$$\limsup_n A_n := \bigcap_m \bigcup_{n\geq m} A_n = \lim_{m\to\infty} \bigcup_{n\geq m} A_n$$
$$\liminf_n A_n := \bigcup_m \bigcap_{n\geq m} A_n = \lim_{m\to\infty} \bigcap_{n\geq m} A_n$$

We can interpret these as follows. The lim sup of A_n is the set of events that are contained in infinitely many of the A_n (but not necessarily in all the A_n past a point – the event could "bounce in and out" of the A_n)

$$\left\{\limsup_n A_n\right\} = \left\{\omega \in \Omega : \omega \in A_n \text{ i.o.}\right\}$$

Similarly, the lim inf of A_n is the set of events that eventually appear in an A_n and then in *all* A_n past that point.

$$\left\{\liminf_n A_n\right\} = \left\{\omega \in \Omega : \omega \in A_n \text{ ev.}\right\}$$

Clearly, if an event is in the lim inf, it also appears infinitely often (because it appears in all the A_n past a point, and so $\liminf_n A_n \subseteq \limsup_n A_n$.

Remark: Take $\left\{A_n, \text{ev.}\right\} = \left\{\liminf_n A_n\right\}$. Then

$$\begin{aligned}\left\{A_n,\ \text{ev.}\right\}^c &= \left[\bigcup_m \bigcap_{n\geq m} A_n\right]^c \\ &= \bigcap_m \left\{\left(\bigcap_{n\geq m} A_n\right)^c\right\} \\ &= \bigcap_m \bigcup_{n\geq m} A_n^c \\ &= \liminf_n A_n^c \\ &= \left\{A_n^c,\ \text{i.o.}\right\}\end{aligned}$$

- ***Proposition***: Let $\left\{A_n\right\}$ be a sequence of measurable sets, then
  - $\mathbb{P}\left(\liminf_n A_n\right) \leq \liminf_n \mathbb{P}\left(A_n\right)$
  - $\mathbb{P}\left(\limsup_n A_n\right) \geq \limsup_n \mathbb{P}\left(A_n\right)$

  (The first statement can be thought of as a generalization of Fatou's Lemma for probabilities).

  Proof of (i): Let us define $B_m = \bigcap_{n\geq m} A_n$. Then we know that $\cdots \subseteq B_m \subseteq B_{m+1} \subseteq \cdots$. In other words, the sets B_m increase monotonically to

$$B = \bigcup_m B_m = \bigcup_m \bigcap_{n\geq m} A_n = \liminf_n A_n$$

  Since the events are increasing, a simple form of monotone convergence gives us that $\mathbb{P}\left(B_n\right) \nearrow \mathbb{P}\left(B\right)$. But we also have that

$$\mathbb{P}\left(B_m\right) \leq \mathbb{P}\left(A_n\right) \qquad n \geq m$$

$$\begin{aligned}\mathbb{P}\left(\bigcap_{n\geq m} A_n^c\right) &= \prod_{n\geq m} \mathbb{P}\left(A_n^c\right) \\ &= \prod_{n\geq m} 1-\mathbb{P}\left(A_n\right) \\ &\leq \prod_{n\geq m} \exp\left(-\mathbb{P}(A_n)\right) \\ &= \exp\left(-\sum_{n\geq m} \mathbb{P}(A_n)\right) \\ &= 0\end{aligned}$$

But we have that

$$\begin{aligned}\mathbb{P}\left(\bigcup_m \bigcap_{n\geq m} A_n^c\right) &\leq \sum_m \mathbb{P}\left(\bigcap_{n\geq m} A_n^c\right) \\ &= 0\end{aligned}$$

As required. ■

- *Notions of convergence*
  - ***Definition (convergence almost surely)*** $X_n \to X$ almost surely if

    $$\mathbb{P}\left\{\omega \in \Omega : X_n(\omega) \to X(\omega)\right\} = 1$$

    More generally, $\left\{X_n\right\}$ converges almost surely if

    $$\mathbb{P}\left(\limsup_n X_n = \liminf_n X_n\right) = 1$$

    This is the notion of convergence that most closely maps to convergence concepts in real analysis.

