1 Probability

Theorem 1.1. *Suppose all n men at a party throw their hats in the center of the room. Each man then randomly selects a hat. Show that the probability that none of the n men selects this own hat is*

$$\sum_{k=2}^{n} \frac{(-1)^k}{k!}.$$

Example 1.2. The probability of winning on a single toss of the dice is p. A starts, and if he fails, he passes the dice to B, who then attempts to win on her toss. They continue tossing the dice back and forth until one of them wins. What are their respective probabilities of winning?

$$P(A \text{ wins}) = \sum_{n=0}^{\infty} P(A \text{ wins on} (2n+1)^{\text{th}} \text{ toss}) = \sum_{n=0}^{\infty} p(1-p)^{2n} = \frac{1}{2-p}.$$

Example 1.3. A and B play until one has 2 more points than the other. Assuming that each point is independently won by A with probability p, what is the probability they will play a total of $2n$ points? What is the probability that A will win? (In each trial, one point either is won by A or by B.)

$$P(2n \text{ points are needed}) = (2p(1-p))^{n-1} \left(p^2 + (1-p)^2\right), \ n \geq 1.$$

$$P(A \text{ wins}) = p^2 \sum_{n=1}^{\infty} (2p(1-p))^{n-1} = \frac{p^2}{1-2p(1-p)}.$$

Example 1.4. A deck of 52 playing cards, containing all 4 aces, is randomly divided into 4 piles of 13 cards each.

$$P(\text{each pile includes one ace}) = 1 \cdot \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49}.$$

(We know 51 cards have not been put, and the second one can just be put in 39 positions.)

Example 1.5. A fair coin is continually flippes. What is the probability that the pattern T, H, H, H occurs before the pattern H, H, H, H? $\frac{15}{16}$, since the only way in which the pattern H, H, H, H can appear before patten T, H, H, H is if the first four flips all land heads.

Example 1.6. Suppose that each coupon obtained is, independent of what has been previously obtained, equally likely to be any of n different types. Find the expected number of coupons one need to obtain in order to have at least one of each type.

Let X_i denote the number of additional coupons collected until the collector has $i+1$ types. Then $X_i \sim \text{Geo}(1 - i/n)$. Then $E\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} E[X_i] = \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j}$.

$$P(X_{(i)} \leqslant x) = \sum_{k=i}^{n} \binom{n}{k} F^k(x) \left(1 - F(x)\right)^{n-k}, f_{X_{(i)}} = \frac{n!}{(n-i)!(i-1)!} f(x) F^{i-1}(x) \left(1 - F(x)\right)^{n-i}.$$

$$f_{X(1),\ldots,X_{(n)}}(x_1, \ldots, x_n) = \begin{cases} n! \prod_{k=1}^{n} f(x_k), & x_1 < \cdots < x_n \\ 0, & \text{otherwise} \end{cases}.$$

Example 1.7. Let $X \sim \text{Binom}(n,p)$, then for $1 \leqslant m \leqslant n$, $E[X(X-1)\cdots(X-m+1)] = \frac{n!}{(n-m)!}p^m$.

Example 1.8. Consider n independent flips of a coin having probability p of landing heads. Say a changeover occurs whenever an outcome dif and only ifers from the cone preceding it. For instance, if the results of the flips are $H\ H\ T\ H\ T\ H\ H\ T$, then there are a total of five changeovers. If $p = \frac{1}{2}$, what is the probability there are k changeovers?

Each flip after the first will, independently, result in a changeover with probability $\frac{1}{2}$. Therefore,

$$P(k \text{ changeovers}) = \binom{n-1}{k}\left(\frac{1}{2}\right)^{n-1}.$$

$$\text{number of changeovers} = \sum_{i=2}^{n} \mathbb{1}_{\{\text{if a change over results from the } i^{\text{th}} \text{ flip}\}} = \sum_{i=1}^{n} X_i, \quad E[X_i] = 2p(1-p).$$

Example 1.9. An urn contains $n+m$ balls, of which n are red and m are black. They are withdrawn from the urn, one at a time and without replacement. Let X be the number of red balls removed before the first black ball is chosen. Let

$$X_i = \mathbb{1}_{\{\text{if red ball } i \text{ is taked before any black ball is chosen}\}}.$$

Then

$$X = \sum_{i=1}^{n} X_i, \quad E[X_i] = P(X_i = 1) = \frac{1}{n+1}$$

since each of these $n+1$ balls is equally likely to be the one chosen earliest.

Example 1.10. A total of r keys are to be put, one at a time, in k boxes, with each key independently being put in box i with probability p_i, $\sum_{i=1}^{k} p_i = 1$. Each time a key is put in a nonempty box, we say that a colllision occurs. Find the expected number of collisions.
Let N_i denote the number of keys in box $i, \forall i \in [k]$. Then, with X equal to the number of collisions

$$X = \sum_{i=1}^{k}(N_i - 1)^+ = \sum_{i=1}^{k}(N_i - 1 + \mathbb{1}_{\{N_i=0\}}).$$

Example 1.11. Let $a_1 < a_2 < \cdots < a_n$ denote a set of n numbers, and consider any permutation of these numbers. We say there is an inversion of a_i and a_j in the permutation if $i < j$ and a_j precedes a_i. Consider now a random permutation of $a_1, \ldots, a_n$−in the sense that each of the $n!$ permutations is equally likely to be chosen−and let N denote the number of inversions in this permutation. Also, let $N_i = \{\text{number of } k :\ k < i, a_i \text{ precedes } a_k \text{ in the permutation}\}, \forall i \in [n]$. Then $N = \sum_{i=1}^{n} N_i$. Show $N_1, \ldots, N_n$ are independent random variables.
Knowing the values of $N_1, \ldots, N_j$ is equivalent to knowing the relative ordering of the elements $a_1, \ldots, a_j$. The independence result follows for clearly the number of $a_1, \ldots, a_j$ that follow a_{j+1} does not probabilistically depend on the relative ordering of $a_1, \ldots, a_j$.
Also,

$$P(N_i = k) = \frac{1}{i}, \forall k = 0, \ldots, i-1,$$

which follows since of the elements $a_1, \ldots, a_{j+1}$, the elements a_{j+1} is equally likely to be the first or second or $\cdots$ or $(i+1)^{\text{th}}$.

