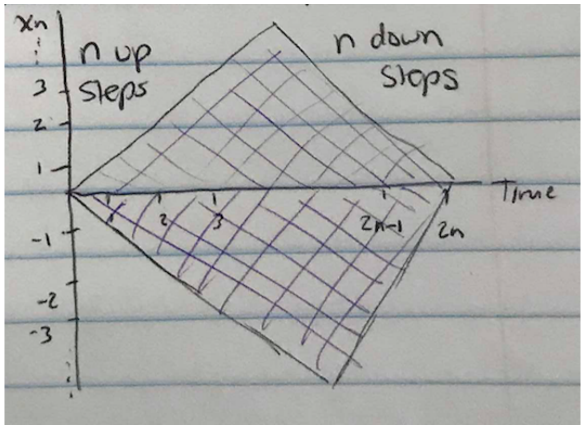
In other words from any state, we can either jump up or down one state. This is often referred to as the *Random Walk*. Characterize the behaviour of this DTMC in terms of its equivalence classes, periodicty, and transience/recurrence.

Solution. First of all since $0 < p < 1$ all states clearly communicate with each other. This implies that $\{X_n, n \in \mathbb{N}\}$ is an irreducible DTMC.
Hence we can determine its periodicity (and likewise its transcience/recurrence) by analyzing any state we wish.
Let us select state 0. Starting from state 0, note that state 0 cannot possibly visited in an odd number of transitions since we are guaranteed to have the number of up(down) jumps exceed the number of down(up) jumps (try to brute force this to see why).
Thus $P_{0,0}^{(1)} = P_{0,0}^{(3)} = P_{0,0}^{(5)} = \ldots = 0$ or equivalent

$$P_{0,0}^{(2n-1)} = 0 \quad \forall n \in \mathbb{Z}^+$$

However since it is clearly possible to return to state 0 in an even number of transitions, it immediately follows that $P_{0,0}^{(2n)} > 0 \ \forall n \in \mathbb{Z}^+$.
Hence $d(0) = gcd\{n \in \mathbb{Z}^+ \mid P_{0,0}^{(n)} > 0\} = gcd\{2, 4, 6, \ldots\} = 2$.
Finally to determine whether state 0 is transient or recurrent, let us consider

$$\sum_{n=1}^{\infty} P_{0,0}^{(n)} = \sum_{n=1}^{\infty} P_{0,0}^{(2n)} = \sum_{n=1}^{\infty} \binom{2n}{n} p^n (1-p)^n$$



That is we need to take n up steps with probability p and n down steps with probability $1-p$.
Note that $\binom{2n}{n}$ accounts for the number of ways to arrange this.
Recall: ratio test for series.
Suppose that $\sum_{n=1}^{\infty} a_n$ is a series of positive terms and

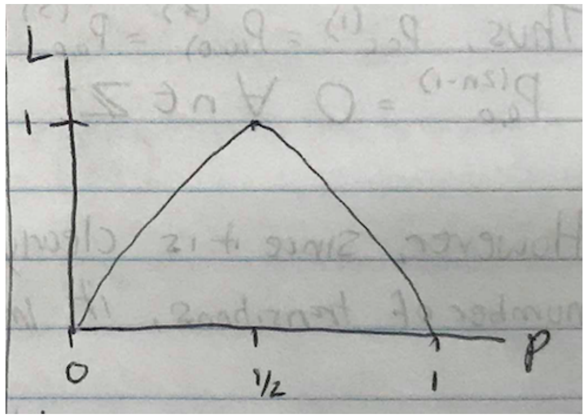
$$L = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$$

1. If $L < 1$ then the series converges.
2. If $L > 1$ then the series diverges.
3. If $L = 1$ then the test is inconclusive.

In our case we obtain

$$\begin{aligned}
L &= \lim_{n\to\infty} \frac{\frac{(2(n+1))!}{(n+1)!(n+1)!} \cdot p^{n+1}(1-p)^{n+1}}{\frac{2n!}{n!n!} \cdot p^n(1-p)^n} \\
&= \lim_{n\to\infty} \frac{(2n+2)!}{(n+1)!(n+1)!} \cdot \frac{n!n!}{2n!} \cdot p(1-p) \\
&= \lim_{n\to\infty} \frac{(2n+2)(2n+1)}{(n+1)(n+1)} \cdot p(1-p) \\
&= \lim_{n\to\infty} \frac{4n^2+6n+2}{n^2+2n+1} \cdot p(1-p) \\
&= 4p(1-p)
\end{aligned}$$

A plot of $L = 4p(1-p)$ reveals the following shape



Note that if $p \neq \frac{1}{2}$ then $L < 1$ which means state 0 is transient (and so is DTMC).
However, if $p = \frac{1}{2}$ then $L = 1$ and the ratio test is inconclusive. We need another approach to handle $p = \frac{1}{2}$ (this approach can handle both $p = \frac{1}{2}$ and $p \neq \frac{1}{2}$).
Recall that $f_{i,j} = P(\text{DTMC ever makes a future visit to state } j \mid X_0 = i)$.
Letting $q = 1 - p$ and conditioning on the state of the DTMC at time 1 we have

$$\begin{aligned}
f_{0,0} &= P(\text{DTMC ever makes a future visit to state } 0 \mid X_0 = 0) \\
&= P(X_1 = -1 \mid X_0 = 0) \cdot P(\text{DTMC ever makes a future visit to state } 0 \mid X_1 = -1, X_0 = 0) \\
&+ P(X_1 = 1 \mid X_0 = 0) \cdot P(\text{DTMC ever makes a future visit to state } 0 \mid X_1 = 1, X_0 = 0) \\
&\text{by Markov property (only most recent state matters) and stationarity (we can treat time 1 as time 0)} \\
&= q \cdot f_{-1,0} + p \cdot f_{1,0}
\end{aligned}$$

Using precisely the same logic it follows that

$$f_{1,0} = q \cdot 1 + p \cdot f_{2,0} \tag{16.1}$$

Consider $f_{2,0}$: in order to visit state 0 from state 2, we first have to visit state 1 again, which happens with probability $f_{2,1}$. This is **equivalent to** $f_{1,0}$, so $f_{2,1} = f_{1,0}$ (probability of ever making a visit to the state one step down).
Given you make a visit back to state 1, then we must make a visit back to state 0, which happens to with probability $f_{1,0}$.
Putting this together we have

$$f_{2,0} = f_{2,1} \cdot f_{1,0} = f_{1,0}^2$$

Thus from equation 16.1 we have

$$f_{1,0} = q + p \cdot f_{1,0}^2$$

which is a quadratic in the form

$$p f_{1,0}^2 - f_{1,0} + q = 0$$

Applying the quadratic formula we get

$$\begin{aligned}
f_{1,0} &= \frac{1 \pm \sqrt{1-4pq}}{2p} \\
&= \frac{1 \pm \sqrt{(p+q)^2 - 4pq}}{2p} \\
&= \frac{1 \pm \sqrt{p^2 - 2pq + q^2}}{2p} \\
&= \frac{1 \pm |p-q|}{2p}
\end{aligned}$$

There can only be **one unique solution** for $f_{1,0}$ which means that one of

$$\begin{aligned}
r_1 &= \frac{1+|p-q|}{2p}; \text{ or} \\
r_2 &= \frac{1-|p-q|}{2p}
\end{aligned}$$

must be inadmissible (intuitively: we see that if $p < 0.5$ then r_1 will be bigger than 1 which is not possible for probabilities so r_1 does not work for all p).
To determine which it is, suppose that $q > p$. Then $|p-q| = -(p-q)$ and the 2 roots become

$$\begin{aligned}
r_1 &= \frac{1-(p-q)}{2p} = \frac{q+q}{2p} = \frac{q}{p} > 1 \\
r_2 &= \frac{1+(p-q)}{2p} = \frac{p+p}{2p} = 1
\end{aligned}$$

Thus we must have

$$f_{1,0} = \frac{1-|p-q|}{2p}$$

Note: using the exact same approach

$$f_{-1,0} = \frac{1-|p-q|}{2q}$$

With knowledge of $f_{1,0}$ and $f_{-1,0}$ we find that

$$\begin{aligned}
f_{0,0} &= qf_{-1,0} + pf_{1,0} \\
&= q(\frac{1-|p-q|}{2q}) + p(\frac{1-|p-q|}{2p}) \\
&= 1 - |p-q|
\end{aligned}$$

Suppose $p > q$ i.e. $1-q > q$ or $2q < 1$:

$$\begin{aligned}
f_{0,0} &= 1-(p-q) \\
&= 1-p+q \\
&= 2q < 1
\end{aligned}$$

since $f_{0,0} < 1$ we have that state 0 is transient.
Similarly for $p < q$ we get $f_{0,0} = 2p < 1$ so state 0 is transient.

When $p = q$, then we have $f_{0,0} = 1$ so state 0 is recurrent under $p = q = \frac{1}{2}$.

17 February 15, 2018

17.1 Example 3.9 solution

Consider the DTMC with TPM

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Determine if $\lim_{n\to\infty} P^{(n)}$ exists (i.e. the limiting behavior of the DTMC).

Solution. There are 2 equivalence classes, namely $\{0, 2\}$ and $\{1\}$.
Each class is recurrent with periods 2 and 1, respectively.
For $n \in \mathbb{Z}^+$, note that

$$P^{(2n)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which follows by taking $P^{(2)}$ and multiplying by itself arbitrarily many times. Also

$$P^{(2n-1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

which follows by taking $P^{(1)}$ and multiply it arbitrarily many times by $P^{(2)}$ (identity).
As such, $\lim_{n\to\infty} P^{(n)}$ does not exist since $P^{(n)}$ alternates between those two matrices..
However, while $\lim_{n\to\infty} P^{(n)}_{0,0}$ and $\lim_{n\to\infty} P^{(n)}_{0,2}$ (and many other) do not exist, note that some limits do exist such as

$$\lim_{n\to\infty} P^{(n)}_{0,1} = 0$$

and

$$\lim_{n\to\infty} P^{(n)}_{1,1} = 1$$

17.2 Example 3.10 solution

Consider the DTMC with TPM

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Determine if $\lim_{n\to\infty} P^{(n)}$ exists.

Solution. There is only one equivalence class and so the DTMC is irreducible. Also it is straightforward to show that the DTMC is aperiodic and recurrent.
It can be shown that

$$\lim_{n\to\infty} P^{(n)} = \begin{bmatrix} \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \end{bmatrix}$$

(we can find this by using some software and repeatedly applying matrix exponentiation; we will later demonstrate a way to solve this for certain DTMCs (see Example 3.10 (continued) solution (finding limiting probability using BLT))).

Note that this matrix has **identical rows**. This implies that $P_{i,j}^{(n)}$ converges to a value as $n \to \infty$ which is the same for all initial states i. In other words, there is a limiting probability the DTMC will be in state j as $n \to \infty$ and this probability is indepdendent of the initial states.

