Compilation and Program Analysis

By Francescomaria Faticanti (and Théophile Dubuc)

Contents

Part I
Introduction

1 Probability

Def 1 : Event

Let Ω be a simple space. An **event** is a set $E \subseteq \Omega$.
Let E, $F \subseteq \Omega$, then there are 3 possible operations: $E \cup F$, $E \cap F$, $\overline{E}$.
Two events E and F are **exclusive** if $E \cap F = \varnothing$.
$E_1, ..., E_n$ **partition** F if they are mutually exclusive (i.e. $\forall i, j$, $E_i \cap E_j = \varnothing$) and $\bigcup\limits_{i=1}^{n} E_i = F$.

Def 2 : Union of probability

$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$

Property 1 :

$\mathbb{P}(E \cap F) \leqslant \mathbb{P}(E) + \mathbb{P}(F)$.
Moreover, if $E \cap F = \varnothing$ then $\mathbb{P}(E \cap F) = \mathbb{P}(E) + \mathbb{P}(F)$.

Def 3 : Conditional probability

The **conditional probability** of an event E given F is: $\mathbb{P}(E|F) = \mathbb{P}_F(E) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$

Def 4 : Independancy

Two event E and F are **independant** if $\mathbb{P}(E \cap F) = \mathbb{P}(E).\mathbb{P}(F)$
Therefore, $\mathbb{P}(E|F) = \mathbb{E}$

Def 5 : Conditionnaly independancy

E and F are **conditionnaly independant given** G if $\mathbb{P}(E \cap F|G) = \mathbb{P}(E|G).P(F|G)$

Theorem 1 : Low of Probability

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E \cap F) + \mathbb{P}(E \cap \overline{F}) \\ &= \mathbb{P}(E|F).\mathbb{P}(F) + \mathbb{P}(E|\overline{F}).\mathbb{P}(\overline{F})\end{aligned}$$

Property 2 :

Let $F_1, ..., F_n$ ba a partition of Ω then $\mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E \cap F_i) = \sum_{i=1}^{n} \mathbb{P}(E|F_i).\mathbb{P}(F_i)$

Theorem 2 : Extended Boyes Law

Let $F_1, ..., F_n$ be a partition of Ω then $\mathbb{P}(F|E) = \frac{\mathbb{P}(E|F).\mathbb{P}(F)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|F).\mathbb{P}(F)}{\sum_{i=1}^{n} \mathbb{P}(E|F_i).\mathbb{P}(F_i)}$

2 Random Variable

2.1 Definition

Def 6 : Discrete Random Variable

Let X be a **discrete random variable**. The **probability mass function** is $P_X()$ is defined by $P_X(e) = \mathbb{P}(X = e)$ where $\sum_{x \in X(\Omega)} P_X(x) = 1$.
The **cumulative distrbution function** of X is $F_X(e) = \mathbb{P}(X \leqslant e)$.

List of common discrete random variables:
- $X \sim \text{Bernoulli}(p)$ if $\mathbb{P}(X = 1) = p$
- $X \sim \text{Bernoulli}(n, p)$ if $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
- $X \sim \text{Geometric}(p)$ if $\mathbb{P}(X = k) = (1-p)^{k-1} p$
- $X \sim \text{Poisson}(\lambda)$ if $\mathbb{P}(X = k) = \frac{e^{\lambda} \lambda^k}{k!}$

Def 7 : p.d.f. for Continous Random Variable

The **probability density function** of a continuous R.V. is a non negative function f_X where $\mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f_X(x)\,dx$ where $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

We suppose that we have a generator of uniform $\mathcal{U}(0, 1)$ random variables.