2 Conditional probability

Definition 2.1. The conditional variance of X given that $Y = y$ is defined by

$$\mathrm{Var}(X|Y=y) = E\left[(X - E[X|Y=y])^2 \middle| Y = y\right], \mathrm{Var}(X|Y) = E\left[(X - E[X|Y])^2 \middle| Y\right],$$

$$\mathrm{Var}(X|Y=y) = E[X^2|Y=y] - (E[X|Y=y])^2, \mathrm{Var}(X|Y) = E[X^2|Y] - (E[X|Y])^2.$$

Example 2.2. At a party n men take off their hats. The hats are then mixed up and each man randomly selects one. We say that a match occurs if a man selects his own hat. What is the probability of no mathes? What is the probability of exactly k matches?

$$P_n = P(E) = P(E|M)P(M) + P(E|M^c)P(M^c) = P(E|M^c)\frac{n-1}{n}.$$

$$P(E|M^c) = P_{n-1} + \frac{1}{n-1}P_{n-2}, \quad P_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^n}{n!}.$$

We consider any fixed group of k men. The probability that they, and only they, select their own hat is $\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-(k-1))}P_{n-k} = \frac{(n-k)!}{n!}P_{n-k}$. There are $\binom{n}{k}$ choices of a set of k men, the desired probability of exactly k matches is $\frac{P_{n-k}}{k!} = \frac{\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{n-k}}{(n-k)!}}{k!}$, which for n large, $\approx \frac{e^{-1}}{k!}$.

Example 2.3 (The Ballot Problem)**.** In an election, candidate A receives n votes, and candidate B receives m votes where $n > m$. Assuming that all orderings are equally likely, show that the probability that A is always ahead in the count of votes is $\frac{n-m}{n+m}$.
Let $P_{n,m}$ denote the desired probability. Then

$$\begin{aligned} P_{n,m} &= P(A \text{ always ahead}|A \text{ receives last vote})\frac{n}{n+m} + P(A \text{ always ahead}|B \text{ receives last vote})\frac{m}{n+m} \\ &= \frac{n}{n+m}P_{n-1,m} + \frac{m}{n+m}P_{n,m-1}, \end{aligned}$$

Consider successive flips of a coin that always land on "heads" with probability p,

$$\begin{aligned} P(\text{first time equal } 2n) &= P(\text{first time equal } 2n|n \text{ heads in first } 2n)\binom{2n}{n}p^n(1-p)^n. \\ &= \left(\frac{n}{n-1+n} + \frac{n-1}{n-1+n}\right)P_{n,n-1}\binom{2n}{n}p^n(1-p)^n = \frac{\binom{2n}{n}p^n(1-p)^n}{2n-1}. \end{aligned}$$

Example 2.4. If X is a discrete r.v. and Y is a continuous r.v., we have

$$f_{Y|X=i}(y) = \frac{P(X=i|Y=y)f_Y(y)}{P(X=i)}.$$

Example 2.5. An individual whose level of exposure to a certain pathogen is x will contract the disease caused by this pathogen with robability $P(x)$. If the exposure level of a randomly chosen member of the population has pdf f, determine the conditional probability density of the exposure level of that member given that he or she (1) has the disease; (2) does not have the disease.

$$f(x|\text{disease}) = \frac{P(\text{disease}|x)f(x)}{\int P(\text{disease}|x)f(x)dx} = \frac{P(x)f(x)}{\int P(x)f(x)dx}, f(x|\text{no disease}) = \frac{(1-P(x))\,f(x)}{\int (1-P(x))\,f(x)dx}.$$

Example 2.6. A coin having probability p of coming up heads is successively flipped until two of the the most recent three flips are heads. Let N denote the number of flips. (Note that if the first two flips are heads, then $N = 2$.) Find $E[N]$.
Let X denote the first time a head appears. Then

$$\begin{aligned} E[N|X] &= E[N|X,h,h]p^2 + E[N|X,h,t]pq + E[N|X,t,h]pq + E[N|X,t,t]q^2 \\ &= (X+1)p^2 + (X+1)pq + (X+2)pq + ((X+2) + E[N])\, q^2. \end{aligned}$$

Example 2.7. You have two opponents with whom you alternate play. Whenever you play A you win with probability p_A; whenever you play B, you win with probability p_B, where $p_B > p_A$. If your objective is to minimize the expected number of games you need to play to win two in a row (two consective success), should you start with A or B?
Let N_A and N_B denote the number of games needed given that you start with A and B.

$$E[N_A] = E[N_A|w]p_A + E[N_A|l](1-p_A),$$

$$E[N_A|w] = E[N_A|ww]p_B + E[N_A|wl](1-p_B) = 2 + (1-p_B)E[N_A], \quad E[N_A|l] = 1 + E[N_B].$$

Example 2.8. Each element in a sequence of binary data is either 1 with probability p or 0 with probability $1-p$. A maximal subsequence of consecutive values having identical outcomes is called a run. For instance, if the outcome sequence is $1,1,0,1,1,1,0$, the first run is of length 2, the second is of length 1, and the third is of length 3.