    Let us understand these two statements intuitively

    - The first statement tells us that for *every single outcome* in the sample space Ω, the sequence of random variables X_n tends to X.
    - The second statement is a shorthand for the statement that

      $$\mathbb{P}\left(\left\{\omega : \limsup_n X_n(\omega) = \liminf_n X_n(\omega)\right\}\right) = 1$$

      This is, again, simply the statement in real analysis that

      $$\limsup_n x_n = \liminf_n x_n \text{ if and only if } \lim_n x_n \text{ exists}$$

      Thus, we require the limit to exist for *every outcome* ω.
  - ***Example***: Consider the sequence $X_n = \frac{1}{n} U[0,1]$. We claim that $X_n \to 0$ almost surely.

    Proof: In this case, $\Omega = [0,1]$. For any ω we might drawn, we will find

    $$0 \leq X_n(\omega) \leq \tfrac{1}{n} \to 0$$

    As required. ■ □

this happening decreases with the number of X generated, but it is still possible that in certain experiments, it will happen.

- ***Example***: Let $X_n = \mathbb{I}_{\{U_n \le \frac{1}{n}\}}$, where the U_n are IID $U[0,\ 1]$ random variables. Let us fix $\varepsilon \in (0,1)$, and consider the definition

$$\begin{aligned}\mathbb{P}\left(\left|X_n - 0\right| > \varepsilon\right) &= \mathbb{P}\left(X_n > \varepsilon\right) \\ &= \mathbb{P}\left(U_n \le \tfrac{1}{n}\right) \\ &= \frac{1}{n} \\ &\to 0\end{aligned}$$

This proves the X_n do indeed converge to 0 in probability. Consider, however, that

$$\sum_n \mathbb{P}\left(X_n > \varepsilon\right) = \sum_n \tfrac{1}{n} = \infty$$

But since our X_n are independent, we know by the second BC Lemma that $\left\{X_n > \varepsilon,\ \text{i.o.}\right\}$ almost surely. So X_n cannot be converging to 0 almost surely. Note, on the other hand, that if we replaced our $1/n$ by $1/n^2$, we would find (using the first Borel-Cantelli Lemma) that X_n does indeed converge to 0 almost surely. □

- ***Claim***: $X_n \to X$ a.s. $\Rightarrow X_n \to_p X$ as $n \to \infty$. As our example shows, however, the converse is false.

Proof: Recall that convergence almost surely can be written as

$$\mathbb{P}\left(\left|X_n - X\right| > \varepsilon \quad \text{i.o.}\right) = 0$$

or in other words

$$\mathbb{P}\left(\limsup_n \left\{\left|X_n - X\right| > \varepsilon\right\}\right) = 0$$

We have, however, that $\liminf_n A_n \subseteq \limsup_n A_n$. As such

$$\begin{gathered}\mathbb{P}\left(\liminf_n \left\{\left|X_n - X\right| > \varepsilon\right\}\right) = 0 \\ \mathbb{P}\left(\left|X_n - X\right| > \varepsilon \text{ e.v.}\right) = 0\end{gathered}$$

This naturally implies that *eventually* the probability of large deviations falls to 0 – this is the definition of convergence in probability. ■

Now, pick some k

$$\begin{aligned}\liminf X_n &\geq \lim_{n\to\infty} \mathbb{E}\left(Y_m\right) \\ &\geq \lim_{n\to\infty} \mathbb{E}\left(\min[Y_m, k]\right)\end{aligned}$$

The absolute value of the variable of which we are taking an expectation is now bounded by k, because the X_n are positive, and so Y_m is positive. So we can apply bounded convergence:

$$\begin{aligned}\liminf X_n &\geq \mathbb{E}\left(\lim_{n\to\infty} \min[Y_m, k]\right) \\ &= \mathbb{E}\left(\min[Y, k]\right)\end{aligned}$$

As $k \to \infty$, the last line tends to $\mathbb{E}(Y) = \mathbb{E}(\liminf_n X_n)$. ■

- ***Theorem (DOM – Dominated Convergence)***: Let $\{X_n\}$ be a sequence of random variables such that $|X_n| \leq Y$ with $\mathbb{E}(Y) < \infty$ and $X_n \to X$ almost surely, then $\mathbb{E}(X_n) \to \mathbb{E}(X)$.

