EE126: Probability and Random Processes

Lecture 20: Markov Processes

Abhay Parekh

UC Berkeley

April 5, 2011

Midterm

Will cover Sections 3.4-6.2 of the book, i.e. what we covered in Lecture from

- No questions will explicitly be on earlier material
- However, there are many dependencies!
- Same cheatsheet, calculator rules as last times.
- Read lecture notes and book. Do problems.
- Review Sessions: Sunday (4-6, Room TBA) and Monday. Stay tuned for bspace announcements.

Logistics

- Lecture notes should be updated.
- Old exams have been posted.
- HW 6 returned today, HW 7 on Thursday.

Example: Spam Folders

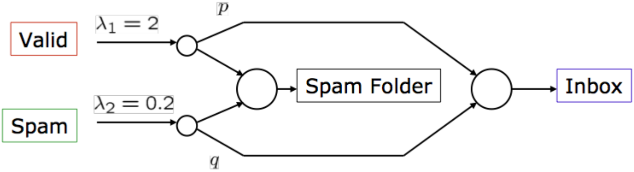
Valid mail and Spam mail both arrive to your inbox according to independent Poisson Processes of $\lambda_v = 4$ (4 per hour) and $\lambda_s = 2$ respectively.

Your spam filter tags a spam email correctly with prob $p = 0.8$; a valid email as spam with prob $q = 0.1$

Example: Spam Folders

Valid mail and Spam mail both arrive to your inbox according to independent Poisson Processes of $\lambda_v = 4$ (4 per hour) and $\lambda_s = 2$ respectively.

Your spam filter tags a spam email correctly with prob $p = 0.8$; a valid email as spam with prob $q = 0.1$



1. Prob an email in your inbox is spam:
2. Prob an email in your spam filter is valid
3. How often should you check your spam folder if you'd like to see one new valid email each time you check?

Conditional Distribution of arrival times

Given that $N(t) = n$, what is the joint distribution of the n arrival times $Y_1, Y_2, ..., Y_n$? Let us call this distribution $f(t_1, t_2, \ldots, t_n)$:

$$f(t_1, ..., t_n) = \frac{n!}{t^n}.$$

We showed that if $X_1, ..X_n$ are iid uniform over $[0, t]$:

$$f_{X_{(1)}, X_{(2)}, \cdots, X_{(n)}}(x_1, ..., x_n) = \frac{n!}{t^n}$$

So given that n arrivals have occurred, $Y_1, .., Y_n$ considered as unordered random variables are iid uniform.

Random Incidence

Pick a random time t. What is the distribution of the inter-arrival time that includes t?
By similar reasoning as for Bernouli, the interval is distributed as an Erlang of order 2.

Random Incidence

Pick a random time t. What is the distribution of the inter-arrival time that includes t?

By similar reasoning as for Bernouli, the interval is distributed as an Erlang of order 2.

What about a general argument? i.e. what if interarrivals are not exponential, but distributed iid according to $f_Y(y)$

Random Incidence

Pick a random time t. What is the distribution of the inter-arrival time that includes t?

By similar reasoning as for Bernouli, the interval is distributed as an Erlang of order 2.

What about a general argument? i.e. what if interarrivals are not exponential, but distributed iid according to $f_Y(y)$

Let w be the length of the interarrival interval that includes t:

The random time is more likely to fall in a larger interval:

Random Incidence

Pick a random time t. What is the distribution of the inter-arrival time that includes t?

By similar reasoning as for Bernouli, the interval is distributed as an Erlang of order 2.

What about a general argument? i.e. what if interarrivals are not exponential, but distributed iid according to $f_Y(y)$

Let w be the length of the interarrival interval that includes t:

The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) =$$

Random Incidence

Pick a random time t. What is the distribution of the inter-arrival time that includes t?

By similar reasoning as for Bernouli, the interval is distributed as an Erlang of order 2.