$$E[L_1] = E[L_1|1]p + E[L_1|0](1-p) = \frac{1}{1-p}p + \frac{1}{p}(1-p), E[L_2] = p\frac{1}{p} + (1-p)\frac{1}{1-p} = 2.$$

Example 2.9. Let $X_1, X_2, \cdots$ be independent continuous random variables with a common distribution function F and density $f = F'$, and for $k \geq 1$ let

$$N_k = \min\{n \geq k : X_n = k^{\text{th}} \text{ largest of } X_1, \ldots, X_n\}.$$

(a) Show that $P(N_k = n) = \frac{k-1}{n(n-1)}, n \geq k$.
Let A_i denote the event that X_i is the k^{th} largest of $X_1, \ldots, X_i, \forall k \leqslant i \leqslant n$. Then $A_k, \ldots, A_n$ are independent events with $P(A_i) = \frac{1}{i}, \forall k \leqslant i \leqslant n$.

$$P(N_k = n) = P(A_k^c A_{k+1}^c \cdots A_{n-1}^c A_n) = (1 - P(A_k)) \cdots P(A_n) = \frac{k-1}{k}\frac{k}{k+1} \cdots \frac{n-2}{n-1}\frac{1}{n} = \frac{k-1}{n(n-1)}.$$

(b) Argue that

$$f_{X_{N_k}}(x) = f(x)\left(\overline{F}(x)\right)^{k-1} \sum_{i=0}^{\infty} \binom{i+k-2}{i} (F(x))^i.$$

Proof. Since knowledge of the set of values $\{X_1, \ldots, X_n\}$ gives us no information about the order of these random variables it follows that given $N_k = n$, the conditional distribution of X_{N_k} is the same as the distribution of k^{th} **largest** of n random variables having distribution F (!!!).

$$f_{X_{N_k}}(x) = \sum_{n=k}^{\infty} \frac{k-1}{n(n-1)} \frac{n!}{(n-k)!(k-1)!} (F(x))^{n-k} \left(\overline{F}(x)\right)^{k-1} f(x).$$

Then make the change of variable $i = n - k$.

3 Discrete time markov chains

Definition 3.1. State j is called recurrent if $f_i = P_i(\eta_i < \infty) = 1$. Positive if $E[\eta_j|X_0 = j] < \infty$.

Corollary 3.2. Let $N_i = \sum_{n=0}^{\infty} \mathbb{1}_{\{X_n=i\}}$. The state i is recurrent if

$$\sum_{n=0}^{\infty} P_{ii}^n = E[N_i|X_0 = i] = \infty.$$

Proof. $P(N_i = 1|X_0 = i) = P(\eta_i = \infty|X_0 = i) = 1 - f_i 0$. by the strong Markov property,

$$P(N_i = 2|X_0 = i) = P(\eta_i < \infty|X_0 = i)P(\eta_i = \infty|X_0 = i) = f_i(1 - f_i).$$

$$P(N_i = k|X_0 = i) = f_i^{k-1}(1 - f_i), \forall k \in \mathbb{Z}^+, \quad P(N_i \geq k|X_0 = i) = f_i^{k-1}, \forall k \in \mathbb{Z}^+.$$

$$E[N_i|X_0 = i] = \sum_{k=1}^{\infty} P(N_i \geq k|X_0 = i) = \sum_{k=1}^{\infty} f_i^{k-1} = \sum_{k=0}^{\infty} f_i^k.$$ □

Definition 3.3.

$$f_k(i, j) = P(\eta_j = k|X_0 = i).$$

$$f(i, j) = P(X_n = j \text{ for some } n > 0|X_0 = i) = P(\eta_1 < \infty|X_0 = i) = \sum_{k=1}^{\infty} f_k(i, j).$$

$$m_j = E[\eta_j|X_0 = j] = \sum_{k=1}^{\infty} k f_k(j, j),$$

Theorem 3.4. *If the* MC *is irreducible and* ***recurrent****, then for any initial state* $\pi_j = \frac{1}{m_j}$.

Proposition 3.5. Suppose $f : E \to [0, \infty)$ is a nonnegative function. Moreover, suppose our MC is irreducible and positive recurrent, then for any initial state

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} f(X_n) = \sum_{j\in E} f(j)\pi_j, \text{ w.p.1..}$$

Proposition 3.6. Suppose $\{X_n\}$ is irreducible and recurrent. Then for each state $i \in E$, **all** invariant measures of $\{X_n\}$ are scalar multilples of $\nu^{(i)}$. Moreover, if $\sum_{j\in E} \nu_j^{(i)} = E_i[\eta_i] < \infty$, i.e, i is positive recurrent, a **unique stationary** distribution π exists, and

$$\pi_j = \frac{\nu_j^{(i)}}{\sum_{j\in E} \nu_j^{(i)}} = \frac{E_i\left[\sum_{n=0}^{\eta_i-1} \mathbb{1}_{\{X_n=j\}}\right]}{E_i[\eta_i]}.$$

Theorem 3.7. *(a) If j is transient or recurrent null, then for any $i \in E$,* $\lim_{n\to\infty} P_{ij}^n = 0$.