2.2 Inverse-Transform Method

Assumptions:
- (i) We know the c.d.f. $F_X(x) = \mathbb{P}(X \leqslant x)$ of random variable X that we want to simulate
- (ii) This distribution is easily invertible, i.e. we can get x from $F_X(x)$

Continuous case:

Idea: map $u \in \mathcal{U}(0,1) \to x$ instantce if the RV X whose cdef is $F_X(x)$

Assumption: $X \in [0, +\infty[$

$(0, \mathcal{U}) \mapsto (0, x)$ mapping function g^{-1}, g is increasing

A: a value $[0, x]$ should be output with probability $F_X(x)$.

Q: probability that g^{-1} outputs a value in $(0, x)$. Then the probability of generating $(0, u)$ is u which is equal to $\mathbb{P}(0 < \mathcal{U} < u) = \mathcal{P}(0 < X < x) = F_X(x)$. Then $x = F_X^{-1}(u)$.

Inverse-Transform Method to generate RV X :

- Generate $u \in \mathcal{U}(0,1)$
- Return $X = F_X^{-1}(u)$

Example: Generate $X \equiv Exp(\lambda)$

Remember, $F_X(x) = 1 - e^{-\lambda x}$. Then if $u = F_X(x)$ it means that $x = -\frac{1}{\lambda} ln(1-u)$ given a $u \in \mathcal{U}(0,1)$.

Set $x = -\frac{1}{\lambda} ln(1-u)$ which is an instance of $X \equiv Exp(\lambda)$.

Discrete case:

$$\text{Generate a discrete R.V. ssuch that } X = \begin{cases} x_0 \text{ with prob } p_0 \\ \vdots \\ x_k \text{ with prob } p_k \end{cases}$$

First, arrange $x_0, ..., x_k$ such that $x_0 < ... < x_k$

Then, generate $u \in \mathcal{U}(0,1)$

- If $0 < u < p_0$ then output x_0
- If $p_0 < u \leqslant p_1$ then output x_1
- If $p_0 + p_1 < u \leqslant p_0 + p_1 + p_2$ then output x_2
- General case: If $\sum_{i=0}^{l-1} p_i < u < \sum_{i=0}^{l} p_i$ then output x_l

2.3 Accept-reject Method

Discrete case:

Given: efficient method for generating RV Q with p.m.f. $\{q_J,\ J \text{ discrete}\}$ where $q_J = \mathbb{P}(Q = J)$

Requirement: $\forall J$, we must have $q_J > 0 \Leftrightarrow p_J > 0$

Output: RV P with p.m.f. $\{p_J,\ J \text{ discrete}\}$ where $p_J = \mathbb{P}(P = J)$

Suppose that we want to generate

$$P = \begin{cases} 1 \text{ with prob } p_1 = 0.36 \\ 2 \text{ with prob } p_2 = 0.24 \\ 3 \text{ with prob } p_3 = 0.40 \end{cases} \text{ and } Q = \begin{cases} 1 \text{ with prob } q_1 = 0.333... \\ 2 \text{ with prob } q_2 = 0.333... \\ 3 \text{ with prob } q_3 = 0.333... \end{cases}$$

Accept-reject Algo to generate discrete RV P:

Let c be a normalized cinstant such that $\forall J$ such that $p_J > 0$, $\frac{p_J}{q_J} \leqslant c$. Notice that $c > 1$. In fact, $c = max\{\frac{p_J}{q_J}, \forall J\}$.

Here is the sequence of operation for the algo:

- (0) Find RV Q such that $q_j > 0 \Leftrightarrow p_j > 0$
- (1) Generate an instance of Q, call it J
- (2) Generate RV $\mathcal{U} \in [0,1]$

- (3) If $\mathcal{U} < \frac{p_J}{c.q_J}$ then return $P = J$ and stop, otherwise return to step (1)

Then $\mathbb{P}(P$ ends up being set to $J) = p_J$

$$\mathbb{P}(P \text{ is set to } P_J) = \frac{\textit{fraction of times J is generated and J is accepted}}{\textit{fraction of times any thing is accpeted}}$$
$$= \frac{\mathbb{P}(\textit{J is generated})\mathbb{P}(\textit{J is accepted} \mid \textit{J is generated})}{\sum_J \textit{fraction of times J is generated and adopted}}$$
$$= \frac{q_J \cdot \frac{p_J}{c.q_J}}{\sum_J \frac{p_J}{c}} = \frac{\frac{p_J}{c}}{\frac{1}{c}} = p_J$$