What about a general argument? i.e. what if interarrivals are not exponential, but distributed iid according to $f_Y(y)$

Let w be the length of the interarrival interval that includes t:

The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) = \frac{w f_Y(w)}{E[Y]}$$

Random Incidence

Let w be the length of the interarrival interval that includes t:
The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) = \frac{w f_Y(w)}{E[Y]}$$

- $E[W] =$

Random Incidence

Let w be the length of the interarrival interval that includes t:
The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) = \frac{w f_Y(w)}{E[Y]}$$

- $E[W] = \frac{E[Y^2]}{E[Y]}$. Random incidence effects exist unless variance is zero.

Random Incidence

Let w be the length of the interarrival interval that includes t:
The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) = \frac{w f_Y(w)}{E[Y]}$$

- $E[W] = \frac{E[Y^2]}{E[Y]}$. Random incidence effects exist unless variance is zero.
- Let v be the time to the next arrival after t

$$f_{v,w}(v, w) = f_{V|W}(v|w) f_W(w)$$

Random Incidence

Let w be the length of the interarrival interval that includes t:
The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) = \frac{w f_Y(w)}{E[Y]}$$

- $E[W] = \frac{E[Y^2]}{E[Y]}$. Random incidence effects exist unless variance is zero.
- Let v be the time to the next arrival after t

$$f_{v,w}(v,w) = f_{V|W}(v|w) f_W(w) = \frac{1}{w} \frac{w f_Y(w)}{E[Y]} = \frac{f_Y(w)}{E[Y]}$$

Random Incidence

Let w be the length of the interarrival interval that includes t:
The random time is more likely to fall in a larger interval:

$$f_w(w) \propto w f_Y(w) \Rightarrow f_w(w) = \frac{w f_Y(w)}{E[Y]}$$

- $E[W] = \frac{E[Y^2]}{E[Y]}$. Random incidence effects exist unless variance is zero.
- Let v be the time to the next arrival after t

$$f_{v,w}(v,w) = f_{V|W}(v|w) f_W(w) = \frac{1}{w}\frac{w f_Y(w)}{E[Y]} = \frac{f_Y(w)}{E[Y]}$$

- $$f_V(v) = \int_v^{\infty} f_{v,w}(v,w)dw = \frac{1 - F_Y(v)}{E[Y]}$$

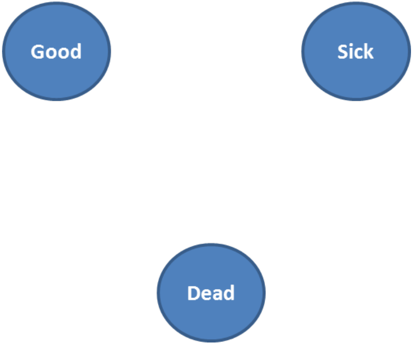
Random Incidence Example

$E[W] = \frac{E[Y^2]}{E[Y]}$ Compare distributions when $E[Y] = 0.5$

Uniform [0,1]:	$\frac{2}{3}$
Fixed at 0.5	0.5
Exponential $\lambda = 2$	1
K-Erlang $\lambda = 2k$	$\frac{1}{2} + \frac{1}{2k}$

Markov Processes

- Future **depends** on past. However, future value depends only the current state of the system.
- Example: Machine Health



Discrete Time Markov Chains

At each integer time, j, the process is in a state X_j, where X_j is in a finite **state space** $S = \{1, 2, ..., m\}$.

Discrete Time Markov Chains

At each integer time, j, the process is in a state X_j, where X_j is in a finite **state space** $S = \{1, 2, ..., m\}$.

Transition Probabilities:

$$p_{ij} = P(X_{n+1} = j | X_n = i) \quad i, j \in S$$

Discrete Time Markov Chains

At each integer time, j, the process is in a state X_j, where X_j is in a finite **state space** $S = \{1, 2, ..., m\}$.