(b) If j is positive recurrent and aperiodic, then $\pi_j = \lim_{n\to\infty} P_{jj}^n > 0$, $\lim_{n\to\infty} P_{ij}^n = f_{ij}\pi_j$.

$$E = \{0, \dots, m, m+1, \dots, M\},$$

where the states $T = \{0, \dots, m\}$ are all transient states, and the states $R = \{m+1, \dots, M\}$ are all recurrent states. The transition matrix P for this DTMC is then of the form

$$P = (Q, U; 0, S)\,.$$

$$E_i[N_j] = \sum_{n=0}^{\infty} P_i(X_n = j) = \sum_{n=0}^{\infty} P_{i,j}^n = \sum_{n=0}^{\infty} Q_{i,j}^n = \left[(I-Q)^{-1}\right]_{i,j}\,, \quad \eta = \min\{n \geq 1 : X_n \in R\}.$$

$$P_i(X_\eta = j) = \sum_{k \in T} P_i(X_\eta = j | X_1 = k) P_{ik} + \sum_{k \in R} P_i(X_\eta = j | X_1 = k) P_{ik} = \sum_{k \in T} Q_{ik} P_k(X_\eta = j) + u_{ij},$$

or equivalently, if we let $H = [P_k(X_\eta = j)]_{k \in T, j \in R}$ be a matrix containing all such probabilities, we see instead that

$$H = QH + U, \quad H = (I-Q)^{-1}U.$$

$$\alpha_i = P(X_0 = i), \forall i \in T, \quad \sum_{i \in T} \alpha_i = 1.$$

$$P(\eta > n) = \sum_{j \in T} P(X_n = j) = \sum_{j \in T} \sum_{i \in T} P(X_n = j | X_0 = i)\alpha_i = \sum_{i \in T} \alpha_i \sum_{j \in T} P_{ij}^n = \alpha Q^n \mathbb{1}_{m+1}, \forall n \in \mathbb{Z}^+.$$

Remark.

$$W_n = \begin{cases} X_n, & n < N \\ A, & n \geq N \end{cases}.$$

$$Q_{i,j} = P_{i,j},\ i \notin \mathcal{A}, j \notin \mathcal{A}, \quad Q_{i,A} = \sum_{j \in \mathcal{A}} P_{i,j},\ i \notin \mathcal{A}, \quad Q_{A,A} = 1.$$

Because the original MC will have entered a state in $\mathcal{A}$ by time m if and only if the state at time m of the new MC is A, we see that

$$P(X_k \in \mathcal{A} \text{ for some } k = 1, \dots, m | X_0 = i) = P(W_m = A | X_0 = i) = P(W_m = A | W_0 = i) = Q_{i,A}^m.$$

In a sequence of independent flips of a fair coin, let N denote the number of flips until there is a run of three consecutive heads. Find

(a) $P(N \leqslant 8)$.
Define a MC with states $0, 1, 2, 3$ where for i ($0 \leqslant i < 3$) means that we currently are on a run of i consecutive heads, and where state 3 means that a run of three consecutive heads has already occurred. (!!! 3 is an **absorbing state**). Because there would be a run of three consecutive heads within the first eight flips if and only if $X_8 = 3$, the desired probability is $Q_{0,3}^8$.

Remark.

$$T = \min\{n \geq 0 : X_n \in \{0, N\}\}, \quad P_i := P\left(X_T = N | X_0 = i\right).$$

$$P_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)^N}, & p \neq 1/2 \\ i/N, & p = 1/2 \end{cases}.$$

4 Poisson processes

$$P\left(X_2 > t | X_1 = s\right) = P\left(N(s, s+t] = 0 | \ X_1 = s\right) = P\left(N(s, s+t] = 0\right) = e^{-\lambda t}., \forall s > 0.$$

$$T_{N(t)} = \sup\{T_n : T_n \leqslant t\}, \quad T_{N(t)+1} = \inf\{T_n : T_n > t\}.$$

Example 4.1 (The coupon collecting problem)**.** There are m different types of coupons. Each time a person collects a coupon it is, independently of ones previously obtained, a type j coupon with probability p_j, $\sum_{j=1}^{m} p_j = 1$. Let N denote the number of coupons one needs to collect in order to have a complete collection of at least one of each type. Find $E[N]$.
Thinking of each L_n as the type of coupon collected during the n^{th} trival,$\forall n \in \mathbb{Z}^+$.

$$N_j = \min\{n \geq 1 : L_n = j\}, \forall j \in [m], \quad N = \max_{1 \leqslant j \leqslant m} N_j.$$

Introduce a Poisson process $\{N(t); t \geq 0\}$ with rate $\lambda = 1$, and assign to each point T_n of $\{N(t); t \geq 0\}$ the label L_n, the n^{th} coupon collected variable from our iid sequence of interest.

$$N_k(t) = \sum_{n=1}^{\infty} \mathbb{1}_{\{T_n \leqslant t, L_n = k\}}, \forall t > 0, \quad T = \max_{1 \leqslant j \leqslant m} T_1^j.$$

Then $\{N_1(t); t \geq 0\}, \ldots, \{N_m(t); t \geq 0\}$ are indep PP. $\{T_n^1\}_{n \in \mathbb{Z}^+}, \ldots, \{T_n^m\}_{n \in \mathbb{Z}^+}$ are indep.

$$T_1^j \sim \exp(1/p_j) \Rightarrow P(T < t) = P\left(\max_{1 \leqslant j \leqslant m} T_1^j < t\right) = P(T_1^j < t, \forall j \in [m]) = \prod_{j=1}^{m} \left(1 - e^{-p_j t}\right).$$

$$E[T] = \int_0^{\infty} P(T > t) dt = \int_0^{\infty} \left[1 - \prod_{j=1}^{m} \left(1 - e^{-p_j t}\right)\right] dt.$$

$$T = \sum_{i=1}^{N} T_i \text{ and } T_i \perp\!\!\!\perp N \Rightarrow E[T] = E[N]E[T_1] = E[N].$$

Example 4.2.

$$P\left(T_n^1 < T_m^2\right) = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n+m-1-k}.$$