Continuous case:

- Find continuous RV Y such that $f_Y(t) > 0 \Leftrightarrow f_X(t) > 0$.
  Let c be a constant such that $\forall t$, such that $f_X(t) > 0$ $\frac{f_X(t)}{f_Y(t)}$, we take the smallest c possible.
- Generate an instance t of Y.
- With prob $\frac{f_X(t)}{c.f_Y(t)}$, return $X = t$ else reject and return to previous step.

3 Stochastic Process

Def 8 : Stochastic Processes

A **stochastic process** is an orderred collection of RVs $\{X_t,\ t \in T\}$. Values assured by $X_t \to$ state space of process (S). t is index of the process.

$\forall i \in \{1, 2\}$, $t = t_i$, SP is a standard RV X_{t_i} of p.m.f. $f_X(t_i)(x)$.
$f_X(t_1)(x) \neq f_X(t_2)(x)$.
We denote by $X_{s1}(t)$ the (deterministique) function of time.
There is 4 cases:

- CS-CT: $s \in \mathbb{R}$, $t \in \mathbb{R}$
- CS-DT: $s \in \mathbb{R}$, $t \in \mathbb{Z}$
- DS-CT: $s \in \mathbb{Z}$, $t \in \mathbb{R}$
- DS-DT: $s \in \mathbb{Z}$, $t \in \mathbb{Z}$

When there is 'DS', we call the stochastic process **chains**.

Property 3 :

Let $(X_t)_t$ be a *Poisson* process of intensity λ, i.e. stochastic process made of Random variables such that $X(t + t') - X(t) \equiv Poisson(\lambda.t')$. $(X_t)_t$ is a Poisson process iff for any $\varepsilon \leftarrow 0$,

$$\mathbb{P}(X(t+\varepsilon) - X(t) = J) = \begin{cases} 1 - \lambda.\varepsilon + o(\varepsilon) & \text{if } J = 0 \\ \lambda.\varepsilon + o(\varepsilon) & \text{if } J = 1 \\ o(\varepsilon) & \text{if } J \geqslant 2 \end{cases}$$

Property 4 :

Let $(N_t^X)_t)$ and $(N_t^Y)_t$ be two independant *Poisson* processes of intensity λ and μ. Let $(N_t^Z)_t$ be the process obtained by addition of $(N_t^X)_t$ and $(N_t^Y)_t$, i.e. $N_t^Z = N_t^X + N_t^Y$. Then, $(N_t^Z)_t)$ is a *Poisson* process of intensity $\lambda + \mu$.

3.1 Markov Chains

Def 9 : DTMC

A DTMC is a stochastic process $\{X_n,\ n \in \mathbb{N}\}$ where X_n is the state (discrete) time step n such that $\forall n \in \ltimes$, $\forall i, j$, $\forall i_0, ..., i_{n-1}$ we have:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) \quad = \quad \mathbb{P}(X_{n+1} = j \mid X_n = i) \quad = \quad P_{i,j}$$

Def 10 : Markovian property

The conditional distribution of any future state X_{n+1}, given the past states $X_0, ..., X_{n-1}$ and the current state X_n is independant of the past and only depend on the present state X_n.

Def 11 :

The **transition probability matrix** associated with any DTMC P whose entry (i, j) denoted by $P_{i,j}$ represents the probability of moving to state j on the next transition given that the current state is i. Then $\forall i$, $\sum\limits_j P_{i,j} = 1$

Property 5 :

$$\lim_{n \to +\infty} P^n_{i,j} \quad = \quad (\lim_{n \to +\infty} P^n)_{i,j}$$

Def 12 : Limiting probability and distribution

Let $\Pi_j = \lim\limits_{n \to +\infty} P^n_{i,j}$. Then Π_j represents the **limiting probability** that the chain is in state j (independently of the state i). For a M-state DTMC with state $1, ..., M$, $\vec{\Pi} = (\Pi_1, ..., \Pi_M)$ where $\sum\limits_{i=1}^{M} \Pi_i = 1$ represents the **limiting distribution** of being in each state.