Transition Probabilities:

$$p_{ij} = P(X_{n+1} = j | X_n = i) \quad i, j \in S$$

As long as we know the initial probabilities, $P(X_0 = i)$, the transition probabilities are sufficient to study the system since:

$$\boxed{P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)}$$

for all times and states. **This is the Markov Property.**

Discrete Time Markov Chains

At each integer time, j, the process is in a state X_j, where X_j is in a finite **state space** $S = \{1, 2, ..., m\}$.

Transition Probabilities:

$$p_{ij} = P(X_{n+1} = j | X_n = i) \quad i, j \in S$$

As long as we know the initial probabilities, $P(X_0 = i)$, the transition probabilities are sufficient to study the system since:

$$\boxed{P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)}$$

for all times and states. **This is the Markov Property.**

Clearly:

$$\sum_j p_{ij} = 1$$

The matrix $[p]_{ij}$ is called the **transition matrix**.

Example: Machine Failure and Replacement

A machine is either broken or working. If it is working it will break down with prob b. If it is broken down it will be fixed and working the next day with prob f. If it is broken for 4 days it is replaced with a new (working) machine.

Reflecting Random Walk

Bob is in a tiled corridor of length m tiles. In each time step, he walks to next numbered tile with prob p unless he lands on tile m. He walks to the next lowered numbered tile with prob $1 - p$ unless he lands on tile 1.

Finite Buffer Queue

Packets arrive according to Poisson process with rate λ at a node with a buffer size of m packets. The time taken to serve a packet is exponentially distributed with rate μ. Model this system as a MC.

Finite Buffer Queue

Packets arrive according to Poisson process with rate λ at a node with a buffer size of m packets. The time taken to serve a packet is exponentially distributed with rate μ. Model this system as a MC. Consider a very small interval. Then, prob of an arrival is $\lambda\delta$. The probability of a departure, given that there is at least one packet present is $\mu\delta$.

Spiders and Fly

A spider moves along a line to left with probability .3 and to the right with prob .4. If it reaches either end of the line it is eaten by one of two spiders. If the line is of length 4:

2-Step Transition Probabilities

The transition matrix tells us what can happen in 1 step. What about in two steps? If there are m states then for all states i, j

$$P(X_2 = j | X_0 = i) = \sum_{k=1}^{m} p_{ik} p_{kj}$$

2-Step Transition Probabilities

The transition matrix tells us what can happen in 1 step. What about in two steps? If there are m states then for all states i, j

$$P(X_2 = j | X_0 = i) = \sum_{k=1}^{m} p_{ik} p_{kj}$$

Therefore for any l:

$$r_{ij}(2) = P(X_{l+2} = j | X_l = i) = \sum_{k=1}^{m} p_{ik} p_{kj}$$

2-Step Transition Probabilities

The transition matrix tells us what can happen in 1 step. What about in two steps? If there are m states then for all states i, j

$$P(X_2 = j | X_0 = i) = \sum_{k=1}^{m} p_{ik} p_{kj}$$

Therefore for any l:

$$r_{ij}(2) = P(X_{l+2} = j | X_l = i) = \sum_{k=1}^{m} p_{ik} p_{kj}$$

Now let's look at the ij^{th} entry of P^2: Product of the i^{th} row and j^{th} column:

$$p_{ij}^{(2)} = \sum_{k=1}^{m} p_{ik} p_{kj}$$

So

$$r_{ij}(2) = p_{ij}^{(2)}$$

n-Step Transition Probabilities

We have the recursion:

$$\begin{aligned} P(X_n = j | X_0 = i) &= \sum_{k=1}^{m} P(X_{n-1} = k | X_0 = i) p_{kj} \\ r_{ij}(n) &= \sum_{k=1}^{m} r_{ik}(n-1) p_{kj} \end{aligned}$$

n-Step Transition Probabilities

We have the recursion:

$$\begin{aligned} P(X_n = j | X_0 = i) &= \sum_{k=1}^{m} P(X_{n-1} = k | X_0 = i) p_{kj} \\ r_{ij}(n) &= \sum_{k=1}^{m} r_{ik}(n-1) p_{kj} \end{aligned}$$

And

$$p_{ij}^{(n)} = \sum_{k=1}^{m} p_{ik}^{(n-1)} p_{ki}$$

So P^n gives the n-step transition probabilities.