18 Tutorial 6

18.1 Determining $f_{i,i}$ and recurrence using definition (infinite summation) from TPM

Consider a DTMC $\{X_n, n = 0, 1, 2, \ldots\}$ with

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{5} & \frac{1}{2} & \frac{3}{10} & 0 & 0 & 0 \\ \frac{2}{5} & 0 & \frac{3}{5} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 1 \end{bmatrix} \end{array}$$

It is easy to show there are three equivalence classes $\{0, 1, 2, 3\}$, $\{4\}$, and $\{5\}$. Calculate $f_{0,0}$ to show that state 0 is recurrent (and hence its equivalence class is recurrent).

Solution. Recall that

$$f_{i,i}^{(n)} = P(X_n = i, X_{n-1} \neq i, \ldots, X_2 \neq i, X_1 \neq i \mid X_0 = i)$$

As $f_{0,0} = \sum_{n=1}^{\infty} f_{0,0}^{(n)}$ we consider the possible paths starting from state 0 that requires exactly n steps to return to step 0 for the first time.

$$\begin{aligned} n = 1 \quad &: \quad 0 \to 0 \\ n = 2 \quad &: \quad 0 \to 1 \to 0 \\ n \geq 3 \quad &: \quad 0 \to 1 \to \ldots \to 1 \to 0 \\ & \quad\;\; 0 \to 1 \to \ldots \to 1 \to 2 \to \ldots \to 2 \to 0 \\ & \quad\;\; 0 \to 3 \to 2 \to \ldots \to 2 \to 0 \end{aligned}$$

Thus we can compute $f_{0,0}^{(n)}$

$$
\begin{aligned}
f_{0,0}^{(1)} &= P_{0,0} = \frac{1}{4} \\
f_{0,0}^{(2)} &= P_{0,1}P_{1,0} = \frac{1}{2}\frac{1}{5} = \frac{1}{10} \\
f_{0,0}^{(n)} &= P_{0,1}P_{1,1}^{n-2}P_{1,0} + \sum_{k=0}^{n-3} P_{0,1}P_{1,1}^{k}P_{1,2}P_{2,2}^{n-3-k}P_{2,0} + P_{0,3}P_{3,2}P_{2,2}^{n-3}P_{2,0} \\
&= \frac{1}{2}\left(\frac{1}{2}\right)^{n-2}\frac{1}{5} + \frac{1}{2}\frac{3}{10}\frac{2}{5}\sum_{k=0}^{n-3}\left(\frac{1}{2}\right)^{k}\left(\frac{3}{5}\right)^{n-3-k} + \frac{1}{4}\frac{2}{5}\left(\frac{3}{5}\right)^{n-3} \\
&= \frac{1}{10}\left(\frac{1}{2}\right)^{n-2} + \frac{3}{50}\sum_{k=0}^{n-3}\left(\frac{1}{2}\right)^{k}\left(\frac{3}{5}\right)^{n-3-k} + \frac{1}{10}\left(\frac{3}{5}\right)^{n-3} \qquad n \geq 3
\end{aligned}
$$

So we have

$$
\begin{aligned}
f_{0,0} &= \sum_{n=1}^{\infty} f_{0,0}^{(n)} \\
&= \frac{1}{4} + \frac{1}{10} + \sum_{n=3}^{\infty}\left\{\frac{1}{10}\left(\frac{1}{2}\right)^{n-2} + \frac{3}{50}\sum_{k=0}^{n-3}\left(\frac{1}{2}\right)^{k}\left(\frac{3}{5}\right)^{n-3-k} + \frac{1}{10}\left(\frac{3}{5}\right)^{n-3}\right\} \qquad n \geq 3 \\
&= \frac{1}{4} + \frac{1}{10} + \frac{1}{20}\sum_{m=0}^{\infty}\left(\frac{1}{2}\right)^{m} + \frac{3}{50}\sum_{m=0}^{\infty}\sum_{k=0}^{m}\left(\frac{1}{2}\right)^{k}\left(\frac{3}{5}\right)^{m-k} + \frac{1}{10}\sum_{m=0}^{\infty}\left(\frac{3}{5}\right)^{m} \\
&= \frac{1}{4} + \frac{1}{10} + \frac{1}{20}\frac{1}{1-\frac{1}{2}} + \frac{3}{50}\sum_{k=0}^{\infty}\sum_{m=k}^{\infty}\left(\frac{1}{2}\right)^{k}\left(\frac{3}{5}\right)^{m-k} + \frac{1}{10}\frac{1}{1-\frac{3}{5}} \\
&= \frac{1}{4} + \frac{1}{10} + \frac{1}{10} + \frac{3}{50}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^{k}\sum_{m=k}^{\infty}\left(\frac{3}{5}\right)^{m-k} + \frac{1}{4} \\
&= \frac{1}{4} + \frac{1}{10} + \frac{1}{10} + \frac{3}{50}\frac{1}{1-\frac{1}{2}}\frac{1}{1-\frac{3}{5}} + \frac{1}{4} \\
&= \frac{9}{20} + \frac{3}{10} + \frac{1}{4} \\
&= 1
\end{aligned}
$$

as desired.

18.2 Making random walk recurrent with a reflecting boundary

Consider the DTMC $\{X_n, n \in \mathbb{N}\}$ represent the random walk model, but now with a reflect boundary at state 0 such that the state space contains only the non-negative integers. Specifically we are now assuming the TPM for the DTMC satisfies

$$P_{i,i-1} = 1-p \quad P_{i,i+1} = p \quad \forall i \in \mathbb{Z}^+$$

and $P_{0,1} = 1$ where $0 < p < 1$. Determine the values of p that result in the DTMC being recurrent.

Solution. It is clear to see that the DTMC is irreducible and has one equivalence class containing all the states. Thus we only need to focus on a single state. Suppose state 0 is recurrent, then $f_{0,0} = 1$. Note that state 0 transitions to state 1 with a probability of 1 ($P_{0,1} = 1$), therefore $f_{0,0} = f_{1,0} = 1$. Letting $q = 1 - p$ we can show that

$$f_{1,0} = q + pf_{2,0} = q + pf_{1,0}^2$$

where we are condition on the state at time 2 and $f_{2,0} = f_{2,1}f_{1,0} = f_{1,0}^2$. Remember that we showed in class the quadratic equation

$$pf_{1,0}^2 - f_{1,0} + q = 0$$

has one solution where

$$f_{1,0} = \frac{1 - |p - q|}{2p}$$

We plug in values for p, q. Specifically when $p \leq q$ or $|p - q| = q - p$

$$f_{1,0} = f_{0,0} = \frac{1 - (q - p)}{2p} = \frac{2p}{2p} = 1$$

For $p > q$ we have $|p - q| = p - q$ so

$$f_{0,0} = \frac{1 - (p - q)}{2p} = \frac{2q}{2p} < 1$$

Thus state 0 and the entire DTMC will be recurrent if $p \leq \frac{1}{2}$.

18.3 Determining pmf of $N_i \sim f_{i,i}^{(n)}$ and the mean number of transitions between re-visit

Consider the DTMC with the TPM

$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{bmatrix} \end{array}$$

and assume $\alpha, \beta \in (0, 1)$ so the DTMC is irreducible (thus both states are recurrent). Find the pmf of $N = \min\{n \in \mathbb{Z}^+ \mid X_n = 0\}$ conditional on $X_0 = 0$, use it to show that $f_{0,0} = 1$ and find the mean number of transitions between successive visits to state 0.

Solution. By definition the pmf of $N \mid (X_0 = 0)$ is given as $P(N = n \mid X_0 = 0) = f_{0,0}^{(n)}$ for $n \in \mathbb{Z}^+$. Thus

$$\begin{aligned} P(N = 1 \mid X_0 = 0) &= f_{0,0}^{(1)} = P_{0,0} = \alpha \\ P(N = n \mid X_0 = 0) &= f_{0,0}^{(n)} \\ &= P_{0,1}P_{1,1}^{n-2}P_{1,0} \\ &= (1 - \alpha)(1 - \beta)^{n-2}\beta \quad n \geq 2 \end{aligned}$$

To find $f_{0,0}$ note that

$$\begin{aligned}
f_{0,0} &= \sum_{n=1}^{\infty} f_{0,0}^{(n)} \\
&= \alpha + \beta(1-\alpha)\sum_{n=2}^{\infty}(1-\beta)^{n-2} \\
&= \alpha + \beta(1-\alpha)\sum_{m=0}^{\infty}(1-\beta)^{m} \\
&= \alpha + \beta(1-\alpha)\frac{1}{1-(1-\beta)} \\
&= \alpha + (1-\alpha) \\
&= 1
\end{aligned}$$

Also note that the mean number of transitions m_i between state i is defined as

$$\begin{aligned}
m_0 = E[N \mid X_0 = 0] &= \sum_{n=1}^{\infty} n f_{0,0}^{(n)} \\
&= \alpha + (1-\alpha)\beta\sum_{n=2}^{\infty} n(1-\beta)^{n-2} \\
&= \alpha + (1-\alpha)\beta\sum_{m=1}^{\infty}(m+1)(1-\beta)^{m-1} \\
&= \alpha + (1-\alpha)\beta\frac{1}{\beta}\left(\sum_{m=1}^{\infty} m(1-\beta)^{m-1}\beta + \sum_{m=1}^{\infty}(1-\beta)^{m-1}\beta\right) \\
&\text{by expectation and total probability of } GEO(\beta) \text{ we have} \\
&= \alpha + (1-\alpha)\left(\frac{1}{\beta}+1\right) \\
&= \frac{1-\alpha+\beta}{\beta}
\end{aligned}$$

Let $\alpha = 1/2$ and $\beta = 1/4$. Evaluate $E[N \mid X_0 = 0]$ and find $\lim_{n\to\infty} P^{(n)}$. What is the relationship between the mean number of transitions and the limiting probability of being in state 0?

Solution. Note that we get $m_0 = \frac{1-1/2+1/4}{1/4} = 3$. Furthermore note that we have

$$\lim_{n\to\infty} P^{(n)} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \left[\begin{array}{cc} 1/3 & 2/3 \\ 1/3 & 2/3 \end{array}\right] \end{array}$$

from R. This coincides with the Basic Limit Theorem (BLT) we proved in class for irreducible, aperiodic and recurrent DTMCs where the limiting probability π_i of being in some state i is equal to $\frac{1}{m_i}$ (intuitively, since the mean time of getting back to state 0 is 3 transitions, we are in some other state for the other two transitions. So the limiting probability of observing a DTMC in state 0 is equal to the fraction of time we spend in that state 0).

19 February 27, 2018

19.1 Example 3.11 solution

Consider the DTMC with TPM

$$P = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 \end{bmatrix}$$

Examine $\lim_{n\to 0} P^{(n)}$ and explain why the limiting probability of being in a state depends on the initial state of this DTMC.