  Proof: Since $|X_n| \leq Y$, we have that $Y + X_n \geq 0$ and $Y - X_n \geq 0$. Applying Fatou's Lemma to both variables, we obtain

  $$\liminf_n \mathbb{E}\left(Y + X_n\right) \geq \mathbb{E}\left(Y + X\right) \qquad \liminf_n \mathbb{E}\left(Y - X_n\right) \geq \mathbb{E}\left(Y - X\right)$$

  Since Y has finite expectation, we can subtract $\mathbb{E}(Y)$ from both sides above, and obtain

  $$\begin{aligned}&\liminf_n \mathbb{E}\left(X_n\right) \geq \mathbb{E}\left(X\right) && \liminf_n \mathbb{E}\left(-X_n\right) \geq \mathbb{E}\left(-X\right) \\ &\liminf_n \mathbb{E}\left(X_n\right) \geq \mathbb{E}\left(X\right) && \limsup_n \mathbb{E}\left(X_n\right) \leq \mathbb{E}\left(X\right)\end{aligned}$$

  Together, the last two lines imply that $\lim_n \mathbb{E}(X_n) \to \mathbb{E}(X)$. ■

- ***Theorem (MON – Monotone Convergence)***: Let $\{X_n\}$ be a sequence of random variables such that $0 \leq X_1 \leq X_2 \leq \cdots$ almost surely, then
  - $X_n \nearrow X$ (possibly ∞) almost surely.
  - $\mathbb{E}(X_n) \nearrow \mathbb{E}(X)$ (possibly ∞).

  Proof: Since $X_n \geq 0$, we can apply Fatou's Lemma, to get

  $$\liminf_n \mathbb{E}(X_n) \geq \mathbb{E}(\liminf_n X) = \mathbb{E}(X)$$

  However, the fact that $X_n \leq X$ also gives $\mathbb{E}(X_n) \leq \mathbb{E}(X) \Rightarrow \liminf_n \mathbb{E}(X_n) \leq \mathbb{E}(X)$. These two statements together imply that $\lim_n \mathbb{E}(X_n) \to \mathbb{E}(X)$. ■

This is the characteristic function of $N(0,\sigma^2)$, and it is continuous at $\theta = 0$. Thus, all the conditions of Levy's Theorem hold, which proves the central limit theorem.

All we now need to do is to prove that $nR_n \to 0$ as $n \to \infty$. To do this, first note the following standard result from deterministic analysis proof?!?

$$\left| e^{ix} - \sum_{m=0}^{n} \frac{(ix)^n}{m!} \right| \le \left(\frac{\left|x\right|^{n+1}}{(n+1)!} \wedge \frac{2\left|x\right|^n}{n!} \right)$$

Here, $a \wedge b = \min(a,b)$. Let us apply expectations

$$\mathbb{E}\left| e^{ix} - \sum_{m=0}^{n} \frac{(ix)^n}{m!} \right| \le \mathbb{E}\left(\frac{\left|x\right|^{n+1}}{(n+1)!} \wedge \frac{2\left|x\right|^n}{n!} \right)$$

Now, use Jensen's Inequality on the LHS to obtain

$$\left| \mathbb{E}\left(e^{i\theta X}\right) - \mathbb{E}\left(\sum_{m=0}^{2} \frac{(iX\theta)^m}{m!} \right) \right| \le \mathbb{E}\left(\frac{\left|X\theta\right|^3}{3!} \wedge \frac{2\left|X\theta\right|^2}{2} \right) = \mathbb{E}\left[f(X,\theta)\right]$$

Now, observe that if $\theta \to 0$, $f(X,\theta) \to 0$ almost surely, since $\mathbb{E}\left|X\right|$ is bounded. Similarly, note that $f(X,\theta)$ is bounded by $\left|X\right|^2 \theta^2$, and since $\mathbb{E}\left|X\right|^2 = \sigma^2 < \infty$, we can say that $f(X,\theta) \le Y$ with $\mathbb{E}\left|Y\right| < \infty$. We can therefore apply the dominated convergence theorem to conclude that

$$\mathbb{E}\left[f(X,\theta)\right] \to 0 \qquad \text{as } \theta \to 0$$

More details?