Chapman Kolmogorov Equation

The $n-$step transition probabilities for an m-state DT Markov Chain are given by:

$$r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)p_{kj}$$

for $n > 1$ and all i, j, starting with

$$r_{ij}(1) = p_{ij}$$

The $n-$step transition probabilities are given by P^n.

Example

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}$$

Steady State Dynamics

What happens as $n \to \infty$?

- Does the limit $r_{ij} = \lim_{n\to\infty} r_{ij}(n)$ exist?
- If so, does r_{ij} depend on i? I.e. the initial state?
- If we observe the MC at some random time n (n large), what is the probability that we find it in state i?

Simple Example

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}, \quad P^2 = \begin{pmatrix} 0.6 & 0.4 \\ 0.56 & 0.44 \end{pmatrix},$$

Simple Example

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}, \quad P^2 = \begin{pmatrix} 0.6 & 0.4 \\ 0.56 & 0.44 \end{pmatrix}, \quad P^3 = \begin{pmatrix} 0.58 & 0.42 \\ 0.588 & 0.412 \end{pmatrix},$$

Simple Example

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}, \quad P^2 = \begin{pmatrix} 0.6 & 0.4 \\ 0.56 & 0.44 \end{pmatrix}, \quad P^3 = \begin{pmatrix} 0.58 & 0.42 \\ 0.588 & 0.412 \end{pmatrix},$$

$$P^4 = \begin{pmatrix} 0.584 & 0.416 \\ 0.5824 & 0.4176 \end{pmatrix}, \quad P^5 = \begin{pmatrix} 0.5833 & 0.4167 \\ 0.5833 & 0.4167 \end{pmatrix},$$

Simple Example

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}, \quad P^2 = \begin{pmatrix} 0.6 & 0.4 \\ 0.56 & 0.44 \end{pmatrix}, \quad P^3 = \begin{pmatrix} 0.58 & 0.42 \\ 0.588 & 0.412 \end{pmatrix},$$

$$P^4 = \begin{pmatrix} 0.584 & 0.416 \\ 0.5824 & 0.4176 \end{pmatrix}, \quad P^5 = \begin{pmatrix} 0.5833 & 0.4167 \\ 0.5833 & 0.4167 \end{pmatrix},$$

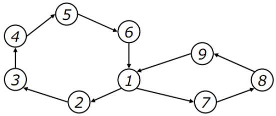
$r_{j1}(n) \to 0.5833$ and $r_{j2} \to 0.4167$ independent of X_0.
Is this true in general?

Periodic States

1. 3 states:

2. 4 states:

3. k classes:



Example:

Absorbing States

Spider and Fly:

Absorbing States

Spider and Fly:If the fly starts in states 1 or 4 it never visits any other state. I.e.,

$$r_{1j} = 0,\ j = 2, 3, 4 \quad r_{4j} = 0,\ j = 1, 2, 3$$

Absorbing States

Spider and Fly:If the fly starts in states 1 or 4 it never visits any other state. I.e.,

$$r_{1j} = 0,\ j = 2, 3, 4 \quad r_{4j} = 0,\ j = 1, 2, 3$$

So the limit depends on the initial state.

Communicating States

Accessible States: If i and j are connected by a directed path from i to j in the MC then $r_{ij}(n) > 0$ for large enough n and j is accessible from i.

A state in a MC is either Recurrent or Transient:

1. **Recurrent States:** Let $A(i)$ be the states accessible from i. Then i is recurrent if for every state $j \in A(i)$ it holds that $i \in A(j)$.
   If i is recurrent then every state in $A(i)$ must be recurrent as well.

Communicating States

Accessible States: If i and j are connected by a directed path from i to j in the MC then $r_{ij}(n) > 0$ for large enough n and j is accessible from i.

