Stochastic Processes

MICHAEL CRAWFORD - 260 454 391

Background Material

The **conditional probability** of $Y = y$ given $X = x$ is
$P_{Y|X=x}(y|x) = \frac{P(X=x,Y=y)}{P(X=x)} = P(Y = y|X = x)$
For **descrete** RVs, the **conditional expectation** of Y given $X = x$ is
$E(Y|X = x) = \sum_y yP_{Y|X=x}(y|x) = \sum_y yP(Y = y, X = x)$
For **continuous** RVs, the **conditional expectation** of Y given $X = x$ is
$E(Y|X = x) = \int_{\mathbb{R}} yf_{Y|X=x}(y|x)$
The **conditional PDF** of Y given $X = x$ is
$f_{Y|X=x}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$
The **Total Law of Prob** says that

$$F_Y(y) = \begin{cases} \sum_x F_{Y|X=x}(y|x)P_X(x) & \text{(discrete case)} \\ \int_{\mathbb{R}} F_{Y|X=x}(y|x)f_X(x) & \text{(continuous case)} \end{cases}$$

It also says that
$P_Y(y) = \sum_x P_{Y|X=x}(y|x)P_X(x), \text{and}$
$f_Y(y) = \int_{\mathbb{R}} f_{Y|X=x}(y|x)f_X(x)$
The **Law of Total Exp** is summarized by
$E(Y) = E_X(E(Y|X))$
which can be broken up into discrete/cts cases:

$$E(Y) = \begin{cases} \sum_x E(Y|X = x)P_X(x) & \text{discrete case} \\ \int_{\mathbb{R}} E(Y|X = x)f_X(X)dx & \text{continuous case} \end{cases}$$

To compute $E(Y|X)$: a) compute $E(Y|X = x)$ and call this $g(x)$; b) then we let $E(Y|X) = g(X) = g(x)|_X$
Variance form of the law of total expectation:
$V(Y) = E_X[V(Y|X)] + V_X[E(Y|X)]$
If $X_i, i = 1, 2, \ldots, N$ and N are all indep and the X_i are i.i.d. with mean μ, then $E[\sum_{i=1}^N] = \mu E(N)$

Probability Generating Functions

The pgf of a discrete RV X is
$\phi_X(z) = E(z^X) = \sum_{k=0}^\infty z^k P_X(k) = \sum_{k=0}^\infty z^k P(X = x)$
Properties of pgf's:
a) $\phi_X(0) = P_X(0) = P(X = 0)$
b) $\phi_X(1) = 1$
c) $\phi_X(z)$ is continuous, increasing and convex
d) $\phi'_X(0) = P_X(1)$
e) In general, $P_X(k) = \frac{\phi^{(k)}(0)}{k!}$, where $\phi^{(k)} = \frac{d^k}{dz^k}\phi$ For $S_n = \sum_{i=1}^n X_i$ where X_i are indep, we have
$\phi_{S_n}(z) = \prod_{i=1}^n \phi_{X_i}(z) = (\text{if also ident. dist}) = (\phi_X(z))^n$

General Into to Stochastic Processes

Stochastic $\Rightarrow$ Random/non-deterministic
Formally, a **stochastic process indexed by a set** T is a family of RVs $\{X_t : t \in T\}$ such that for each $t \in T$, X_t is a RV
In order to **specify** a stochastic process, one only needs to specify the joint CFD $F_{X_{t_x},\ldots,X_{t_n}}(x_{t_1}, \ldots, x_{t_n})$ for all of the subsets $(t_1, \ldots, t_n)$ of t_i's in T
The set T is called the **parameter space**
The set of values that X_t can assume is called the **state space**
T can be either countable or uncountable, in which case the stoch. pro. is called eiter a **discrete paramater process** or a **continuous parameter process** (resp)
The state space $\{X_t\}_{t\in T}$ can be either countable or uncountable, in which case the stoch. pro. is called either a **discrete state space process** or a **continuous state space process** (resp) Using these two classifiers, we combine them to get a total of **four classes** of stochastic processes