But we have

$$\begin{aligned} R_n &= \mathbb{E}\left(f(X,\tfrac{\theta}{\sqrt{n}})\right) \\ &= n\mathbb{E}\left\{ \frac{\left|X\frac{\theta}{\sqrt{n}}\right|^3}{3!} \wedge \left|X\right|^2 \frac{\theta^2}{n} \right\} \\ &= \mathbb{E}\left\{ \frac{\left|X\right|^3 \frac{\theta^3}{\sqrt{n}}}{3!} \wedge \left|X\right|^2 \frac{\theta^2}{n} \right\} \end{aligned}$$

- ***Definition*** (Stopping time): Suppose T is a non-negative integer-valued random variable. Then T is said to be a *stopping time* with respect to an underlying sequence $\{X_n\}$ if, for each $k \geq 0$,

$$\mathbb{I}_{\{T=k\}} = f_k\left(X_0, \cdots, X_k\right)$$

Where f_k is a deterministic function. In other words, we require

$$\{T = k\} \in \mathcal{F}_k \qquad \forall k$$

- ***Example*** (Hitting Times): Let $T = \inf\{n \geq 0 : X_n \in A\}$. We then have

$$\mathbb{I}_{\{T=k\}} = \mathbb{I}_{\{X_0 \notin A, \cdots, X_{k-1} \notin A, X_k \in A\}}$$

Similarly, we would define $T = \inf\{n \geq 0 : S_n \in A\}$, and we would then have

$$\mathbb{I}_{\{T=k\}} = \mathbb{I}_{\{S_0 \notin A, \cdots, S_{k-1} \notin A, S_k \in A\}}$$

- ***Proposition*** (Wald's First Identity): Let S_n be a random walk $S_n = \sum_{i=1}^n X_i$, with $S_0 = 0$, and let T be a stopping time with respect to the sequence $\{\mathcal{F}_n\}$ (where $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$).
    - If $X_i \geq 0$, then $\mathbb{E}(S_T) = \mathbb{E}(T)\mathbb{E}(X_i)$ [this could, of course, we infinity].
    - If $\mathbb{E}|X_i| < \infty$ and $\mathbb{E}(T) < \infty$, then $\mathbb{E}(S_T) = \mathbb{E}(T)\mathbb{E}(X_0)$

Proof: Let $S_T = \sum_{i=1}^T X_i = \sum_{i=1}^\infty X_i \mathbb{I}_{\{i \leq T\}}$. We then have

$$\mathbb{E}(S_T) = \mathbb{E}\left(\sum_{i=1}^\infty X_i \mathbb{I}_{\{i \leq T\}}\right)$$

Now, let us do both parts:

- ***First part:*** $X_i \geq 0$, and indicators are always positive, so by Fubini I, we can interchange the expectation and the sum:

$$\mathbb{E}(S_T) = \sum_{i=1}^\infty \mathbb{E}\left(X_i \mathbb{I}_{\{i \leq T\}}\right)$$

Consider, however, that

$$\begin{aligned} \mathbb{I}_{\{T \geq i\}} &= 1 - \mathbb{I}_{\{T < i\}} \\ &= 1 - \mathbb{I}_{\{T \leq i-1\}} \end{aligned}$$

This implies that $\mathbb{I}_{\{T \geq i\}} \in \mathcal{F}_{i-1}$. Going back to our sum

- ***Example***: Let $\{X_n\}$ be IID, with $\mathbb{E}(X_i) = 1$ and $\mathbb{E}|X_i| < \infty$. And let $M_n = \prod_{i=1}^{n} X_i$. This is a martingale:
  - Condition 1 satisfied.
  - $\mathbb{E}|M_n| = \prod_{i=1}^{n} \mathbb{E}|X_i| < \infty$
  - Conditioning

$$\begin{aligned}
\mathbb{E}\left(M_n \mid \mathcal{F}_{n-1}\right) &= \mathbb{E}\left(\prod_{i=1}^{n} X_i \mid \mathcal{F}_{n-1}\right) \\
&= \mathbb{E}\left(X_n \prod_{i=1}^{n-1} X_i \mid \mathcal{F}_{n-1}\right) \\
&= M_{n-1}\mathbb{E}\left(X_n \mid \mathcal{F}_{n-1}\right) \\
&= M_{n-1}
\end{aligned}$$

  It is quite astounding, therefore, that even this process can be written as the sum of uncorrelated increments!