3.2 Stationary distributions

Def 13 : Stationary distribution

A probabilistic distribution $\vec{\Pi} = (\Pi_1, ..., \Pi_M)$ is said to be **stationary** for the M.C. if $\vec{\Pi}.P = \vec{\Pi}$ and $\sum\limits_{i=1}^{M} \Pi_i = 1$. These is the **stationary equations**.

Theorem 3 : Stationary distribution

Given a finite state DTMC with M states, let $\Pi_j = \lim\limits_{n \to +\infty} P^n_{i,j} > 0$ be the limiting probability of being in state j and let $\vec{\Pi} = (\Pi_1, ..., \Pi_M)$ where $\sum\limits_{i=1}^{M} \Pi_i = 1$ be the limiting distribution. Assuming that the limiting distribution exists, $\vec{\Pi}$ is also a stationary distribution and it is unique.

Proof :

$\Pi_j.P = \lim_{n\to+\infty} P_{i,j}^{n+1} = \lim_{n\to+\infty} \sum_{k=1}^{M} P_{i,k}^n.P_{k,j} = \sum_{k=1}^{M} \lim_{n\to+\infty} P_{i,k}^n.P_{k,j} = \sum_{k=1}^{M} \Pi_k.P_{k,j} = \Pi_j$

Then, $\vec{\Pi}.P = \vec{\Pi}$.

Let $\vec{\Pi}'$ be a stationary distribution. Let $\vec{\Pi}$ is the limiting distribution. Let's show that $\vec{\Pi}'$ and $\vec{\Pi}$ are equals. Let's assume we start at time $\varnothing$ with distribution $\vec{\Pi}'$.

$\Pi'_j = \mathbb{P}(X_0 = j) = \mathbb{P}(X_n = j)$.

Then, $\forall n \in \mathbb{N}$, $\Pi'_j = \sum_{i=1}^{M} \mathbb{P}(X_n = j \mid X_0 = i).\mathbb{P}(X_0 = i) = \sum_{i=1}^{M} P_{i,j}^n.\Pi'_i$

Nevertheless $\Pi'_j = \lim_{n\to+\infty} \Pi'_j = \lim_{n\to+\infty} \sum_{i=1}^{M} P_{i,j}^n.\Pi'_i = \sum_{i=1}^{M} \lim_{n\to+\infty} P_{i,j}^n.\Pi'_i = \sum_{i=1}^{M} \Pi_j.\Pi'_i = \Pi_j \sum_{i=1}^{M} \Pi_i = \Pi_j$

Then, the stationary distribution is unique.

■

Summary: Finding the limiting probability in a finite state DTMC

By the previous theorem [3.2], given that the limiting distribution $\{\Pi_j\ ,\ 1 \leqslant j \leqslant M\}$ exist, we can obtain it by solving the sattionary equations $\vec{\Pi}.P = \vec{\Pi}$ and $\sum_{i=1}^{M} \Pi_i = 1$ where $\vec{\Pi} = (\Pi_1, ..., \Pi_M)$.

Def 14 : Stationary Markov Chains

A MC for which the limiting probabilities exist is said to be **stationary** (or in **steady state**) if the initial state is chosen according to the stationary probabilities.

3.3 Infinite state DTMC

Now, $\vec{\Pi} = (\Pi_0, ...,)$ and $\Pi_j = \lim_{n\to+\infty} P_{i,j}^n$ and $\sum_{j=0}^{+\infty} \Pi_j = 1$.

The result of therorem 3.2 still holds for infinite state DTMC.