Solution. We have three equivalence classes $\{0\}, \{1\}, \{2\}$. Since $P_{i,i} > 0\ \forall i = 0, 1, 2$ each state is aperiodic.

$$\sum_{n=1}^{\infty} P_{0,0}^{(n)} = \sum_{n=1}^{\infty} P_{2,2}^{(n)} = \sum_{n=1}^{\infty} = \infty$$

which implies state 0 and 2 are recurrent.
We can also show that

$$\sum_{n=1}^{\infty} P_{1,1}^{(n)} = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{2} \cdot \frac{1}{1-\frac{1}{2}} = 1 < \infty$$

which means state 1 is transient.
In fact, states 0 and 2 are examples of **absorbing states** (i.e. $P_{0,0} = P_{2,2} = 1$).
If one begins in state 1, note that with probability $\frac{1}{2}$ one can end up in one of the two recurrent classes in the next transition i.e. the DTMC will eventually be absorbed into either state 0 or 2 so state 1 is *transient*.
It can be shown that

$$\lim_{n\to 0} P^{(n)} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix}$$

Intuitively this makes sense: eventually state 1 will go into 0 or 2. The relative probability of going into state 0 and going into state 2 is equivalent to our limiting probability entries.

Remark 19.1. Unlike the previous example, it *does matter* from which state one begins in this DTMC. Furthermore the second column of the above contains all zeros.

19.2 Theorem 3.6 solution

Theorem 19.1. For any state i and transient state j of a DTMC, $\lim_{n\to\infty} P_{i,j}^{(n)} = 0$.

Proof. Recall that we defined the "first visit probability in n steps" as

$$f_{i,j}^{(n)} = P(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i)$$

Furthermore we defined

$$f_{i,j} = P(\text{DTMC ever makes a future visit to state } j \mid X_0 = i) = \sum_{n=1}^{\infty} f_{i,j}^{(n)}$$

We can then use the fact that

$$P_{i,j}^{(n)} = \sum_{k=1}^{n} f_{i,j}^{(k)} P_{j,j}^{(n-k)}$$

So we have

$$\begin{aligned}
\sum_{n=1}^{\infty} P_{i,j}^{(n)} &= \sum_{n=1}^{\infty}\sum_{k=1}^{n} f_{i,j}^{(k)} P_{j,j}^{(n-k)} \\
&= \sum_{k=1}^{\infty} f_{i,j}^{(k)} \sum_{n=k}^{\infty} P_{j,j}^{(n-k)} \\
&= \sum_{k=1}^{\infty} f_{i,j}^{(k)} \sum_{l=0}^{\infty} P_{j,j}^{(l)} && l = n-k \\
&= (\sum_{k=1}^{\infty} f_{i,j}^{(k)})(1 + \sum_{l=1}^{\infty} P_{j,j}^{(l)}) && P_{i,i}^{(0)} = 1 \quad \forall i \\
&= f_{i,j}(1 + \sum_{l=1}^{\infty} P_{j,j}^{(l)}) \\
&\le 1 + \sum_{l=1}^{\infty} P_{j,j}^{(l)} && f_{i,j} \le 1 \text{ since it's a probability} \\
&< \infty \text{ since } j is transient so the summation is finite
\end{aligned}$$

Recall the **nth term test**: If $\lim_{n\to 0} a_n \neq 0$ or if the limit does not exist, then $\sum_{n=1}^{\infty} a_n$ diverges.
We can show by contradiction that if the series converges then the limit is 0.
By the **nth term test** for infinite series

$$\lim_{n\to 0} P_{i,j}^{(n)} = 0$$

□

19.3 Example 3.10 (continued) solution (finding limiting probability using BLT)

Recall we were given the DTMC with TPM

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Find the limiting probability of this DTMC.

Solution. Clearly DTMC is irreducible, aperiodic and positive recurrent. The BLT conditions are met and $\pi = (\pi_0, \pi_1, \pi_2)$ exists and satisfies

$$\begin{aligned}
&\pi = \pi P \\
&\pi e' = 1 && e' = (1, 1, \ldots, 1)^T
\end{aligned}$$

Thus we have

$$(\pi_0, \pi_1, \pi_2) = (\pi_0, \pi_1, \pi_2) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Expanding the matrix multiplication we get

$$\begin{aligned}
\pi_0 &= \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1 \\
\pi_1 &= \frac{1}{2}\pi_0 + \frac{1}{4}\pi_1 + \frac{1}{3}\pi_2 \\
\pi_2 &= \frac{1}{4}\pi_1 + \frac{2}{3}\pi_2 \\
1 &= \pi_0 + \pi_1 + \pi_2
\end{aligned}$$

(we can disregard the $\pi_1 = \dots$ equation since it requires solving for the most terms and we have 4 equations for 3 unknowns). Solving the system we get

$$\begin{aligned}
\pi_0 = \pi_1 &= \frac{4}{11} \\
\pi_2 &= \frac{3}{11}
\end{aligned}$$

Recall we had

$$\lim_{n\to\infty} P^{(n)} = \begin{bmatrix} \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \\ \frac{4}{11} & \frac{4}{11} & \frac{3}{11} \end{bmatrix} = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 \\ \pi_0 & \pi_1 & \pi_2 \\ \pi_0 & \pi_1 & \pi_2 \end{bmatrix}$$

We observe that each row of $P^{(n)}$ converges to π as $n \to \infty$.

20 March 1, 2018

20.1 Example 3.12 solution (Gambler's Ruin) (applying limiting probability)

Consider a gambler who, at each play of a game, has probability $p \in (0,1)$ of winning one unit and probability $q = 1 - p$ of losing one unit. Assume that successive plays of the game are independent. If the gambler initially begins with i units, what is the probability that the gambler's fortune will reach N units ($N < \infty$) before reaching 0 units? This problem is often referred to as the Gambler's Ruin Problem, with state 0 representing bankruptcy and state N representing the jackpot.

Solution. For $n \in \mathbb{N}$, define X_n as the gambler's fortune after the nth play of the game. Therefore $\{X_n, n \in \mathbb{N}\}$ is a DTMC with TPM

$$P = \begin{bmatrix}
1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\
q & 0 & p & 0 & \dots & 0 & 0 & 0 \\
0 & q & 0 & p & \dots & 0 & 0 & 0 \\
\vdots & & & & & & & \\
0 & 0 & 0 & 0 & \dots & q & 0 & p \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 1
\end{bmatrix}$$

Where the TPM has entries P_{ij} where $i,j \in [0,N]$ (for each $i \in [1, N-1]$, we have $P_{i(i-1)} = q$ and $P_{i(i+1)} = p$. $P_{00} = P_{NN} = 1$ and all other $P_{mn} = 0$).
Note: States 0 and N are treated as **absorbing states**, therefore they are both recurrent.
States $\{1, 2, \dots, N-1\}$ belong to the same equivalence class and it is straightforward to see it is a transient class.
Goal: Determine $P(i)$ for $i = 0, 1, \dots, N$ which represents the probability that starting with i units, the gambler's fortune will eventually reach N units.

It follows that (for the limiting TPM)

$$\lim_{n\to\infty} P^{(n)} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1-P(1) & 0 & 0 & 0 & \dots & 0 & 0 & P(1) \\ 1-P(2) & 0 & 0 & 0 & \dots & 0 & 0 & P(2) \\ \vdots & & & & & & & \\ 1-P(N-1) & 0 & 0 & 0 & \dots & 0 & 0 & P(N-1) \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}$$

where the $(0,0)$ entry is 1 (0 state is recurrent), $(0,N)$ is 0 (can never reach state N from 0) and similarly for (N,N) and $(N,0)$.
Entries $(i,N) = P(i)$ by definition where $P(0) = 0$ and $P(N) = 1$. Since the sum of the entries for a given row must be 0, we get for entry $(i,0)$ as $1 - P(i)$.
Conditioning on the outcome of the very first game $i = 1, 2 \dots, N-1$, we have

$$P(i) = p \cdot P(i+1) + q \cdot P(i-1)$$

(if we go up by 1 with probability p, then the probability of winning afterwards is $P(i+1)$ i.e. $P(i+1) = P(i \mid \text{win})$ and $p = P(\text{win})$. Similarly for q and $P(i-1)$).
Note that $p + q = 1$, so we have

$$\begin{aligned} P(i) &= pP(i) + qP(i) \\ \Rightarrow p(P(i+1) - P(i)) &= q(P(i) - P(i-1)) \\ P(i+1) - P(i) &= \frac{q}{p}(P(i) - P(i-1)) \end{aligned}$$

To determine if an explicit solution exists, consider several values of i

$$\begin{aligned} &i = 1 \Rightarrow P(2) - P(1) = \frac{q}{p}(P(1) - P(0)) = \frac{q}{p}P(1) \\ &i = 2 \Rightarrow P(3) - P(2) = \frac{q}{p}(P(2) - P(1)) = \left(\frac{q}{p}\right)^2 P(1) \\ &i = 3 \Rightarrow P(4) - P(3) = \frac{q}{p}(P(3) - P(2)) = \left(\frac{q}{p}\right)^3 P(1) \\ &\vdots \\ &i = k \Rightarrow P(k+1) - P(k) = \frac{q}{p}(P(k) - P(k-1)) = \left(\frac{q}{p}\right)^k P(1) \end{aligned}$$

Note: the above k equations are linear in terms of the unknowns $P(1), P(2), \dots, P(k+1)$. Summing these k

equations yields

$$\begin{aligned}
P(k+1) - P(1) &= \sum_{i=1}^{k} \left(\frac{q}{p}\right)^i P(1) \\
\Rightarrow P(k+1) &= \sum_{i=0}^{k} \left(\frac{q}{p}\right)^i P(1) \quad k = 0, 1, \dots, N-1 \\
\Rightarrow P(k) &= \sum_{i=0}^{k-1} \left(\frac{q}{p}\right)^i P(1) \quad k = 1, 2, \dots, N
\end{aligned}$$

Applying the formula for a finite geometric series

$$P(k) = \begin{cases} \frac{1-\left(\frac{q}{p}\right)^k}{1-\frac{q}{p}} P(1) & p \neq \frac{1}{2} \\ kP(1) & p = \frac{1}{2} \end{cases}$$

Letting $k = N$, we obtain for $p \neq \frac{1}{2}$

$$1 = P(N) = \frac{1-\left(\frac{q}{p}\right)^N}{1-\frac{q}{p}} P(1) \Rightarrow P(1) = \frac{1-\frac{q}{p}}{1-\left(\frac{q}{p}\right)^N}$$

Similarly for $p = \frac{1}{2}$, we have $P(1) = \frac{1}{N}$.
So for $k = 0, 1, \dots, N$ we have

$$P(k) = \begin{cases} \frac{1-\left(\frac{q}{p}\right)^k}{1-\left(\frac{q}{p}\right)^N} & p \neq \frac{1}{2} \\ \frac{k}{N} & p = \frac{1}{2} \end{cases}$$

Think about what happens when $N \to \infty$ for each case.