Theorem 4.3. *Let $A_1, \ldots, A_n$ be disjoint intervals with union B, let $a_1, \ldots, a_n$ be their respective lengths, and set $b = a_1 + \cdots + a_n$. Then for $k_1 + \cdots + k_n = k$, $k_1, \ldots, k_n \in \mathbb{N}$,*

$$P\left(N_{A_1} = k_1, \ldots, N_{A_n} = k_n \mid N_B = k\right) = \frac{k!}{k_1! \cdots k_n!} \left(\frac{a_1}{b}\right)^{k_1} \cdots \left(\frac{a_n}{b}\right)^{k_n}.$$

Theorem 4.4. *Given that $N(t) = n$, the n arrival times $T_1, \ldots, T_n$ have the same distribution as the order statistic corresponding to n independent random variables uniformly distributed on the interval $(0, t)$, i.e.,*

$$f_{T_1, \ldots, T_n | N(t) = n}(t_1, \ldots, t_n) = \begin{cases} \frac{n!}{t^n}, & 0 < t_1 < \cdots < t_n; \\ 0, & \textit{otherwise} \end{cases}.$$

Example 4.5. Let $X, Y_1, \ldots, Y_n$ be independent exponential random variables; X having rate λ, and Y_i having rate μ. Let A_j be the event that the j^{th} smallest of these $n+1$ random variables is one of the Y_i. Find $p = P(X > \max_i Y_i)$ by using the identity

$$p = P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cdots A_{n-1}).$$

$$P(A_1) = \frac{n\mu}{\lambda + n\mu}, \quad P(A_j|A_1 \cdots A_{j-1}) = \frac{(n-j+1)\mu}{\lambda + (n-j+1)\mu}, \forall 1 < j \leqslant n.$$

Theorem 4.6. *If $X_1, \ldots, X_n$ are independent exponential,, then $\min_i X_i$ indep the order of X_i's.*

$$E\left[X_{(2)}\right] = E\left[X_{(1)}\right] + E\left[X_{(2)} - X_{(1)}\right] = \frac{1}{\mu_1 + \mu_2} + \frac{1}{\mu_2}\frac{\mu_1}{\mu_1 + \mu_2} + \frac{1}{\mu_1}\frac{\mu_2}{\mu_1 + \mu_2}.$$

Example 4.7. Customers arrive at a two-server service station according to a Poisson process with rate λ. Whenever a new customer arrives, any customer that is in the system immediately departs. A new arrival enters service first with server 1 and then with server 2. If the service times at the servers are independent exponentials with respective rates μ_1 and μ_2, what proportion of entering customers completes their service with server 2?
Let S_i denote the service time at server $i, i = 1, 2$ and let X denote the time until the next arrival. Then, with p denoting the proportion of customers that are served by both servers, we have

$$p = P(X > S_1 + S_2) = P(X > S_1)P(X > S_1 + S_2|X > S_1) = \frac{\mu_1}{\mu_1 + \lambda}\frac{\mu_2}{\mu_2 + \lambda}.$$

Example 4.8. Consider an n-server parallel queuing system where customers arrive according to a Poisson process with rate λ, where the service times are exponential random variables with rate μ, and where any arrival finding all servers busy immediately departs without receiving any service. If an arrival finds all servers busy, find

(a) the expected number of busy servers found by the next arrival, $N|T = t \sim \text{Binom}\left(n, e^{-\mu T}\right)$.

$$E[N] = \int_0^\infty E[N|T = t]\lambda e^{-\lambda t} dt = \int_0^\infty n e^{-\mu t}\lambda e^{-\lambda t} dt = \frac{n\lambda}{\lambda + \mu}.$$

(b) the probability that the next arrival finds all servers free, $P(N = 0) = \prod_{j=1}^n \frac{(n-j+1)\mu}{\lambda + (n-j+1)\mu}$.

(c) the probability that the next arrival finds exactly i of the servers free. Conditioning on T,

$$P(N = n - i) = \frac{\lambda}{\lambda + (n-i)\mu}\prod_{j=1}^{i}\frac{(n-j+1)\mu}{\lambda + (n-j+1)\mu}.$$

Example 4.9. Consider a single server queuing system where customers arrive according to a Poisson process with rate λ, service times are exponential with rate μ, and customers are served in the order of their arrival. Suppose that a customer arrives and finds $n-1$ others in the system. Let X denote the number in the system at the moment that customer departs. Find the probability mass function of X. $P(X = m) = \binom{n+m-1}{n-1} p^n (1-p)^m$. with $p = \frac{\mu}{\lambda + \mu}$.

Example 4.10. A cable car starts off with n riders. The times between successive stops of the car are independent exponential random variables with rate λ. At each stop one rider gets off. This takes no time, and no additional riders get on. After a rider gets off the car, he or she walks home. Independently of all else, the walk takes an exponential time with rate μ. Then distribution of the time S_n at which the last rider departs the car is Gamma(n, λ).
Suppose the last rider departs the car at time t. What is the probability that all the other riders are home at that time?
Use the result that given $S_n = t$, the set of times at which the first $n-1$ riders departed are independent uniform$(0,t)$ random variables. Therefore, each of these riders will not be walking at time t with probability

$$p = \int_0^t \left(1 - e^{-\mu(t-s)}\right) ds/t = \frac{1-e^{-\mu t}}{\mu t}, \quad \text{so } p^n.$$

5 Renewal processes

$$T_n(\omega) = \inf\{t \geq 0 : N(t,\omega) \geq n\}, \quad N(t,\omega) = \sup\{n \geq 1 : T_n(\omega) \leqslant t\}.$$
$$T_{N(t)} = \sup\{T_n : T_n \leqslant t\}, \quad T_{N(t)+1} = \inf\{T_n : T_n > t\}, \quad T_{N(t)} \leqslant t < T_{N(t)+1}.$$

(a) $h(t), t \geq 0$ is bounded on fintie intervals and F is a cdf satisfying $F(0) = 1$ and $\lim_{t\to\infty} F(t) = 1$.