A state in a MC is either Recurrent or Transient:

1. **Recurrent States:** Let $A(i)$ be the states accessible from i. Then i is recurrent if for every state $j \in A(i)$ it holds that $i \in A(j)$.
   If i is recurrent then every state in $A(i)$ must be recurrent as well.
2. **Transient States:** There is a state $j \in A(i)$ that cannot access i.

Example: Spiders and Fly...

MC State Properties

1. **Any state j of a MC must be accessible to at least one recurrent state:** If j is recurrent then...

MC State Properties

1. **Any state j of a MC must be accessible to at least one recurrent state:** If j is recurrent then...
   If j is transient: Pick a state in $A(j)$ that cannot access j. Call this j_1 if j_1 is not recurrent then pick a state in $A(j_2)$ that cannot access j_1. Can $j_2 = j$?

MC State Properties

1. **Any state j of a MC must be accessible to at least one recurrent state:** If j is recurrent then...
   If j is transient: Pick a state in $A(j)$ that cannot access j. Call this j_1 if j_1 is not recurrent then pick a state in $A(j_2)$ that cannot access j_1. Can $j_2 = j$?
   $j_2 \neq j_1$. Since there are only a finite number of states, this process must terminate at some state j_n, that is recurrent.

MC State Properties

1. **Any state j of a MC must be accessible to at least one recurrent state:** If j is recurrent then...
   If j is transient: Pick a state in $A(j)$ that cannot access j. Call this j_1 if j_1 is not recurrent then pick a state in $A(j_2)$ that cannot access j_1. Can $j_2 = j$?
   $j_2 \neq j_1$. Since there are only a finite number of states, this process must terminate at some state j_n, that is recurrent.
2. **If a recurrent state j is visited once, it must be revisited an infinite number of times:** Once j is visited the first time, the MC stays in $A(i)$.
3. **If a state is Transient it can only visited a finite number of times.**

Markov Chain Decomposition

- A MC has at least one recurrent class
- If a MC has more than two recurrent classes two states from two different recurrent classes are inaccessible
- At least one recurrent state is accessible from a transient state.

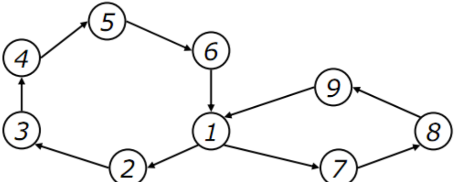
Recurrent Class Periodicity

The states in the Recurrent Class of a Markov Chain are periodic if they can be grouped into $d > 1$ groups so that all transitions from one group lead to the next group.
The period of a state i is defined by:

$$\gcd\{n : r_{ii}(n) > 0\}$$

If a recurrent class is not periodic then for any state i in the class then there must be a time n beyond which

$$r_{ii}(n) > 0$$



Steady State Convergence

Let's look at a MC with one aperiodic recurrent class and one or more transient classes. As $n \to \infty$, $r_{ij}(n) = 0$ for all the transient states. The states in the recurrent class will tend to to a positive limit.

Proof: Typically uses results from Linear Algebra:

Steady State Convergence

Let's look at a MC with one aperiodic recurrent class and one or more transient classes. As $n \to \infty$, $r_{ij}(n) = 0$ for all the transient states. The states in the recurrent class will tend to to a positive limit.

Proof: Typically uses results from Linear Algebra:Another way (See Problem 19 of the book.) Show that

$$|r_{ij}(n) - r_{kj}(n)| \leq c\gamma^n$$

for positive $c, \gamma < 1$. Then that the sequence $r_{ij}(n)$ converges to a limit.

Steady State Convergence

Let's look at a MC with one aperiodic recurrent class and one or more transient classes. As $n \to \infty$, $r_{ij}(n) = 0$ for all the transient states. The states in the recurrent class will tend to to a positive limit.

Proof: Typically uses results from Linear Algebra:Another way (See Problem 19 of the book.) Show that

$$|r_{ij}(n) - r_{kj}(n)| \leq c\gamma^n$$

for positive $c, \gamma < 1$. Then that the sequence $r_{ij}(n)$ converges to a limit.