21 Tutorial 7

21.1 DTMCs and N_i (minimum n for $X_n = i$), transience and recurrence, limit probabilities, number of transitions

The state space S of DTMC is the set of all non-negative integers, namely $S = \{0, 1, 2, \dots\}$. For each $i \in S$, the one-step transition probabilities have the form

$$P_{i,i+1} = \frac{\alpha}{2+i} \text{ and } P_{i,0} = \frac{2+i-\alpha}{2+i}$$

resulting in the TPM

$$P = \begin{array}{c c} & \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & \dots \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \end{array} & \begin{bmatrix} \frac{2-\alpha}{2} & \frac{\alpha}{2} & 0 & 0 & 0 & \dots \\ \frac{3-\alpha}{3} & 0 & \frac{\alpha}{3} & 0 & 0 & \dots \\ \frac{4-\alpha}{4} & 0 & 0 & \frac{\alpha}{4} & 0 & \dots \\ \frac{5-\alpha}{5} & 0 & 0 & 0 & \frac{\alpha}{5} & \ddots \\ \frac{6-\alpha}{6} & 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \end{array}$$

1. Explain why we must restrict α to the interval $0 \le \alpha \le 2$ in order for P to be a TPM when defined as above.
2. Let $N_0 = \min\{n \in \mathbb{Z}^+ \mid X_n = 0\}$ be the time until the first return to state 0. For each value of $\alpha \in [0, 2]$, find an expression for the pmf of N_0, conditional on $X_0 = 0$.
3. For each value of $\alpha \in [0, 2]$, discuss the transience and recurrence of each state of S.
4. Letting $0 < \alpha \le 2$, find the mean recurrent time of state 0 and determine if the states of this DTMC are positive or null recurrent.
5. For what values of $\alpha \in [0, 2]$ does the DTMC have a limiting distribution on S?
6. For the values of α where the Basic Limit Theorem is satisfied, find the limit probabilities $\{\pi_j\}_{j\in S}$.
7. Letting $0 < \alpha \le 2$, determine the expected number of of transitions to reach state 2, conditional on $X_0 = 0$.

Solution.

1. Note that all transitions have some probability bounded by $0 \le P_{i,j} \le 1$.

   So for $P_{i,i+1}$ we have
   $$0 \le \frac{\alpha}{2+i} \le 1 \Rightarrow 0 \le \alpha \le 2 + i$$
   since the smallest value i can be is 0, then $0 \le \alpha \le 2$.

2. Note for each value $\alpha \in [0, 2]$ and by the structure of the DTMC we have
   $$\begin{aligned} f_{0,0}^{(1)} &= P_{0,0} = 1 - \frac{\alpha}{2} \\ f_{0,0}^{(2)} &= P_{0,1}P_{1,0} = \frac{\alpha}{2} \cdot (1 - \frac{\alpha}{3}) \\ f_{0,0}^{(3)} &= P_{0,1}P_{1,2}P_{2,0} = \frac{\alpha}{2} \cdot \frac{\alpha}{3} \cdot (1 - \frac{\alpha}{4}) \\ f_{0,0}^{(n)} &= P_{0,1}P_{1,2} \dots P_{n-2,n-1}P_{n-1,0} = \frac{\alpha^{n-1}}{n!} \cdot (1 - \frac{\alpha}{n+1}) \end{aligned}$$
   for $n \in \mathbb{Z}^+$.

3. When $\alpha = 0$, it is clear that only state 0 is recurrent and all other states are transient (since $P_{i,0} = 1\ \forall i \in Z^+$).

   When $\alpha \ne 0$, we note that all states are in one equivalence class, so we only focus on the transience one state e.g. 0.

We find $f_{0,0}$ when $\alpha \neq 0$

$$\begin{aligned}
f_{0,0} = \sum_{n=1}^{\infty} f_{0,0}^{(n)} &= \sum_{n=1}^{\infty} \frac{\alpha^{n-1}}{n!}(1 - \frac{\alpha}{n+1} = \sum_{n=1}^{\infty} \frac{\alpha^{n-1}}{n!} - \sum_{n=1}^{\infty} \frac{\alpha^n}{(n+1)!} \\
&= \frac{1}{\alpha}\left(\sum_{n=1}^{\infty} \frac{\alpha^n}{n!} - \sum_{n=2}^{\infty} \frac{\alpha^n}{n!}\right) \\
&= \frac{1}{\alpha}\frac{\alpha}{1} \\
&= 1
\end{aligned}$$

So state 0 is recurrent and so are the other states, thus the DTMC is recurrent.

4. We can find the mean recurrent time m_0 of state 0 from

$$\begin{aligned}
m_0 = E[N_0 \mid X_0 = 0] &= \sum_{n=1}^{\infty} n f_{0,0}^{(n)} \\
&= \sum_{n=1}^{\infty} \frac{\alpha^{n-1}}{(n-1)!} - \sum_{n=1}^{\infty} \frac{n\alpha^n}{(n+1)!} \\
&= e^{\alpha}\left(\sum_{n=0}^{\infty} \frac{\alpha^n e^{-\alpha}}{n!} - \sum_{n=1}^{\infty} \frac{(n+1-1)\alpha^n e^{-\alpha}}{(n+1)!}\right) \\
&= e^{\alpha}\left(1 - \sum_{n=1}^{\infty} \frac{\alpha^n e^{-\alpha}}{n!} + \sum_{n=1}^{\infty} \frac{\alpha^n e^{-\alpha}}{(n+1)!}\right) \\
&= e^{\alpha}\left(1 - (1 - e^{-\alpha} + \frac{1}{\alpha}\sum_{n=2}^{\infty} \frac{\alpha^{n+2} e^{-\alpha}}{(n+2)!}\right) \\
&= e^{\alpha}(e^{-\alpha} + \frac{1}{\alpha}(1 - e^{-\alpha} - \alpha e^{-\alpha}) \\
&= \frac{e^{\alpha} - 1}{\alpha} < \infty
\end{aligned}$$

for $0 < \alpha \leq 2$. This implies state 0 is positive recurrent. Since the DTMC is irreducible, this implies every state is positive recurrent.

5. From part (3) we have two cases.

If $\alpha = 0$, then we have

$$P = P^{(n)} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & \dots \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & \ddots \\ 1 & 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \end{array}$$

Therefore the limit of the n-step TPM as $n \to 0$ has identical rows, implying there is a unique solution to the limiting probabilities equal to $\pi = (1, 0, 0, \dots, 0, \dots)$.

If $0 < \alpha \leq 2$, we know the DTMC is irreducible and positive recurrent. Also note that $d(0) = gcd\{2, 3, 4, \ldots\} = 1$ implying that state 0 is aperiodic. Thus the conditions of the Basic Limit Theorem hold true and a unique limiting distribution $\{\pi_j\}_{j \in S}$ exists.

Putting both cases together a limiting distribution exists for all $0 \leq \alpha \leq 2$.

6. From part (e), $0 < \alpha \leq 2$. From part (4) we see that $\pi_0 = \frac{\alpha}{e^\alpha - 1}$.

   We can also derive π_0. From the Basic Limit Theorem, we have

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{i,j}$$
$$\sum_{i=0}^{\infty} \pi_1 = 1$$

   Note $P_{i-1,i} > 0$ and $P_{i,0} > 0$, and all other $P_{i,j} = 0$. Thus we have

$$\begin{aligned}
\pi_0 &= \sum_{i=0}^{\infty} \pi_i P_{i,0} = \sum_{i=0}^{\infty} \pi_i \frac{2+i-\alpha}{2+i} \\
\pi_1 &= \pi_0 P_{0,1} = \frac{\alpha}{2}\pi_0 \\
\pi_2 &= \pi_1 P_{1,2} = \frac{\alpha^2}{3!}\pi_0 \\
\pi_3 &= \pi_2 P_{2,3} = \frac{\alpha^3}{4!}\pi_0 \\
&\vdots \\
\pi_i &= \frac{\alpha^i}{(i+1)!}\pi_0 \quad i \in S
\end{aligned}$$

   Along with $1 = \sum_{i \in S} \pi_i$, we have enough equations to solve for each π_i.

   Solving for π_0 we have

$$\begin{aligned}
1 = \sum_{i \in S} \pi_i &= \pi_0 + \frac{\alpha}{2!}\pi_0 + \frac{\alpha^2}{3!}\pi_0 + \ldots \\
&= \pi_0 \cdot \frac{1}{\alpha} \sum_{i=1}^{\infty} \frac{\alpha^i}{i!} \\
&= \frac{\pi_0}{\alpha}\left(\sum_{i=0}^{\infty} \frac{\alpha^i}{i!} - 1\right) \\
&= \frac{\pi_0}{\alpha}(e^\alpha - 1) \\
\Rightarrow \pi_0 &= \frac{\alpha}{e^\alpha - 1}
\end{aligned}$$

   which we recognize as $\frac{1}{m_0}$ from (4). Therefore we have

$$\pi_i = \frac{\alpha^{i+1}}{(i+1)!(e^\alpha - 1)} \quad \forall i \in S$$

7. Let T_i denote the number of transitions to state 2 starting from state i. We want to find $E[T_0]$.

   Conditioning on the state at step 1, we have

$$\begin{aligned}
E[T_0] &= E[T_0 \mid X_1 = 0]P(X_1 = 0 \mid X_0 = 0) + E[T_0 \mid X_1 = 1]P(X_1 = 1 \mid X_0 = 0) \\
&= \frac{2-\alpha}{2}(1 + E[T_0]) + \frac{\alpha}{2}(1 + E[T_1]) \\
\Rightarrow &\frac{\alpha}{2}E[T_0] = 1 + \frac{\alpha}{2}E[T_1] \\
E[T_0] &= \frac{2}{\alpha} + E[T_1]
\end{aligned}$$

   Likewise

$$\begin{aligned}
E[T_1] &= E[T_1 \mid X_1 = 0]P(X_1 = 0 \mid X_0 = 1) + E[T_1 \mid X_1 = 2]P(X_1 = 2 \mid X_0 = 1) \\
&= \frac{3-\alpha}{3}(1 + E[T_0]) + \frac{\alpha}{3}(1) \\
&= \frac{3-\alpha}{3}(1 + \frac{2}{\alpha} + E[T_1]) + \frac{\alpha}{3} \\
\Rightarrow &\frac{\alpha}{3}E[T_1] = 1 + \frac{2(3-\alpha)}{3\alpha} \\
\Rightarrow &E[T_1] = \frac{\alpha + 6}{\alpha^2}
\end{aligned}$$

   Thus

$$\begin{aligned}
E[T_0] &= \frac{2}{\alpha} + \frac{\alpha+6}{\alpha^2} \\
&= \frac{3(\alpha+2)}{\alpha^2}
\end{aligned}$$

22 March 6, 2018

22.1 Example 3.11 (continued) solution (showing absorption probability equal limiting probabilities)

Recall the earlier DTMC with TPM

$$P = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 \end{bmatrix}$$

We previously claimed

$$\lim_{n \to 0} P^{(n)} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix}$$

Show that the absoprtion probabilities from transient state 1 into states 0 and 2 are equal to $\lim_{n\to\infty} P_{1,0}^{(n)}$ and $\lim_{n\to\infty} P_{1,2}^{(n)}$, respectively.