(b) See if F is nonarithmetic and h is dRI on $[0,\infty)$. Then $\lim_{t\to\infty} H(t) = \frac{\int_0^\infty h(s)ds}{\int_0^\infty x dF(x)}$.

Proposition 5.1. A necessary condition for $h(t), t \geq 0$ being directly Riemann integrable is h is bounded and continuous a.e. w.r..t lebesgue measure. Some sufficient conditions are

(a) h is nonnegative and nonincreasing and Riemann integrable on $[0,\infty)$.

(b) h is continuous a.e. on $[0,\infty)$, and there exists a positive function $b(t), t \geq 0$ such that $|h(t)| \leqslant b(t), \forall t \geq 0$ with b being directly Riemann integrable.

Definition 5.2. A stochastic process $\{X(t), t \geq 0\}$ is a regenerative process w.r.t. the sequence of random times $\{\tau_n\}_{n\in\mathbb{Z}^+}$ if the random blocks $\{(\tau_{n+1} - \tau_n, \{X(t); \tau_n \leqslant t < \tau_{n+1}\})\}_{n\in\mathbb{N}}$ are iid.

Example 5.3. Suppose $\{N(t)\}$ is a renewal process with increments $\{X_n\}$ that have cdf F satisfying $F(0) = 0$ and $\lim_{t\to\infty} F(t) = 1$. Show that its renewal function u satisfies $u(t) = E[N(t)] + 1$.

$$\begin{aligned} u(t+a) - u(t) = E[N(t+a) - N(t)] &= \int_0^\infty E\left[N(t+a) - N(t) | Y(t) = y\right] dG_t(y) \\ &= \int_0^a E\left[N(t+a) - N(t) | Y(t) = y\right] dG_t(y) = \int_0^a E\left[1 + N(a-y)\right] dG_t(y) \\ &= \int_0^a u(a-y) dG_t(y) \leqslant u(a) \int_0^a G_t(y) \leqslant u(a). \end{aligned}$$

Example 5.4. If $A(t)$ and $B(t)$ are, respectively, the age and the excess at time t of a renewal process having an interarrival distribution F, calculate

$$\begin{aligned} P\left(B(t) > x | A(t) = s\right) &= P\left(0 \text{ renewals in } (t, t+x] | A(t) = s\right) = P\left(\text{interarrival} > x + s | A(t) = s\right) \\ &= P\left(\text{interarrival} > x + s | \text{interarrival} > s\right) = \frac{1 - F(x+s)}{1 - F(s)}. \end{aligned}$$

6 Continuous time Markov chains

Theorem 6.1. *A jump process $\{X(t); t \geq 0\}$ on E with embedded process $\{X_n\}_{n\geq 0}$ and sojourn times (holding time) $\{T_n\}_{n\geq 1}$ is a CTMC if and only if the following are satisfied:*

(a) $\{X_n\}_{n\geq 0}$ is a DTMC on E with transition matrix $P = (P_{ij})_{i,j\in E}$, where $P_{ii} = 0, \forall i \in E$.

(b) For $m \in \mathbb{Z}^+$ and $t_1, \ldots, t_m \geq 0$, we have

$$P(T_1 \leqslant t_1, \ldots, T_m \leqslant t_m | X_n, n \in \mathbb{N}) = \prod_{n=1}^{m} P(T_n \leqslant t_n | X_{n-1}) = \prod_{n=1}^{m} \left(1 - e^{-\nu_{X_{n-1}} t}\right),$$

which says that the sojourn times $\{T_n\}_{n\in\mathbb{Z}^+}$ are conditionally independent, given the entire $\{X_n\}_{n\in\mathbb{N}}$ and such a sojourn time in state i is exponentilly distributed with rate ν_i.

Example 6.2 ($M/M/s$)**.**

$$P(X_{n+1} = i+1, T_{n+1} > t | X_n = i, (X_j, T_j), j \leqslant n) = e^{-[\min(i,s)\mu+\lambda]t} \frac{\lambda}{\lambda + \min(i,s)\mu}, \forall i \in \mathbb{Z}^+,$$

$$P(X_{n+1} = i-1, T_{n+1} > t | X_n = i, (X_j, T_j), j \leqslant n) = e^{-[\min(i,s)\mu+\lambda]t} \frac{\min(i,s)\mu}{\lambda + \min(i,s)\mu}, \forall i \in \mathbb{Z}^+.$$

Lemma 6.3.

$$\lim_{t\to 0} \frac{1 - P_{ii}(t)}{t} = \nu_i, \quad \lim_{t\to 0} \frac{P_{ij}(t)}{t} = q_{ij}, \forall i \neq j.$$

$$P_{ii}(t) = P(X(t) = i | X(0) = i), \quad 1 - P_{ii}(t) = P(X(t) \neq i | X(0) = i).$$