Let $n \to \infty$, $r_{ij}(n) = \pi_j$ for all states j:

Then from the Chapman Kolmogorov Equations: for each state j:

$$\pi_j = \sum_{k=1}^{m} \pi_k p_{kj}$$

These are called the **Balance Equations**.

Also, $\sum_{k=1}^{m} \pi_k = 1$. Now we have $m+1$ equations and m

Examples

1. $$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}$$

Examples

1. $$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{pmatrix}$$

   $\pi_1 = 0.5\pi_1 + \pi_2 0.7$ and $\pi_1 + \pi_2 = 1$. So $\pi_2 = 5/12 = 0.4167$ and $\pi_1 = 7/12 = .5833$.

2. Transient + Recurrent

Example: Reflecting Random Walk

$$
\begin{aligned}
\pi_1 &= (1-p)\pi_1 + \pi_2(1-p) \\
\pi_2 &= (1-p)\pi_3 + p\pi_1 \\
\pi_3 &= (1-p)\pi_4 + p\pi_2 \\
\pi_4 &= p\pi_4 + p\pi_3 \\
\pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1
\end{aligned}
$$

$$\pi_i = \pi_1 (\frac{p}{1-p})^{i-1} \quad i = 2, 3, 4$$

Uniqueness of Balance Equations

Suppose there is another non-negative solution to the balance equations: $\bar{\pi}_j$.
Then initialize probabilities so that $P(X_0 = j) = \bar{\pi}_j$. Then

$$P(X_1 = j) = \sum_{k=1}^{m} P(X_1 = j | X_0 = k)\bar{\pi}_k = \sum_{k=1}^{m} p_{kj}\bar{\pi}_k = \pi_j$$

Similarly: $P(X_n = j) = \bar{\pi}_j$.
Thus:

$$\bar{\pi}_j = \lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} \sum_{k=1}^{m} \bar{\pi}_k r_{kj}(n)$$

$$\bar{\pi}_j = \sum_{k=1}^{m} \bar{\pi}_k \pi_j = \pi_j$$

Turns out that the balance equations yield unique values of π_j even if the recurrent class is periodic, but obviously that isn't the steady state distribution.

Steady State Convergence

Any MC with a single aperiodic recurrent class must converge in the sense that

1. For each state j:

$$\lim_{n\to\infty} r_{ij} = \pi_j, \quad \text{for all } i$$

2. π_j are given by the system of equations:

$$\begin{aligned} \pi_j &= \sum_{k=1}^{m} \pi_k p_{kj}, \quad j = 1, 2, ..., m \\ 1 &= \sum_{k=1}^{m} pi_k \end{aligned}$$

3. $\pi_j = 0$ for all transient states j.
$\pi_j > 0$ for all recurrent states j.

Birth Death Processes

- $m + 1$ states: $0, 1, 2, \ldots m$.
- $p_{ii+1} = b_i$ for $i = 0, 1, 2, \ldots, m - 1$
- $p_{ii-1} = d_i$ for $i = 1, 2, \ldots, m$
- $d_0 = b_m = 0$
- $p_{ii} = 1 - b_i - d_i$

$$\pi_0 = (1 - b_0)\pi_0 + d_1\pi_1 \Rightarrow \pi_1 = \frac{b_0}{d_1}\pi_0$$

$$\pi_1 = b_0\pi_0 + (1 - b_1 - d_1)\pi_1 + d_2\pi_2 \Rightarrow \pi_2 = \frac{b_0 b_1}{d_1 d_2}\pi_0$$

etc

$$pi_i = \frac{b_0 b_1 \ldots b_{i-1}}{d_1, \ldots, d_i}\pi_0, \quad i = 1, 2, \ldots, m$$

Also $\sum_i \pi = 1$, so the stationary distribution can be obtained.
Example: $p_i = d_i$ for $i = 1, 2, \ldots, m1$: $p_i = \frac{b_0}{d_i}\pi_0$