Solution. First of all, relabel the states of the DTMC is as follows

$$
\begin{aligned}
0^* &= \text{ state 1 of the original DTMC} \\
1^* &= \text{ state 0 of the original DTMC} \\
2^* &= \text{ state 2 of the original DTMC}
\end{aligned}
$$

The new TPM corrresponding to states $\{0^*, 1^*, 2^*\}$

$$
P = \begin{array}{c} \\ 0^* \\ 1^* \\ 2^* \end{array} \begin{array}{c} \begin{array}{ccc} 0^* & 1^* & 2^* \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}
$$

Based on the notation used in class, we have

$$
\begin{aligned}
Q &= \begin{bmatrix} \frac{1}{2} \end{bmatrix} \\
R &= \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \end{bmatrix} \\
O &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
I &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

Using the definition of absorption probability

$$
U_{i,k} = R_{i,k} + \sum_{j=0}^{M-1} Q_{i,j} U_{j,k}
$$

we get

$$
\begin{aligned}
U_{0^*,1^*} &= R_{0^*,1^*} + Q_{0^*,0^*} U_{0^*,1^*} \\
&= \frac{1}{3} + \frac{1}{2} U_{0^*,1^*} \\
\Rightarrow U_{0^*,1^*} &= \frac{2}{3} = \lim_{n\to\infty} P_{1,0}^{(n)}
\end{aligned}
$$

And similarly,

$$
U_{0^*,2^*} = \frac{1}{3} = \lim_{n\to\infty} P_{1,2}^{(n)}
$$

22.2 Example 3.13 solution (solving absorption probabilities)

Consider the DTPMC with TPM

$$
P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}
$$

Suppose the DTMC starts off in state 1. What is the probability that the DTMC ultimately ends up in state 3? How would this probability change if the DTMC begins in state 0 with probability $\frac{3}{4}$ and in state 1 with probability

$\frac{1}{4}$?

Solution. First we wish to calculate $U_{1,3}$. Here,

$$Q = \begin{bmatrix} 0.4 & 0.3 \\ 0.1 & 0.3 \end{bmatrix}$$
$$R = \begin{bmatrix} 0.2 & 0.1 \\ 0.3 & 0.3 \end{bmatrix}$$
$$U = \begin{bmatrix} U_{0,2}, U_{0,3} \\ U_{1,2}, U_{1,3} \end{bmatrix}$$

Since $U = (I - Q)^{-1}R$ (matrix definition of absorption probability matrix) we need the inverse of

$$I - Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.4 & 0.3 \\ 0.1 & 0.3 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.3 \\ -0.1 & 0.7 \end{bmatrix}$$

Recall if

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

provided that $ad - bc \neq 0$. Thus we have

$$(I - Q)^{-1} = \frac{1}{0.42 - 0.03} \begin{bmatrix} 0.7 & 0.3 \\ 0.1 & 0.6 \end{bmatrix} = \begin{bmatrix} \frac{70}{39} & \frac{10}{13} \\ \frac{10}{39} & \frac{20}{13} \end{bmatrix}$$

so we get

$$U = (I - Q)^{-1}R = \begin{bmatrix} 0.589744 & 0.410256 \\ 0.512821 & 0.487179 \end{bmatrix}$$

Thus $U_{1,3} = 0.487179$.
Now under the alternate set of initial conditions

$$\begin{aligned} P(\text{DTMC ultimately in state 3}) &= P(\text{final state } 3 \mid X_0 = 0)P(X_0 = 0) + P(\text{final state } 3 \mid X_0 = 1)P(X_0 = 1) \\ &= \frac{3}{4}U_{0,3} + \frac{1}{4}U_{1,3} \\ &= \frac{3}{4}(0.410256) + \frac{1}{4}(0, 487179) \\ &\approx 0.429487 \end{aligned}$$

Alternatively we could calculate $U_{1,3}$ using the definition of absorption probability as a system of equations

$$U_{i,k} = R_{i,k} + \sum_{j=0}^{M-1} Q_{i,j}U_{j,k}$$

Set $i = 1$ and $k = 3$ to get

$$U_{1,3} = R_{1,3} + Q_{1,0}U_{0,3} + Q_{1,1}U_{1,3} = 0.3 + 0.1U_{0,3} + 0.3U_{1,3}$$

so we also need $U_{0,3}$

$$U_{0,3} = R_{0,3} + Q_{0,0}U_{0,3} + Q_{0,1}U_{1,3} = 0.1 + 0.4U_{0,3} + 0.3U_{1,3}$$

We have to solve two equations with two unknowns

$$\begin{aligned} 0.1U_{0,3} - 0.7U_{1,3} &= -0.3 \\ 0.6U_{0,3} - 0.3U_{1,3} &= 0.1 \end{aligned}$$

Multiplying the first equation by 6 and subtracting one from the other we get

$$3.9U_{1,3} = 1.9 \Rightarrow U_{1,3} = 0.4871719$$

22.3 Example 3.14 solution (absorbing states with absorbing recurrent classes)

Consider the DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 & 0 \\ 0.1 & 0.3 & 0.3 & 0.3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0 & 0.4 & 0.6 \end{bmatrix} \end{array}$$

Suppose the DTMC starts off in state 1. What is the probability that the DTMC ultimately ends up in state 3?

Solution. Goal: find $\lim_{n\to\infty} P_{1,3}^{(n)}$. We can begin grouping states 3 and 4 together into a single state 3^* resulting in the revised TPM

$$P^* = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3^* \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3^* \end{array} \\ \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

This is identical to the TPM from Example 3.13, so we have

$$U_{1,3} = 0.487179$$

Once in 3^*, the DTMC will remain in recurrent class $\{3, 4\}$ with associated TPM

$$\begin{array}{c} \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cc} 3 & 4 \end{array} \\ \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \end{array}$$

Note: for this "smaller" DTMC, the conditions of the BLT are satisfied.
We can solve for limiting probabilities π_3, π_4

$$\begin{aligned} (\pi_3, \pi_4) &= (\pi_3, \pi_4) \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \\ \pi_3 + \pi_4 &= 1 \end{aligned}$$

We can show that $\pi_3 = \frac{4}{7}$ and $\pi_4 = \frac{3}{7}$. Thus

$$\lim_{n\to\infty} P_{1,3}^{(n)} = U_{1,3^*} \cdot \pi_3 = (0.487179)(\frac{4}{7}) = 0.278388$$

Why?

$$\begin{aligned}
\lim_{n\to\infty} P_{1,3}^{(n)} &= \lim_{n\to\infty} P(X_n = 3 \mid X_0 = 1)\\
&\lim_{n\to\infty} \{P(X_n = 3, X_T = 2 \mid X_0 = 1) + P(X_n = 3, X_T = 3 \mid X_0 = 1)\}\\
&P(X_n = 3, X_T = 2) = 0 \text{ since state 2 is absorbing, we have}\\
&\lim_{n\to\infty} P(X_n = 3, X_T = 3 \mid X_0 = 1)\\
&\lim_{n\to\infty} P(X_n = 3 \mid X_T = 3^*, X_0 = 1) P(X_T = 3^* \mid X_0 = 1)\\
&\lim_{n\to\infty} P_{3^*,3}^{n-T} U_{1,3^*}\\
&\lim_{n\to\infty} P_{3,3}^{n} U_{1,3^*}\\
&U_{1,3^*} \dot{\pi}_3
\end{aligned}$$

22.4 Aside: n-step TPM ($P^{(n)}$) for absorbing DTMCs

Recall that we have in Example 3.14 that we have

$$P = \begin{bmatrix} Q & R \\ 0 & A \end{bmatrix}$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.7 & 0.3 \\ 0 & 0.4 & 0.6 \end{bmatrix}$$

So we have

$$P^2 = \begin{bmatrix} Q^2 & QR + RA \\ 0 & A^2 \end{bmatrix} \Rightarrow P^{(n)} = \begin{bmatrix} Q^n & ? \\ 0 & A^n \end{bmatrix}$$

(we solve for the ? quadrant).

Note that A^n is just A but with the lower right 4 elements as a separate matrix to the matrix power of n. Thus if we take the limit as $n \to \infty$ we get our limiting probability matrix (for the lower right 4 elements)

$$\begin{bmatrix} \pi_1 & \pi_2 \\ \pi_3 & \pi_4 \end{bmatrix}$$

22.5 Example 3.11 (continued) solution (mean absorption time w_i)

Recall the modified TPM

$$P = \begin{array}{c c} & \begin{array}{ccc} 0^* & 1^* & 2^* \end{array} \\ \begin{array}{c} 0^* \\ 1^* \\ 2^* \end{array} & \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

What is the mean absoprtion time for this DTMC, given that it begins in state 0^*?

Solution. Rcall that we have (in matrix form)

$$w' = (I - Q)^{-1} e'$$

Thus we have (since we only have 1 transient state)

$$\begin{aligned} w_{0^*} = w' &= (\frac{1}{2})^{-1}(1) \\ &= 2 \end{aligned}$$

Looking at this particular TPM, given that the DTMC initially begins in state 0^*, each transition will return to state 0^* with probability $\frac{1}{2}$ or become absorbed into one of the two absorbing states with probability $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$. Thus

$$T \mid (X_0 = 0^*) \sim GEO(\frac{1}{2}) \Rightarrow E[T \mid X_0 = 0^*] = \frac{1}{\frac{1}{2}} = 2$$

22.6 Example 3.13 (continued) solution (mean absorption time w_i)

Consider the DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

starts off in state 1, how long on average does it take to end up in either state 2 or 3?

Solution. Find w_1 where $w' = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$. Note that

$$w' = (I - Q)^{-1} e' = \begin{bmatrix} \frac{70}{39} & \frac{10}{13} \\ \frac{10}{39} & \frac{20}{13} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow w_1 = \frac{10}{39} + \frac{20}{13} = \frac{70}{39} \approx 1.79$$

22.7 Example 3.13 (continued) solution (average number of visits prior to absorption $S_{i,l}$)

Consider the DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.1 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Given $X_0 = 1$, what is the average number of visits to state 0 prior to absoprtion? Also what is the probability that the DTMC ever makes a visit to state 0?