- *Optional Stopping Theorem for Martingales*
  - ***Question***: If T is a stopping time, when is it true that $\mathbb{E}(M_T) = \mathbb{E}(M_0)$? This is the question that will be concerning us in this section. Let us consider some simple examples.
  - ***Example***: Let $M_n = S_n - n\mu$, with $M_0 = 0$ and $\mathbb{E}(X_1) = \mu$. When is it the case that $\mathbb{E}(M_T) = \mathbb{E}(M_0) = 0$? Effectively, we are asking when it is the case that

$$\mathbb{E}(S_T) = \mu \mathbb{E}(T)$$

  This is precisely the subject matter of Wald's First Identity. $\square$

  - ***Expample***: Let $M_n = S_n^2 - n\sigma^2$, with $\mathbb{E}(X_1) = 0, \mathbb{E}(X_1^2) = \sigma^2 < \infty$. When is it true that $\mathbb{E}(M_T) = \mathbb{E}(M_0) = 0$? Effectively, we are asking when it is the case that

$$\mathbb{E}(S_T^2) = \sigma^2 \mathbb{E}(T)$$

  This is precisely the subject matter of Wald's Second Identity. $\square$

  - ***Proposition***: Let $\{M_n : n \geq 0\}$ be a martingale with respect to $\{\mathcal{F}_n\}$, and let T be a stopping time. Then for each $m \geq 1$

$$\mathbb{E}\left(M_{T \wedge m}\right) = \mathbb{E}\left(M_0\right)$$

  Where $T \wedge m = \min(T, m)$.

LECTURE 6 – *24th February 2011*

- ***Proof***: Let $\left\{D_i\right\}$ be the martingale differences, and write

$$\begin{aligned} M_{T\wedge n} &= \sum_{i=1}^{T\wedge n} D_i + M_0 \\ &= \sum_{i=1}^{n} D_i \mathbb{I}_{\{T\geq i\}} + M_0 \end{aligned}$$

Take expectations

$$\mathbb{E}(M_{T\wedge n}) = \sum_{i=1}^{n} \mathbb{E}\left(D_i \mathbb{I}_{\{T\geq i\}}\right) + \mathbb{E}\left(M_0\right)$$

We know, however, that $\mathbb{I}_{\{T\geq i\}} \in \mathcal{F}_{i-1}$

$$\begin{aligned} \mathbb{E}(M_{T\wedge n}) &= \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}\left(D_i \mathbb{I}_{\{T\geq i\}} \mid \mathcal{F}_{i-1}\right)\right] + \mathbb{E}\left(M_0\right) \\ &= \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{I}_{\{T\geq i\}} \mathbb{E}\left(D_i \mid \mathcal{F}_{i-1}\right)\right] + \mathbb{E}\left(M_0\right) \\ &= \mathbb{E}\left(M_0\right) \end{aligned}$$

As required. ■

Remark: Consider that

- Consider that $\lim_{n\to\infty} M_{T\wedge n} \to M_T$. As such, $\mathbb{E}\left[\lim_{n\to\infty} M_{T\wedge n}\right] = \mathbb{E}\left[M_T\right]$.
- By the theorem above, however, $\mathbb{E}\left[M_{T\wedge n}\right] = \mathbb{E}\left[M_0\right]$, which implies than $\lim_{n\to\infty} \mathbb{E}\left[M_{T\wedge n}\right] = \mathbb{E}\left[M_0\right]$

Thus, every optional stopping theorem boils down to the following interchange argument – if we can make the interchange, then $\mathbb{E}\left[M_T\right] = \mathbb{E}\left[M_0\right]$:

$$\boxed{\mathbb{E}\left[\lim_{n\to\infty} M_{T\wedge n}\right] = \lim_{n\to\infty} \mathbb{E}\left[M_{T\wedge n}\right]}$$