Branching Processes

Specifically, we consider the GALTON-WATSON BRANCHING PROCESS
A stochastic $\{X_t : t = 1, 2, \ldots\}$ is called a **branching process** if it satisfies the following:
a)$X_0 = 1$
b) The indiv. at t can give rise to j indiv's at $t + 1$ with prob $p_j, j = 1, 2, \ldots$
c) The p_j's are assumed to be the same during each time period
d) The prob that an indiv. will give rise to j indiv's is assumed to be independent of how many indiv's there are at that time
In a branching process, $X_n = \#$ OF INDIV'S AT TIME $t = n$
We have two useful identities:
$\phi_{X_n}(z) = \phi_{X_{n-1}}(\phi(z))$
$\phi_{X_n}(z) = \phi(\phi_{X_{n-1}}(z))$
The probability of extinction, η, is the smallest root satisfying the equation $z = \phi(z)$.
$\phi'(1) \le 1 \Rightarrow \eta = 1$
$\phi'(1) > 1 \Rightarrow \eta < 1$
The equation $\mu_n = \mu \cdot \mu_{n-1} \Rightarrow \mu_n = \mu^n$ tells us that if $\mu_n > 1$, then the **average increase** in pop size will be geometric (i.e. super fast) and if $\mu < 1$, we get the converse.

A useful equation: $\sigma_n^2 = \begin{cases} \mu^{n-1}\sigma^2(1-\mu^n)/(1-\mu) & \text{if } \mu \ne 1 \\ n\sigma^2 & \text{if } \mu = 1 \end{cases}$

If there are initially a indiv's, then the probabilty of extinction is
$\eta^2 \cdot [P(\text{EXTINTION OF LINE } 1 \cap \cdots \cap \text{EXTINCTION OF LINE } a) - \text{INDEP. OF LINES}]$
So the probability of non-extinction is $1 - \eta^a$.

Markov Chains (MC)

Attn is restricted to DISCRETE PARAMETER, CONTABLE STATE SPACE stoch pros here
The stoch. pro. $\{X_n : n = 0, 1, \dot{a}\}$ is called a **discrete parameter MC** if
$P[X_{n+1} = j|X_0 = i_0, \ldots, X_n = i_n] = P[X_{n+1} = j|X_n = i] = P_{ij}$
This is called the **Markov Property**. The RHS is caleld the **transition probabilty** P_{ij}. All of the transition probabilities are stored in the **transition matrix** $(P)_{ij}$
A MC is **homogeneous** or **stationary** if
$P[X_{n+1} = j|X_n = i] = P[X_1 = j|X_0 = i]$ - i.e. P_{ij} does not depend on n
Useful equation: $\sum_j P_{ij} = 1$
$\tilde{P}_i = (P_{0,i}, P_{1,i}, dots)$
$P_{ij}^n = P[X_{k+n} = j|X_k = i]$ is the **probability of going from i to j in n steps**
The **Chapman - Kolmogorov Equations** are:
$P_{ij}^n = \sum_{k=1}^\infty P_{ij}^n P_{jk}^m$
$P^{(m+n)} = P^{(n)}P^{(m)}$
$P^{(n)} = P^n$
State i is **accessible** from state j (written $i \to j$) if $\exists\ n > 0$ s.t. $P_{ij}^n > 0$.
States i and j **communicate** (written $i \leftrightarrow j$) iff $i \to j$ and $j \to i$
The relation '$\leftrightarrow$' is an equivalence relation, so we have
a) $i \leftrightarrow i$
b) $i \to j \Rightarrow j \to i$
c) $i \to j$ and $j \to k \Rightarrow i \to k$
Markov Chain Triangle Inequality: $P_{ik}^{(m+n)} \ge P_{ij}^m P_{jk}^n$
The set of all states that communicate with each other is called a **communicating class**
If all states in a MC communicate with each other then the MC is called **irreducible**
f_i is the prob that an MC will return to state i if it starts in state
A state i in an MC is called **recurrent** if $f_i = 1$ and **transient** if $f_i < 1$
If an MC starts in a recurrent state i, then it will re-enter state i infinitely-many times with probability 1
If an MC starts in a transient state i, then
$P[X_k \text{IS IN STATE } i \text{FOR } n \text{TIME PERIODS}] = f_i^{n-1}(1 - f_i)$
If a state is transient, then with probability 1 there will be only a finite number of visits to that state
If i is transient, then the expected $\#$ of visits in state i is $1/(1 = f_i)$ (the mean of the geometric distribution with $p = 1 - f_i$
State i is **recurrent** if $\sum_{n=1}^\infty P_{ii}^n = \infty$
State i is **transient** if $\sum_{n=1}^\infty < \infty$
If $i \leftrightarrow j$, then either a) both i and j are recurrent; or b) both i and j are transient
In an irreducible MC, either all states are transient or all states are recurrent.