Solution. Goal: find $S_{1,0}$ where

$$S = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} S_{0,0} & S_{0,1} \\ S_{1,0} & S_{1,1} \end{bmatrix} \end{array}$$

From our earlier calculations we know

$$S = (I - Q)^{-1} = \begin{bmatrix} \frac{70}{39} & \frac{10}{13} \\ \frac{10}{39} & \frac{20}{13} \end{bmatrix} \Rightarrow S_{1,0} = \frac{10}{39} \approx 0.257$$

Lastly, we calculate (as derived)

$$f_{1,0} = \frac{S_{1,0} - \delta_{1,0}}{S_{0,0}} = \frac{\frac{10}{39} - 0}{\frac{70}{39}} = \frac{1}{7}$$

23 Tutorial 8

23.1 Transforming problem into absorption problem

Recall the DTMC $\{X_n, n \in \mathbb{N}\}$ from Tutorial 7 with state space $S = \{0, 1, 2, \ldots\}$. For each $i \in S$ we had

$$P_{i,i+1} = \frac{\alpha}{2+i} \text{ and } P_{i,0} = \frac{2+i-\alpha}{2+i}$$

resulting in the TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & \ldots \end{array} \\ \begin{bmatrix} \frac{2-\alpha}{2} & \frac{\alpha}{2} & 0 & 0 & 0 & \ldots \\ \frac{3-\alpha}{3} & 0 & \frac{\alpha}{3} & 0 & 0 & \ldots \\ \frac{4-\alpha}{4} & 0 & 0 & \frac{\alpha}{4} & 0 & \ldots \\ \frac{5-\alpha}{5} & 0 & 0 & 0 & \frac{\alpha}{5} & \ddots \\ \frac{6-\alpha}{6} & 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \end{array}$$

In part (7) we determined the number of transitions to reach state 2, conditional on $X_0 = 0$ (letting $0 \le \alpha \le 2$). Show that we can solve this expected value using absorbing DTMC results.

Solution. We can transform our DTMC to an equivalent sub-DTMC with absorbing state 2 and TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} \frac{2-\alpha}{2} & \frac{\alpha}{2} & 0 \\ \frac{3-\alpha}{3} & 0 & \frac{\alpha}{3} \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

This sub-DTMC behaves the same up until the first visit to 2 which is all that we are concerned with. Thus the expected number of transitions to 2 starting from state 0 is equivalent to the mean absorption time for state 0 or w_0. That is

$$w_0 = 1 + Q_{0,0}w_0 + Q_{0,1}w_1 \Rightarrow w_0 = 1 + \frac{2-\alpha}{2}w_0 + \frac{\alpha}{2}w_1 \Rightarrow w_0 = \frac{2}{\alpha} + w_1$$
$$w_1 = 1 + Q_{1,0}w_0 = 1 + \frac{3-\alpha}{3}w_0$$

Thus we have

$$w_0 = \frac{2}{\alpha} + 1 + \frac{3-\alpha}{3}w_0 \Rightarrow w_0 = \frac{3}{\alpha}\frac{2+\alpha}{\alpha} = \frac{3(2+\alpha)}{\alpha^2}$$

as desired.

23.2 Mean recurrent time

Consider DTMC with TPM

$$P = \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{3}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} \end{array}$$

If the process begins in state 0, then on average how many transitions are required to return to 0?

Solution. Note that the DTMC is ergodic and irreducible, thus the BLT applies here.

24 March 13, 2018

24.1 Example 4.1 solution ($P(X_j = \min\{X_1, \dots, X_n\})$ for $X_i \sim EXP(\lambda_i)$)

Let $\{X_i\}_{i=1}^n$ be a sequence of independent random variables where $X_i \sim EXP(\lambda_i)$, $i = 1, 2, \dots, n$. For $j \in \{1, 2, \dots, n\}$, calculate $P(X_j = \min\{X_1, \dots, X_n\})$.

Solution. Note that

$$\begin{aligned}
P(X_j = \min\{X_1, \dots, X_n\}) = \dots& \\
&(n-1)\text{-fold intersection of events (no } X_j < X_j \text{ term)} \\
&= P(X_j < X_1, X_j < X_2, \dots, \widehat{X_j < X_j}, \dots, X_j < X_n) \\
&= \int_0^\infty P(X_j < X_1, X_j < X_2, \dots, X_j < X_n \mid X_j = x)\lambda_j e^{-\lambda_j x} dx \\
&\text{since } X_j \text{ and } \{X_i\}_{i=1, i\neq j}^n \text{ are independent, we have} \\
&= \int_0^\infty P(X_1 > x, X_2 > x, \dots, X_n > x)\lambda_j e^{-\lambda_j x} dx \\
&\text{since } \{X_i\}_{i=1, i\neq j}^n \text{ are independent, we have} \\
&= \int_0^\infty P(X_1 > x)P(X_2 > x)\dots P(X_n > x)\lambda_j e^{-\lambda_j x} dx \\
&= \int_0^\infty e^{-\lambda_1 x} \cdot e^{-\lambda_2 x} \cdot \dots \cdot e^{-\lambda_n x}\lambda_j e^{-\lambda_j x} dx \\
&= \frac{\lambda_j}{\sum_{i=1}^n \lambda_i} \int_0^\infty (\sum_{i=1}^n \lambda_i) e^{-\lambda_1 x} e^{-(\sum \lambda_i)x} dx \\
&= \frac{\lambda_j}{\sum_{i=1}^n \lambda_i}
\end{aligned}$$

24.2 Theorem 4.1 (equivalent memoryless property definition)

Theorem 24.1. X is memoryless iff $P(X > y + z) = P(X > y)P(X > z)$, $\forall y, z \geq 0$.

Proof. **Forwards direction:**
Recall a rv is memoryless iff

$$P(X > y + z \mid X > y) = P(X > z) \quad \forall y, z \geq 0$$

Thus we have

$$P(X > y + z \mid X > y) = \frac{P(X > y + z, X > y)}{P(X > y)} = \frac{P(X > y + z)}{P(X > y)}$$

If X is memoryless, then $P(X > y + z \mid X > y) = P(X > z)$.
The result immediately follows.
Backwards direction:
Conversely, if $P(X > y + z) = P(X > y)P(X > z)\ \forall y, z \geq 0$, note that by conditional probability

$$P(X > y + z \mid X > y) = \frac{P(X > y + z)}{P(X > y)} = \frac{P(X > y)P(X > z)}{P(X > y)} = P(X > z)$$

which implies that X is memoryless. $\square$

24.3 Theorem 4.2 (exponentials are memoryless)

Theorem 24.2. An exponential distribution is memoryless.

Proof. Suppose that $X \sim EXP(\lambda)$. For $y, z \geq 0$

$$P(X > y + z) = e^{-\lambda(y+z)} = e^{-\lambda y}e^{-\lambda z} = P(X > y)P(X > z)$$

By Theorem 4.1, X is memoryless. $\square$

25 March 15, 2018

25.1 Example 4.2 solution (non-identical exponentials problem)

Suppose that a computer has 3 switches which govern the transfer of electronic impulses. These switches operate independently of one another and their lifetimes are exponentially distributed with mean lifetimes of 10, 5, and 4 years, respectively.

1. What is the probability that the time until the very first switch breakdown exceeds 6 years?

   Solution. Let X_i represent the lifetime of switch i where $i = 1, 2, 3$.
   We know $X_i \sim EXP(\lambda_i)$ where $\lambda_1 = \frac{1}{10}$, $\lambda_2 = \frac{1}{5}$, and $\lambda_3 = \frac{1}{4}$.
   The time until the 1st breakdown is defined by the rv $Y = \min\{X_1, X_2, X_3\}$.
   Since the lifetimes are independent of each other, $Y \sim Exp(\lambda)$ where $\lambda = \frac{1}{10} + \frac{1}{5} + \frac{1}{5} = \frac{11}{20}$ (from Example 4.1).
   We wish to calculate $P(Y > 6) = e^{-\frac{11}{20}(6)} = e^{-3.3} = 0.0369$.

2. What is the probability that switch 2 outlives switch 1?

   Solution. We simply want to calculate $P(X_1 < X_2)$. We showed in Example 2.11 that this is

   $$P(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{1/10}{1/10 + 1/5} = \frac{1}{3}$$

3. If switch 3 is known to have lasted 2 years, what is the probability it will last at most *3 more* years?

Solution. We wish to compute $P(X_3 \le 5 \mid X_3 > 2)$, or

$$1 - P(X_3 > 5 \mid X_3 > 2) = 1 - P(X_3 > 2 + 3 \mid X_3 > 2)$$

By the memoryless property of EXP, this is equivalent to

$$1 - P(X_3 > 3) = 1 - e^{-\frac{1}{4}(3)} = 1 - e^{-0.75} \approx 0.5276$$

4. Consider only switches 1 and 2, what is the expected amount of time until they have both suffered a breakdown?

Solution. We want to solve for $E[\max\{X_1, X_2\}]$. Consider

$$\begin{aligned}
E[\max\{X_1, X_2\} \mid X_2 > X_1] &= E[X_2 \mid X_2 > X_1] \\
&= E[X_1 + (X_2 - X_1) \mid X_2 > X_1] \\
&= E[X_1 \mid X_2 > X_1] + E[X_2 - X_1 \mid X_2 > X_1] \\
&\text{by Exercise 4.1 we have} \\
&= E[\min\{X_1, X_2\}] + E[X_2 - X_1 \mid X_2 > X_1] \\
&\text{by memoryless property and remark (2) in notes we have} \\
&= E[\min\{X_1, X_2\}] + E[X_2]
\end{aligned}$$

In a completely similar fashion, we have

$$E[\max\{X_1, X_2\} \mid X_1 > X_2] = E[\min\{X_1, X_2\}] + E[X_1]$$

Thus by total probability and total expectations (we don't have to worry about $X_1 = X_2$ since exponential is continuous and $P(X = c) = 0$).

$$\begin{aligned}
E[\max\{X_1, X_2\}] &= P(X_1 > X_2)E[\max\{X_1, X_2\} \mid X_1 > X_2] + P(X_2 > X_1)E[\max\{X_1, X_2\} \mid X_2 > X_1] \\
&= \frac{2}{3}[E[\min\{X_1, X_2\}] + E[X_1]] + \frac{1}{3}[E[\min\{X_1, X_2\}] + E[X_2]] \\
&= E[\min\{X_1, X_2\}] + \frac{2}{3}E[X_1] + \frac{1}{3}E[X_2] \\
&= \frac{1}{\frac{1}{10} + \frac{1}{5}} + \frac{2}{3}(10) + \frac{1}{3}(5) \\
&= \frac{35}{3} \approx 11.67 \text{ years}
\end{aligned}$$

26 March 20, 2018

26.1 Poisson processes are Poisson distributed

Theorem 26.1. If $\{N(t), t \ge 0\}$ is a Poisson process at rate λ, then $N(t) \sim POI(\lambda t)$.

Proof. We will prove this by showing that their MGFs are equivalent (i.e. the MGF of the Poisson process is that of $POI(\lambda t)$).
For $t \ge 0$, let $\phi_t(u) = E[e^{uN(t)}]$ the MGF of $N(t)$ (where the MGF is parameterized by u).