- ***Corollary I***: Let $\left(M_n : n \geq 0\right)$ be a martingale with respect to $\left\{\mathcal{F}_n\right\}$ and T be a stopping time such that T is bounded (in other words, there exists a $K < \infty$ such that $\mathbb{P}\left(T < K\right) = 1$), then $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.
- ***Corollary II***: Let $\left|M_{T\wedge n}\right| \leq Z$, with $\mathbb{E}(Z) < \infty$. Then by dominated convergence, the interchange holds and $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.
- ***Corollary III***: Let $\left(M_n : n \geq 0\right)$ be a martingale with respect to $\left\{\mathcal{F}_n\right\}$ and T be a stopping time such that $\mathbb{E}(T) < \infty$. Provided the martingale differences are uniformly bounded ($\mathbb{E}\left[\left|D_i\right| \mid \mathcal{F}_{i-1}\right] \leq C < \infty$), then $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.

Proof: Let $M_n = \boldsymbol{f}(X_n)$ for an $\boldsymbol{f}$ that satisfies the assumption of the proposition. The main question here is one of uniqueness, since it is clear that a constant vectors can solve $P\boldsymbol{f} = \boldsymbol{f}$. Let us first show that M_n is a martingale:

- $\mathbb{E}\left|M_n\right| = \mathbb{E}\left|\boldsymbol{f}(X_n)\right| \leq c$
- $M_n \in \mathcal{F}_n = \sigma(X_1, \cdots, X_n)$ (in fact, it only depends on the last X_n).
- $\mathbb{E}\left[M_n \mid \mathcal{F}_{n-1}\right] = \mathbb{E}\left[\boldsymbol{f}(X_n) \mid X_{n-1}\right] = P\boldsymbol{f}(X_{n-1}) = \boldsymbol{f}(X_{n-1})$

As such, $\left\{M_n : n \geq 0\right\}$ is a martingale with respect to $\left\{\mathcal{F}_n\right\}$. Note that $\sup_{n\geq 1} \mathbb{E}\left|M_n\right| < \infty$, because we have assumed that $\boldsymbol{f}$ is bounded. Thus, $M_n \to M_\infty$ almost surely. Suppose there exists $x, y \in \mathcal{S}$ such that

$$\boldsymbol{f}(x) < \boldsymbol{f}(y)$$

Since $\left\{X_n : n \geq 0\right\}$ is irreducible and recurrent

$$\begin{aligned} &\liminf_n f(X_n) \leq \boldsymbol{f}(x) && \text{a.s.} \\ &\limsup_n f(X_n) \geq \boldsymbol{f}(y) && \text{a.s.} \end{aligned}$$

This means, however, that $\liminf_n \boldsymbol{f}(X_n) \neq \limsup_n \boldsymbol{f}(X_n)$, which contradicts the convergence statement. □

Lecture 7 – *2nd March 2011*

- ***Stochastic stability***[4]
  - ***Deterministic motivation***: Consider a dynamical system $X(t)$ for which $\mathrm{d}X(t) = f(X(t))\,\mathrm{d}t$, with $f : \mathbb{R} \to \mathbb{R}$ (this can be thought of as the "equation of motion" of the system).

    Now, consider an "energy" function $g : \mathbb{R} \to \mathbb{R}_+$, with

$$\frac{\mathrm{d}g\left(X(t)\right)}{\mathrm{d}t} \leq -\varepsilon \qquad \varepsilon > 0$$

[4] Some parts of this topic were covered at the end of the previous lecture. For expositional purposes, we chose to present the material here instead.

This is effectively a statement of the fact the energy of the system is "forced" down to 0, since it is "constantly decreasing" As such, $\exists t_0(x,\varepsilon)$ such that $g\left(X(t)\right)=0$ for all $t \geq t_0$; in other words, our dynamical process is "pushed" towards a "stable" state.

- We now need to adapt this idea to a discrete stochastic process. We can write our "equation of motion" as $X_{n+1} - X_n = f(X_n)$. To add stochasticity, we can write $X_{n+1} = \psi(X_n, \varepsilon_{n+1})$, where the ε_i are random variables. This is effectively the definition of a Markov process, provided ε_n is independent of X_0. If we make the ε_i IID, the Markov process becomes time-homogeneous.