Let $\{N_i(t), t \geq 0\}$ be a Poisson process with rate ν_i. Then

$$P(X(t) \neq i | X(0) = i) = P(N_i(t) \geq 1) = \nu_i t + o(t), \text{ as } t \to 0.$$

$$\lim_{t\to 0} \frac{1 - P_{ii}(t)}{t} = \nu_i.$$

$$\begin{aligned}
P_{ij}(t) = P(X(t) = j | X(0) = i) &= P(X(t) = j, T_1 \leqslant t | X(0) = i) + P(X(t) = j, T_1 > t | X(0) = i) \\
&= \sum_{k\neq i} P(X(t) = j, T_1 \leqslant t, X(T_1) = k | X(0) = i) + \mathbb{1}_{\{i=j\}} e^{-\nu_i t} \\
&= \sum_{k\neq i} P(X(t) = j, T_1 \leqslant t | X(T_1) = k, X(0) = i) P_{ik} + \mathbb{1}_{\{i=j\}} e^{-\nu_i t} \\
&= \sum_{k\neq i} \int_0^t P(X(t) = j | T_1 = s, X(T_1) = k, X(0) = i) \nu_i e^{-\nu_i s} ds P_{ik} + \mathbb{1}_{\{i=j\}} e^{-\nu_i t} \\
&= \sum_{k\neq i} \int_0^t P(X(t) = j | X(s) = k) \nu_i e^{-\nu_i s} ds P_{ik} + \mathbb{1}_{\{i=j\}} e^{-\nu_i t} \\
&= \sum_{k\neq i} \int_0^t P_{kj}(t-s) \nu_i e^{-\nu_i s} ds P_{ik} + \mathbb{1}_{\{i=j\}} e^{-\nu_i t} = \int_0^t \sum_{k\neq i} P_{kj}(t-s) \nu_i e^{-\nu_i s} ds P_{ik} + \mathbb{1}_{\{i=j\}} e^{-\nu_i t}.
\end{aligned}$$

$$\lim_{t\to 0}\frac{P_{ij}(t)}{t}=\lim_{t\to 0}\frac{1}{t}\int_0^t\sum_{k\neq i}P_{kj}(t-s)\nu_i e^{-\nu_i s}P_{ik}ds=\lim_{t\to 0}\frac{1}{t}\int_0^t\sum_{k\neq i}P_{kj}(s)\nu_i e^{-\nu_i(t-s)}P_{ik}ds=\sum_{k\neq i}P_{kj}(0)\nu_i P_{ik}=\nu_i P_{ij},\forall i\neq$$

Corollary 6.4.

$$P'_{ij}(0)=\lim_{t\downarrow 0}\frac{P_{ij}(t)-P_{ij}(0)}{t}=\nu_i P_{ij},\forall i\neq j,\ P'_{ii}(0)=\lim_{t\downarrow 0}\frac{P_{ii}(t)-P_{ii}(0)}{t}=-\nu_i,\forall i\neq j,$$

$$\nu_i=\sum_{j\neq i}\nu_i P_{ij}=\sum_{j\neq i}q_{ij},\ P_{ij}=\frac{q_{ij}}{\nu_i}=\frac{q_{ij}}{\sum_{j\neq i}q_{ij}},$$

Example 6.5. The backward equations for the birth and death process become

$$P'_{0j}(t)=\lambda_0 P_{1,j}(t)-\lambda_i P_{0j}(t),\quad P'_{ij}(t)=\lambda_i P_{i+1,j}(t)+\mu_i P_{i-1,j}(t)-(\lambda_i+\mu_i)P_{ij}(t),\forall i>0.$$

$$\begin{aligned}P'_{ij}(t)&=\lim_{s\downarrow 0}\frac{P_{ij}(t+s)-P_{ij}(t)}{s}=\lim_{s\downarrow 0}\frac{\sum_{k\in E}P_{ik}(t)P_{kj}(s)-\sum_{k\in E}P_{ik}(t)P_{kj}(0)}{s}\\&=\sum_{k\in E}P_{ik}(t)\lim_{s\downarrow 0}\frac{P_{kj}(s)-P_{kj}(0)}{s}=\sum_{k\in E}P_{ik}(t)P'_{kj}(0)=\sum_{k\in E}P_{ik}(t)q_{kj},\forall\text{finite } E.\end{aligned}$$

Example 6.6. The forward equations for the birth and death process become

$$P'_{i0}(t)=\mu_1 P_{i1}(t)-\lambda_0 P_{i0}(t),\quad P'_{ij}(t)=\lambda_{j-1}P_{i,j-1}(t)+\mu_{j+1}P_{i,j+1}(t)-(\lambda_j+\mu_j)P_{ij}(t).$$

$$Q^n=S\Lambda^n S^{-1},\quad P(t)=\sum_{k=0}^{\infty}Q^k\frac{t^k}{k!}=\cdots,\ t\geq 0,\quad \xi_i=\inf\{t>0: X(t^-)\neq X(t)=i\},\forall i\in E.$$

$$E_i\left[\int_0^{\xi_i}\mathbb{1}_{\{X(t)=j\}}\right]=\frac{1}{\nu_i}E_i\left[\sum_{n=0}^{\eta_i-1}\mathbb{1}_{\{X_n=j\}}\right],\forall i,j\in E.$$

Theorem 6.7. *Suppose X is irreducible, and let π be a positive measure on E. Then the following statements are equivalent.*

(a) $\{X(t);t\geq 0\}$ is ergodic, i.e., irreducible, and positive recurrent, with stationary distribution π.

(b) π satisfies the balance equation $\pi Q=0$, $\sum_{i\in E}\pi_i=1$.

(c) For a fixed $i\in E$, $E_i[\xi_i]<\infty$ and $\pi_j=\frac{E_i\left[\int_0^{\xi_i}\mathbb{1}_{\{X(t)=j\}}dt\right]}{E_i[\xi_i]},\forall j\in E$.