Simple Random Walk

Let $X_1, X_2, \ldots$ be i.i.d. random variables s.t.

$$P(X_i = x) = \begin{cases} p & x = +1 \\ q = 1 - p & x = -1 \end{cases}$$

If we definte $S_n = X_1 + \cdots + X_n$ then the stoch. pro. $\{S_n : n = 0, 1, 2, \ldots\}$ is a **simple random walk on the intgers**
The probability that our random walk is on position k after n steps is
$P(S_n = k) = \binom{n}{\frac{k+n}{2}} p^{(k+n)/2} q^{(n-k)/2}$
In the SRW, all states communicate, so its an irreducible MC

Long Run Behaviour of Markov Chains

A class A of states is said to be **closed** if

$$P[X_{n+1} \in A|X_n = i] = \begin{cases} 1 & i \in A \\ 0 & i \notin A \end{cases}$$

A MC is **indecomposable** if the state space does not contant 2 or more disjoint classes of closed states
If a chain is IRREDUCIBLE then it is INDECOMPOSABLE
If for every pair of stats i and j either $i \to j$ or $j \to i$ holds, the chain is indecomposable
The **period**, $d(i)$, of a state i of a MC is the GCD of *all* $n > 0$ for which $P_{ii}^n > 0$
If $d(i) = 1$ for all states, then the MC is **aperiodic**
For a recurrent state i, let $f_i i^n$ be the probability of returning to state i *for the first time* in n steps. Then, the probability of *eventual return* to state i from i is $f_{ii} = \sum_{n=1}^{\infty} f_{ii}^n$
The expected time to return to state i for the first time is called the **mean recurrence time** and is denoted m_i
$m_i < \infty \Rightarrow i$ is **positive-recurrent**
$m_i = \infty \Rightarrow i$ is **null-recurrent**
It is not contradictory to simultaneously have a prob. of return equal to 1 and a mean return time equal to 0

Stationarity

We call $\tilde{\pi} = \{\pi_1, \pi_2, \dots\}$ a **stationary initial distribution** of a MC if for $i = 1, 2, \dots$ we have: a) $\sum_{j=1}^{\infty} \pi_i P_{ij}$; and b) $\sum \pi_i = 1$, $\pi_i \geq 0$
In matrix notation, these conditions are written as $\tilde{\pi} = \tilde{\pi} P$ and $\tilde{\pi}\tilde{e} = 1$, where $\tilde{e} = \{1, 1, \dots\}$
THEOREM: Take a MC that is *irreducible, recurrent*, and *aperiodic*: 1) If it is *positive-recurrent*, then $\lim_{n\to\infty} P_{ij}^n = \pi_i = 1/m_i < \infty$ and the π_i are unique sols of the above system of equations
2) If is is *null-recurrent*, then $\lim_{n\to\infty} P_{ij}^n = 0$
3) If it is irreducible, aperiodic and *transient*, then $\lim_{n\to\infty} P_{ij}^n = 0 \ \forall\ i, j$