  Now consider the "energy" function – it would be too strong to ask for the energy of the process to decrease along *every* path. We therefore require it to decrease in expectation:

$$\mathbb{E}\left[g(X_{n+1}) - g(X_n) \mid X_n\right] \leq -\varepsilon$$

  Since we need this be true *whatever state* our process first starts in, we can write $\mathbb{E}_x[\cdot] = \mathbb{E}[\cdot \mid X_0 = x]$, and the condition above becomes

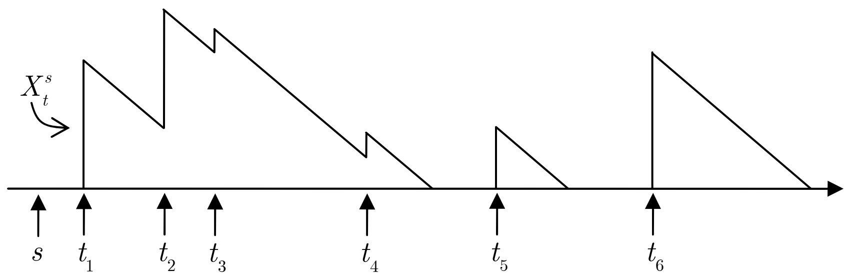
$$\mathbb{E}_x\left[g(X_1)\right] - g(x) \leq -\varepsilon$$

- We now specialize this to a particular stochastic process. Consider a Markov chain $\left\{X_n : n \geq 0\right\}$ that is irreducible. We would like to know whether the chain has a steady state. For a finite state space, all we need is to check for solutions to $\boldsymbol{\pi}^\top P = \boldsymbol{\pi}, \boldsymbol{\pi}^\top \boldsymbol{1} = 1$; indeed, the existence of such a $\boldsymbol{\pi}$ is associated with positive recurrence. If the spate space is countably infinity, however, things get slightly more complicated; the two equations above do not suffice, and we also require $\boldsymbol{\pi} \geq \boldsymbol{0}$, which makes things more complicated. Here, we attempt to find simpler conditions for stability to hold.

- ***Proposition***: Let $\left\{X_n : n \geq 0\right\}$ be an irreducible markov chain on a countable state space $\mathcal{S}^k$. Let $K \subseteq \mathcal{S}$ be a set containing a finite number of states. Then, if there exists a function $g : \mathcal{S} \to \mathbb{R}_+$ such that (recall $\mathbb{E}_x[\cdot] = \mathbb{E}[\cdot \mid X_0 = x]$)

$$\mathbb{E}_x g(X_1) - g(x) \leq -\varepsilon \quad \forall x \in \bar{\mathcal{K}} \text{ and some } \varepsilon{>}0$$

- Set $s < t$ and $\omega \in \Omega$. Define $X_t^s(\omega)$ to be the work in the system at time t given the system is empty at time s. For the point ω depicted above (and assuming $s < t_1$), X_t^s would look like this



- For $s' < s$, it is clear that

$$X_t^{s'}(\omega) \geq X_t^s(\omega)$$

Because this variable is non-decreasing in s, $X_t^*(\omega) = \lim_{s\to-\infty} X_t^s(\omega)$ exists. This is the "steady state" of the system; what we would observe if we had started the system a very, very long time in the past.

- Now, define a time-shift operator θ_τ such that $X_t^s(\theta_\tau \omega) = X_{t+\tau}^{s+\tau}(\omega)$. As such, we have

$$\lim_{s\to-\infty} X_t^s(\theta_\tau \omega) = \lim_{s\to-\infty} X_{t+\tau}^{s+\tau}(\omega) = X_{t+\tau}^*(\omega) = X_t^*(\theta_\tau \omega)$$

As such, $X_t^*(\omega) =_d X_t^*(\theta_\tau \omega)$. Shift invariance also holds for the starred process.

Lecture 8 – *23rd March 2011*

- We have seen a number of properties of X_t^*, including the fact that it exists. It might, however, be equal to infinity. We now look for conditions under which X_t^* is finite.

- ***Proposition***: If $\rho < 1$, then $X_t^* < \infty$ a.s. for all t.