Theorem 6.8. *Suppose X is ergodic, with stationary distribution π. Then for each $i,j\in E$,*

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t P_{ij}(s)ds=\pi_j,\quad \lim_{t\to\infty}P_{ij}(t)=\pi_j.$$

Example 6.9. The birth and death process with parameters $\lambda_n=0$ and $\mu_n=\mu, n\geq 0$ is called a pure death process. Find $P_{ij}(t)$.
Since the death rate is constant, it follows that as long as the system is nonempty, the number of deaths in any interval of length t will be a Poisson random variable with mean μt. Hence,

$$P_{ij}(t)=e^{-\mu t}(\mu t)^{i-j}/(i-j)!,\ 0<j\leqslant i,\quad P_{i0}(t)=\sum_{k=i}^{\infty}e^{-\mu t}(\mu t)^k/k?$$

7 Brownian Motion

Theorem 7.1. *More generally, for a sequence* $0 \leqslant t_1 < \cdots < t_n$,

$$f_{B(t_1),\ldots,B(t_n)}(z_1,\ldots,z_n) = f_{B(t_1)}(z_1) f_{B(t_2-t_1)}(z_2-z_1)\cdots f_{B(t_n-t_{n-1})}(z_n-z_{n-1}),\ z_1,\ldots,z_n \in \mathbb{R}.$$

$$B(t)|B(1)=0 \sim N(0,t-1), \forall t>1.$$

$$f_{B(t)|B(1)=0}(z) = \frac{f_{B(t),B(1)}(z,0)}{f_{B(1)}(0)} = \frac{f_{B(t)}(z) f_{B(1-t)}(0-z)}{f_{B(1)}(0)} = \frac{\frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} \frac{1}{\sqrt{2\pi(1-t)}} e^{-\frac{z^2}{2(1-t)}}}{\frac{1}{\sqrt{2\pi}}} = \frac{1}{\sqrt{2\pi t(1-t)}} e^{-\frac{z^2}{2t(1-t)}}.$$

$$B(t)|B(1) \sim N\left(0, t(1-t)\right), \forall 0<t<1 \Rightarrow P\left(B(t)>a|\tau_a \leqslant t\right) = 1/2.$$

$$\begin{aligned} P\left(B(t)>a|\tau_a \leqslant t\right) &= P\left(B(t)-B(\tau_a)+B(\tau_a)>a|\tau_a \leqslant t\right) = P\left(B(t)-B(\tau_a)>0|\tau_a \leqslant t\right) \\ &= P\left(B(t-\tau_a)>0|\tau_a \leqslant t\right) = P\left(B(t-\tau_a)>0\right) = 1/2 \\ P\left(B(t)>a\right) &= P\left(B(t)>a, \tau_a \leqslant t\right) + P\left(B(t)>a, \tau_a>t\right) = P\left(B(t)>a, \tau_a \leqslant t\right) \\ &= P\left(B(t)>a|\tau_a \leqslant t\right) P\left(\tau_a \leqslant t\right) = 1/2 P\left(\tau_a \leqslant t\right), \forall a>0. \\ P(\tau_a \leqslant t) &= 2P\left(B(t) \geq a\right) = 2\int_a^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx = 2\int_{\frac{a}{\sqrt{t}}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \end{aligned}$$

$$P\left(M(t)<x\right) = P(\tau_x>t).$$

Let A denote the event that the standard BM $B(t)$ has at least one zero in the time interval (t_0, t_1).

$$\begin{aligned} P\left(A|B(t_0)=x\right) &= P\left(\min_{t_0<s<t_1} B(s) \leqslant 0|B(t_0)=x\right) = P\left(\max_{t_0<s<t_1} B(s) \geq 0|B(t_0)=-x\right) \\ &= P\left(\max_{t_0<s<t_1} B(s) \geq x|B(t_0)=0\right) = P\left(\max_{0<s<t_1-t_0} B(s) \geq x|B(0)=0\right) \\ &= P\left(M(t_1-t_0) \geq x\right) = P(\tau_x \leqslant t_1-t_0), \forall x>0. \end{aligned}$$

$$\begin{aligned} P(A) &= \int_{-\infty}^\infty P\left(A|B(t_0)=x\right) f_{B(t_0)}(x) = 2\int_0^\infty P(\tau_x \leqslant t_1-t_0) \frac{1}{\sqrt{2\pi t_0}} e^{-\frac{x^2}{2t_0}} dx \\ &= 2\int_0^\infty \int_0^{t_1-t_0} \frac{x}{\sqrt{2\pi}} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} du \frac{1}{\sqrt{2\pi t_0}} e^{-\frac{x^2}{2t_0}} dx = \frac{1}{\pi\sqrt{t_0}} \int_0^{t_1-t_0} \int_0^\infty u^{-\frac{3}{2}} x e^{-\left(\frac{1}{2u}+\frac{1}{2t_0}\right)x^2} dx du \\ &= \frac{1}{\pi\sqrt{t_0}} \int_0^{t_1-t_0} u^{-\frac{3}{2}} \frac{1}{2} \frac{1}{\frac{1}{2u}+\frac{1}{2t_0}} du = \frac{2}{\pi} \arctan\left(\sqrt{\frac{t_1-t_0}{t_0}}\right). \end{aligned}$$

Theorem 7.2 (Optimal sampling theorem)**.** *Let* $\{X(t); t \geq 0\}$ *be a martingale w.r.t.* $\{Y(t); t \geq 0\}$ *and let* τ *be a stopping time w.r.t.* $\{Y(t); t \geq 0\}$.

$$P(\tau<\infty)=1,\ \lim_{a\to\infty} \sup_{t\geq 0} E\left[|X(t\wedge\tau)| \mathbb{1}_{\{|X(t\wedge\tau)|\geq a\}}\right] = 0, \Rightarrow E\left[X(\tau)\right] = E\left[X(0)\right].$$

$$P\left(B(\tau_{a,b})=a\right) = \frac{b}{b-a}, \quad E[\tau_{a,b}] = |a|b$$