For $h \geq 0$, consider

$$\begin{aligned}
\phi_{t+h}(u) &= E[e^{uN(t+h)}] \\
&= E[e^{u[N(t+h)-N(t)+N(t)]}] \\
&= E[e^{u[N(t+h)-N(t)]}e^{uN(t)}] \\
&= E[e^{u[N(t+h)-N(t)]}]E[e^{uN(t)}] && \text{due to independent increments} \\
&= E[e^{uN(h)}]E[e^{uN(t)}] && \text{due to stationary increments} \\
&= \phi_t(u)\phi_h(u)
\end{aligned}$$

Next note that for $j \geq 2$ we have

$$\begin{aligned}
&0 \leq P(N(h)=j) \leq P(N(h) \geq 2) \\
\Rightarrow &0 \leq \frac{P(N(h)=j)}{h} \leq \frac{P(N(h) \geq 2)}{h} \\
\Rightarrow &0 \leq \lim_{h \to 0} \frac{P(N(h)=j)}{h} \leq \lim_{h \to 0} \frac{P(N(h) \geq 2)}{h} && \text{Let } h \to 0 \\
\Rightarrow &0 \leq \lim_{h \to 0} \frac{P(N(h)=j)}{h} \leq 0 && P(N(h)=j) \in o(h)
\end{aligned}$$

Thus $\lim_{h \to 0} \frac{P(N(h)=j)}{h} = 0$ by Squeeze theorem. Therefore $P(N(h)=j) \in o(h)$ i.e. of order h. Using this result, we get

$$\begin{aligned}
\phi_h(u) &= E[e^{uN(h)}] \\
&= \sum_{j=0}^{\infty} e^{uj} P(N(h)=j) \\
&= P(N(h)=0) + e^u P(N(h)=1) + \sum_{j=2}^{\infty} e^{uj} P(N(h)=j) \\
\sum_{j=2}^{\infty} e^{uj} P(N(h)=j) &= \sum j = 2 c_j o(h) \text{ linear combination of } o(h) \text{ so we have} \\
&= 1 - \lambda h + o(h) + e^u(\lambda h + o(h)) + o(h) \\
&= 1 + \lambda h(e^u - 1) + o(h)
\end{aligned}$$

Returning to $\phi_{t+h}(u)$

$$\begin{aligned}
\phi_{t+h}(u) &= \phi_t(u)\phi_h(u) \\
&= \phi_t(u)[1 + \lambda h(e^u - 1) + o(h)] \\
&= \phi_t(u) + \lambda h(e^u - 1)\phi_t(u) + o(h) \\
\phi_{t+h}(u) - \phi_t(u) &= \lambda h(e^u - 1)\phi_t(u) + o(h) \\
\frac{\phi_{t+h}(u) - \phi_t(u)}{h} &= \lambda(e^u - 1)\phi_t(u) + \frac{o(h)}{h} && \text{divide by } h \\
\frac{d}{dt}(\phi_t(u)) &= \lambda(e^u - 1)\phi_t(u) && \text{take } h \to 0
\end{aligned}$$

This is a differential equation for $\phi_t(u)$, note that (replacing t with s)

$$\begin{aligned}
\frac{\frac{d}{ds}(\phi_s(u))}{\phi_s(u)} &= \lambda(e^u - 1) \\
\frac{d}{ds}(\ln \phi_s(u)) &= \lambda(e^u - 1) \\
d(\ln \phi_s(u)) &= \lambda(e^u - 1)ds \\
\int_0^t d(\ln \phi_s(u)) &= \int_0^t \lambda(e^u - 1)ds \\
\ln \phi_s(u)\Big|_{s=0}^{s=t} &= \lambda(e^u - 1)t \\
\ln \phi_t(u) - \ln \phi_0(u) &= (\lambda t)(e^u - 1) \\
\ln \phi_t(u) &= (\lambda t)(e^u - 1) \qquad \phi_0(u) = E[e^{uN(0)}] = 1 \text{ since } N(0) = 0 \\
\phi_t(u) &= e^{(\lambda t)(e^u - 1)}
\end{aligned}$$

We recognize this as the mgf of a $POI(\lambda t)$ r.v. By the uniqueness property of MGFs, $N(t) \sim POI(\lambda t)$. □

26.2 Interarrival times T_i between Poisson events are Exponential distributed

Theorem 26.2. If $\{N(t), t \geq 0\}$ is Poisson process at rate $\lambda > 0$, then $\{T_i\}_{i=1}^{\infty}$ is a sequence of iid $EXP(\lambda)$ random variables.

Proof. We begin by considering T_1 (first arrival time) for $t \geq 0$. Note that the tail probability of T_1 is

$$\begin{aligned}
P(T_1 > t) &= P(\text{no events before time } t) \\
&= P(N(t) = 0) \\
&= \frac{e^{-\lambda t}(\lambda t)^0}{0!} \\
&= e^{-\lambda t}
\end{aligned}$$

We recognize this as the tail probability of an $EXP(\lambda)$ r.v. thus $T_1 \sim EXP(\lambda)$.
Next for $s > 0$ and $t \geq 0$ consider

$$\begin{aligned}
&P(T_2 > t \mid T_1 = s) \\
=&P(T_2 > t \mid N(w) = 0 \quad \forall w \in [0, s) \text{ and } N(s) = 1) \\
=&P(\text{no events occur in } (s, s+t] \mid N(w) = 0 \quad \forall w \in [0, s) \text{ and } N(s) = 1) \\
=&P(N(s+t) - N(s) = 0 \mid N(w) = 0 \quad \forall w \in [0, s) \text{ and } N(s) = 1) \\
&\text{since conditional interval } [0, s] \text{ does not overlap } (s, s+t] \text{ by independent increments} \\
=&P(N(s+t) - N(s) = 0) \\
=&P(N(t) = 0) \qquad \text{by stationary increments} \\
=&e^{-\lambda t}
\end{aligned}$$

Note that $e^{-\lambda t}$ is independent of s, thus T_1 and T_2 are independent r.v's so $P(T_2 > t \mid T_1 = s) = P(T_2 > t) = e^{-\lambda t}$ thus $T_2 \sim EXP(\lambda)$. Carrying this out inductively leads to the desired result. □

27 March 22, 2018

27.1 Example 4.3 solution

At a local insurance company, suppose that fire damage claims come into the company according to a Poisson process at rate 3.8 expected claims per year.

1. What is the probability that exactly 5 claims occur in the time interval $(3.2, 5]$ (measured in years)?

   Solution. Let $N(t)$ be the number of claims arriving to the company in the interval $[0, t]$.

   Since $\{N(t), t \geq 0\}$ is a Poisson process with $\lambda = 3.8$, we want to find

   $$P(N(5) - N(3.2) = 5) = \frac{e^{-3.8(1.8)}(3.8 \cdot 1.8)^5}{5!} \approx 0.1335$$

2. What is the probability that the time between the 2nd and 4th claims is between 2 and 5 months?

   Solution. Let T be the time between the 2nd and 4th claims, so $T = T_3 + T_4$ where T_3, T_4 are independent $EXP(3.8)$ random variables representing the arrival times of the 3rd and 4th claims.

   We know $T \sim Erlang(2, 3.8)$ which means

   $$P(T > t) = e^{-3.8t} \sum_{j=0}^{2-1} \frac{(3.8t)^j}{j!} = e^{-3.8t}(1 + 3.8t)$$

   for $t \geq 0$.

   We wish to calculate (converted from months to years)

   $$P(\frac{1}{6} < T\frac{5}{12}) = P(T > \frac{1}{6}) - P(T > \frac{5}{12}) \approx 0.3367$$

3. If exactly 12 claims have occurred within the first 5 years, how many claims on average occurred within the first 3.5 years?

   Solution. We want to calculate $E[N(3.5) \mid N(5) = 12]$. Recall for $s < t$ we have $N(s) \mid (N(t) = n) \sim Bin(n, \frac{s}{t})$, thus $N(3.5) \mid (N(5) = 12) \sim BIN(12, 3.5/5)$ so $E[N(3.5) \mid N(5) = 12] = 12(3.5/5) = 42/5 = 8.4$.

   Remark 27.1. Without conditioning, $N(3.5) \sim POI(3.8 \cdot 3.5)$ so $E[N(3.5)] = (3.8)(3.5) = 13.3 \neq 8.4$.

   So conditioning on $N(5)$ **does indeed affect the mean of** $N(3.5)$.

4. At another competing insurance company, suppose fire damage claims arrive according to a Poisson process at rate 2.9 claims per year. What is the probability that 3 claims arrive to this company before 2 claims arrive to the first company? Assume insurance comapnies operate independently of each other.

   Solution. Let $N_1(t)$ denote the number of claims arriving to the 1st company by time t and $N_2(t)$ denote the number of claims arriving at the 2nd company.

   We are assuming $\{N_1(t), t \geq 0\}$ (i.e. Poisson process at rate $\lambda_1 = 3.8$) and $\{N_2(t), t \geq 0\}$ (i.e. Poisson process at rate $\lambda_2 = 2.9$) are independent processes.

   In general

   $$P(S_n^{(1)} < S_m^{(2)}) = \sum_{j=0}^{m-1} \binom{n+j-1}{n-1} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^n \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^j$$

We want the probability of 3 claims to company 2 before 2 claims to company 1 so we have

$$\begin{aligned}
P(S_3^{(2)} < S_2^{(1)} &= 1 - P(S_2^{(1)} < S_3^{(2)}) \\
&= 1 - \sum_{j=0}^{3-1} \binom{2+j-1}{2-1} \left(\frac{3.8}{3.8+2.9}\right)^n \left(\frac{2.9}{2.9+3.8}\right)^j \\
&\approx 0.2191
\end{aligned}$$

(we don't need to take the complement, but for clarity we did).

28 March 27, 2018

28.1 Example 4.3 (continued) solution (Poisson process with classifications)

Suppose that fire damage claims can be categorized as being either commercial, business or residential.
At the first insurance company, past history suggests that 15% are commercial, 25% are business, and the remaining 60% are residential. Over the course of the next 4 years, what is the probability that the company experiences fewer than 5 claims in each of the 3 categories?

Solution. Let $N_c(t)$ bbe the number of commercial claims by time t. Likewise $N_b(t)$ and $N_r(t)$ denote the number of business and residential claims by time t, respectively.
It follows that (recall $\lambda = 3.8$ for the whole process)

$$\begin{aligned}
N_c(4) &\sim POI(3.8 \cdot 0.15 \cdot 4 = 2.28) \\
N_b(4) &\sim POI(3.8 \cdot 0.25 \cdot 4 = 3.8) \\
N_r(4) &\sim POI(3.8 \cdot 0.6 \cdot 4 = 9.12)
\end{aligned}$$

We wish to calculate the joint probability

$$\begin{aligned}
P(N_c(4) < 5, N_b(4) < 5, N_r(4) < 5) &= P(N_c(4) < 5) \cdot P(N_b(4) < 5) \cdot P(N_r(4) < 5) && \text{independence} \\
&= \left(\sum_{i=0}^{4} \frac{e^{-2.28}(2.28)^i}{i!}\right) \left(\sum_{i=0}^{4} \frac{e^{-3.8}(3.8)^i}{i!}\right) \left(\sum_{i=0}^{4} \frac{e^{-9.12}(9.12)^i}{i!}\right) \\
&= (0.91857)(0.66784)(0.05105) \\
&\approx 0.0313
\end{aligned}$$

28.2 Theorem 4.5 (conditional distribution of first arrival time given $N(t) = 1$ is uniform)

Theorem 28.1. Suppose that $\{N(t), t \geq 0\}$ is a Poisson process at rate λ. Given $N(t) = 1$, the conditional distribution of the first arrival time is uniformly distributed on $[0, t]$. That is $S_1 \mid (N(t) = 1) \sim U(0, t)$.