Example:

Suppose an unbounded queue system when at each time step, with probability $p = \frac{1}{40}$, one job arrives and, independently, with probability $q = \frac{1}{30}$, one job departs.

Q: what is the average number of jobs in the system ?

Let's construct an infinite DTMC A such that $\forall i, j \in \mathbb{N}$, $A_{i,j}$ is the probability that there is currently i jobs in the system and it will be j after the time step.

Therefore, $\forall i \in \mathbb{N}^*$, $A_{i-1,i} = p(1-q) = r$, $A_{i,i-1} = q(1-p) = s$, $A_{i,i} = 1 - r - s$ and $A_{0,0} = 1 - r$.

The stationnary equations are:

$\Pi_0 = (1-r)\Pi_0 + s\Pi_1$

$\Pi_1 = rPi_0 + (1-r-s)\Pi_1 + s\Pi_2$

$\vdots$

and $\Pi_0 + \Pi_1 + ... = 1$

Then, $\Pi_1 = \frac{r}{s}\Pi_0$, $\Pi_2 = (\frac{r}{s})^2\Pi_0$, $\Pi_3 = (\frac{r}{s})^3\Pi_0$. In fact $\forall i \in \mathbb{N}$, $\Pi_i = (\frac{r}{s})^i\Pi_0$

To verify that, $\forall i > 0$, Π_i has to be equal to $r\Pi_{i-1} + (1-r-s)\Pi_i + s\Pi_{i+1}$. In fact, that is can be verify through calculus.

To find Π_0, we exploit the last stationnary equation : $\sum_{i=0}^{+\infty} \Pi_i = 1$. Then $\Pi_0 \sum_{i=0}^{+\infty} (\frac{r}{s})^i = 1$. Therefore, $\Pi_0 = 1 - \frac{r}{s}$. Then $\Pi_i = (\frac{r}{s})^i.(1 - \frac{r}{s})$.

Let's call N the number or jobs in the system and $\alpha = \frac{r}{s}$.
Then $\mathbb{E}(N) = \sum_{i=0}^{+\infty} i\Pi_i = \sum_{i=0}^{+\infty} i\alpha^i(1-\alpha) = \alpha.(1-\alpha)\sum_{i=1}^{+\infty}(i.\alpha^{i-1}) = \alpha.(1-\alpha)\frac{d}{dp}(\frac{1}{1-\alpha}) = \frac{\alpha}{1-\alpha}$

3.4 Ergodicity questions

Def 15 : Period

The **period** of a state j is the greatest common divisor (GCD) of the set of integers n such that $P_{i,j}^n > 0$.
A state is **aperiodic** is it has period 1.
A chain is **aperiodic** if all its states are aperiodic.

Def 16 : accessibility

A state j is **accessible** from i if $\exists n \in \mathbb{N}$ such that $P_{i,j}^n > 0$.
Two states i and j communicate if i is accessible from j and conversly.
A MC is **irreductible** if all its states communicate with each other.

Theorem 4 :

For an irreductible aperiodic finite state MC with transition matrix P, $m_{j,j} = \frac{1}{\Pi_j}$???

Def 17 : Recurrent & transiant

f_j is the probability that a chain starting in state j ever return to state j.
If $f_j = 1$, j is said to be **recurrent** otherwise, it is **transient**.
Recurrent MCs fall into two types:

- **Positive recurrent**: The wait time between recurrences is finite.
- **Null recurrent**: The wait time between recurrences is null.

Def 18 : Ergodic DTMC

An **ergodic** DTMC is one that have the 3 following properties:

- is aperiodic
- is irreductible
- is positive recurrent

Theorem 5 : Ergodic thm of DTMC

Given a recurrent, aperiodic and irreductible DTMC, $\Pi_j = \lim_{n\to+\infty} P_{i,j}^n$ exists and $\Pi_j = \frac{1}{m_{j,j}}$
For a positive recurrent, aperiodic and irreductible DTMC, $\forall j$, $\Pi_j > 0$.