Proof. In order to identify the conditional distribution $S_1 \mid (N(t) = 1)$, we consider the cdf of $S_1 \mid (N(t) = 1)$ which we will denote

$$
\begin{aligned}
G(s) &= P(S_1 \le s \mid N(t) = 1) \qquad 0 \le s \le t \\
&= \frac{P(S_1 \le s, N(t) = 1)}{P(N(t) = 1)} \\
&= \frac{P(1 \text{ event in } [0, s] \text{ and } 0 \text{ events in } (s, t])}{P(N(t) = 1)} \\
&= \frac{P(N(s) = 1, N(t) - N(s) = 0)}{P(N(t) = 1)} \\
&= \frac{P(N(s) = 1)P(N(t) - N(s) = 0)}{P(N(t) = 1)} && \text{independent increments} \\
&= \frac{e^{-\lambda s}\lambda^1}{1!} \frac{e^{-\lambda(t-s)}\lambda^0}{0!} \frac{1!}{e^{-\lambda t}\lambda^1} \\
&= \frac{e^{-\lambda s}(\lambda s)^1}{1!} \frac{e^{-\lambda(t-s)}(\lambda(t-s))^0}{0!} \frac{1!}{e^{-\lambda t}(\lambda t)^1} \\
&= \frac{s}{t}
\end{aligned}
$$

which is exactly the cdf of a $U(0, t)$ random variable so $S_1 \mid (N(t) = 1) \sim U(0, t)$. □

28.3 Theorem 4.6 (conditional joint distribution of n arrival times is the n order statistics with $U(0, t)$)

Theorem 28.2. Let $\{N(t), t \ge 0\}$ be a Poisson process at rate λ. Given $N(t) = n$, the conditional joint distribution of the n arrival times is identical to the joint distribution of the n order statistics from the $U(0, t)$ distribution. That is

$$(S_1, S_2, \ldots, S_n) \mid (N(t) = n) \sim (Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)})$$

where $\{Y_i\}_{i=1}^n$ is an iid sequence of $U(0, t)$ random variables.

Proof. For $0 < s_1 < s_2 < \ldots < s_n < t$, consider

$$P(S_1 \le s_1, s_1 < S_2 \le s_2, \ldots, s_{n-1} < S_n \le s_n \mid N(t) = n) = \frac{P(S_1 \le s_1, s_1 < S_2 \le s_2, \ldots, s_{n-1} < S_n \le s_n, N(t) = n)}{P(N(t) = n)}$$

Denote the above as Q. Thus we have

$$
\begin{aligned}
Q &= \frac{P(N(s_1) = 1, N(s_2) - N(s_1) = 1, \ldots, N(s_n) - N(s_{n-1}) = 1, N(t) - N(s_n) = 0)}{P(N(t) = n)} \\
&= \frac{P(N(s_1) = 1)P(N(s_2) - N(s_1) = 1) \ldots P(N(s_n) - N(s_{n-1}) = 1)P(N(t) - N(s_n) = 0)}{P(N(t) = n)} && \text{independent increments} \\
&= \frac{e^{-\lambda s_1}(\lambda s_1)^1}{1!} \cdot \ldots \cdot \frac{e^{-\lambda(s_n - s_{n-1})}(\lambda(s_n - s_{n-1}))^1}{1!} \cdot \frac{e^{-\lambda(t-s_n)}(\lambda(t - s_n))^0}{0!} \cdot \frac{n!}{e^{-\lambda t}(\lambda t)^n} \\
&= \frac{n!}{t^n} s_1(s_2 - s_1) \ldots (s_n - s_{n-1}) && 0 < s_1 < s_2 < \ldots < s_n < t
\end{aligned}
$$

Differentiating this we get our density function

$$\frac{\partial^n Q}{\partial s_1 \partial s_2 \ldots \partial s_n} = \frac{n!}{t^n}$$

which is the density function for our n order statistics from the $U(0,t)$ distribution. □

29 March 29, 2018

29.1 Example 4.4 solution (average waiting time)

Suppose customers arrive to a subway station in accordance with a Poisson process at rate λ. If the subway train departs promptly at time t (for some time $t > 0$) what is the expected total *waiting* time of all customers arriving during the interval $[0,t]$?

Solution. Let $N(t)$ count the number of customers arriving to the station by time t where $N(t) \sim POI(\lambda t)$. Letting S_i denote the arrival time of the ith customer to the station we have that $W_i = t - S_i$ represent the waiting time of the ith customer.
The total waiting time of all customers is

$$W = \sum_{i=1}^{N}(t) = W_i = t - S_i$$

Goal: calculate $E[W]$.

Remark 29.1. We cannot use the result of the random sum of random variables where $E[W] = E[N(t)]E[W_i]$ since two assumptions are violated here (1) $N(t)$ and W_i are not independent and W_i's are not iid.

Note that

$$\begin{aligned}
E[W] &= E\left[\sum_{i=1}^{N(t)}(t - S_i)\right] \\
&= \sum_{n=0}^{\infty} E\left[\sum_{i=1}^{N(t)}(t - S_i) \mid N(t) = n\right] P(N(t) = n) \\
&= \sum_{n=1}^{\infty} E\left[\sum_{i=1}^{n}(t - S_i) \mid N(t) = n\right] P(N(t) = n) && W = 0 \text{ if } N(t) = 0 \\
&= \sum_{n=1}^{\infty} E\left[nt - \sum_{i=1}^{n} S_i \mid N(t) = n\right] P(N(t) = n) \\
&= \sum_{n=1}^{\infty} \left\{nt - E\left[\sum_{i=1}^{n} S_i \mid N(t) = n\right]\right\} P(N(t) = n)
\end{aligned}$$

By theorem 4.6, note that $(S_1, \ldots, S_n) \mid N(t) = n \sim (Y_{(1)}, \ldots, Y_{(n)})$ where $(Y_{(1)}, \ldots, Y_{(n)})$ are the n order statistics from $U(0,t)$ distribution.

Thus

$$\begin{aligned}
E[W] &= \sum_{n=1}^{\infty} \left\{ nt - E\left[\sum_{i=1}^{n} Y_{(i)} \right] \right\} P(N(t) = n) \\
&= \sum_{n=1}^{\infty} \left\{ nt - E\left[\sum_{i=1}^{n} Y_{i} \right] \right\} P(N(t) = n) && \text{summing over all } Y_{(i)} \text{ anyways} \\
&= \sum_{n=1}^{\infty} \left\{ nt - \sum_{i=1}^{n} E[Y_i] \right\} P(N(t) = n) \\
&= \sum_{n=1}^{\infty} \left\{ nt - \sum_{i=1}^{n} \frac{t}{2} \right\} P(N(t) = n) \\
&= \sum_{n=1}^{\infty} \left\{ nt - \frac{nt}{2} \right\} P(N(t) = n) \\
&= \frac{t}{2} \sum_{n=0}^{\infty} nP(N(t) = n) \\
&= \frac{t}{2} E[N(t)] \\
&= \frac{\lambda t^2}{2}
\end{aligned}$$

Note that $E[Y_{(i)}] \neq \frac{t}{2}$ since it's not just any Y_i but the ith order statistics which depends on i, so we must do the conversion in line 2.

29.2 Example 4.5 solution (pdf and joint probability of arrival times)

Satellites are launched at times according to a Poisson process at rate 3 per year.
During the past year, it was observed only two satellites were launched. What is the joint probability that the first of these two satellites was launched in the first 5 months and the second satellite was launched prior to the last 2 months of the year?

Solution. Let $\{N(t), t \geq 0\}$ be the Poisson process at rate $\lambda = 3$ governing satellite launches. We are interested in calculating

$$P(S_1 \leq \frac{5}{12}, S_2 \leq \frac{5}{6} \mid N(1) = 2)$$

Given $N(1) = 2$, we use Theorem 4.6 to obtain the conditional joint pdf of $(S_1, S_2) \mid (N(1) = 2) \sim (Y_{(1)}, Y_{(2)})$ where $\{Y_i\}_{i=1}^{2}$ are iid $U(0, 1)$.
Recall we have for the 2 order statistics pdf at time 1

$$g(s_1, s_2) = \frac{2!}{1^2} = 2$$

for $0 < s_1 < s_2 < 1$. We use this conditional joint pdf and integrate it over the follow region

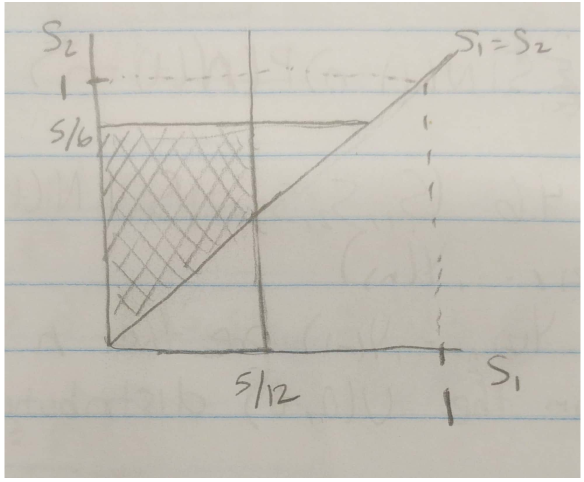


Figure 29.1: The support of our pdf is the top left triangle in the diagram. We want to integrate over the shaded region.

So

$$\begin{aligned}
P(S_1 \le \frac{5}{12}, S_2 \le \frac{5}{6} \mid N(t) = 2) &= \int_0^{5/12} \int_{s_1}^{5/6} 2 \quad ds_2 \quad ds_1 \\
&= 2 \int_0^{5/12} (s_2) \Big|_{s_1}^{5/6} \quad ds_1 \\
&= 2 \int_0^{5/12} (\frac{5}{6} - s_1) \quad ds_1 \\
&= 2 \left(\frac{5s_1}{6} - \frac{s_1^2}{2} \right) \Big|_0^{5/12} \\
&= \frac{25}{48} \\
&\approx 0.5208
\end{aligned}$$

29.3 Example 5.1 solution

A Poisson process $\{X(t), t \ge 0\}$ at rate λ is an example of a CTMC. Determine the values of $v_i, i \in N$ and construct the TPM of its embedded Markov chain.

Solution. Clearly $X(t) \in \mathbb{N}$ for all $t \ge 0$. Moreover, by the way a Poisson process is constructed, we have $T_i \sim EXP(v_i = \lambda)$, $i \in \mathbb{N}$.
The corresponding TPM of the embedded Markov chain looks like

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & \dots \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \end{array}$$