Proof: Fix $t \in \mathbb{R}$ and $\omega \in \Omega$ and define

$$T_t^s(\omega) = \sup\left\{\tau < t : X_\tau^s(\omega) = 0\right\}$$

Intuitively, this is the last empty time before time t (this is clearly not a stopping time; just a random time):

Intuitively, this states that the average amount of work in the system at any given time is equal to the average number of arrivals per unit time multiplied by the average sojourn time.

Proof:

$$\int_0^t N(s)\ \mathrm{d}s = \int_0^t A(s) - D(s)\ \mathrm{d}s$$

Note that we can write

$$\begin{aligned} A(s) &= \sum\nolimits_{n\in\mathbb{Z}} \mathbb{1}_{\{t_n \le s\}} \\ D(s) &= \sum\nolimits_{n\in\mathbb{Z}} \mathbb{1}_{\{d_n \le s\}} \end{aligned}$$

As such

$$N(s) = \sum\nolimits_{n\in\mathbb{Z}} \mathbb{1}_{\{t_n \le s\}} - \mathbb{1}_{\{d_n \le s\}} = \sum\nolimits_{n\in\mathbb{Z}} \mathbb{1}_{\{t_n \le s \le d_n\}}$$

(The last equality follows because any jobs that arrive after s won't be counted at all, and any events that arrive and leave before s will be counted by both indicators and therefore cancel out). By swapping the summation and integration (valid by Fubini), we obtain

$$\begin{aligned} \int_0^t N(s)\ \mathrm{d}s &= \sum\nolimits_{n\in\mathbb{Z}} \int_0^t \mathbb{1}_{\{t_n \le s \le d_n\}}\ \mathrm{d}s \\ &= \sum\nolimits_{n\in\mathbb{Z}} \left\{ \begin{matrix} \text{Amount of time job } n \text{ was} \\ \text{in the system during } [0,t] \end{matrix} \right\} \end{aligned}$$

We can bound this above by considering the sojourn time of all arrivals up to and including time t (though some of them may overrun past t) and lower bound it by considering the sojourn time of all job that depart before time t (even though some jobs that leave after t do spend *some* time in the system before t). This gives

$$\sum\nolimits_{i=0}^{D(t)} \theta_{n_i} \le \int_0^t N(s)\ \mathrm{d}s \le \sum\nolimits_{n=0}^{A(t)} \theta_n$$

Where n_i is the index of the i^{th} job to leave the system (since we have no assumed FIFO processing discipline, we cannot assume that $\theta_i = i$).

Before we continue, we will need the following claim

Claim: Under the two assumptions above for in this case, because M_n has some structure (the fact it's non-decreasing) that w_n does not. However, since this is a Markov chain, a stationary distribution is all we could really want].

- If the random walk has positive drift – in other words, if

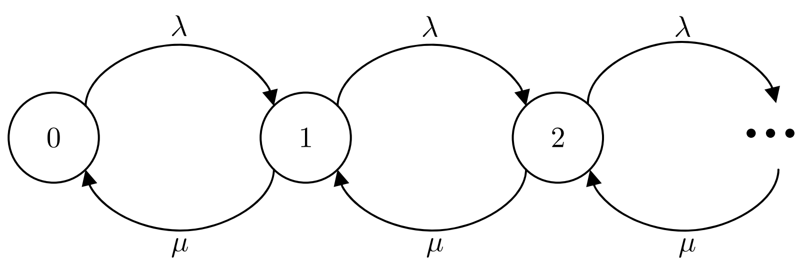
$$\mathbb{E}(Z_n) = \mathbb{E}(S_{n-1}) - \mathbb{E}(\tau_n) > 0$$

– then the random walk drifts to infinity and the waiting times get infinitely large. This is consistent with our findings in the $G/G/1$ queue, since $\mathbb{E}(\tau_{n+1}) > \mathbb{E}(S_n) \Leftrightarrow \rho < 1$. On the other hand, if the random walk has negative drift, the chain is stable and the waiting times return to 0 infinitely often almost surely (we motivated this result in homework 2 using a simpler reflected random walk).

- ***The single-server $M/M/1$ Markovian queue***
  - We now consider the most tractable of all single-server queue models; the $M/M/1$ in queue. In that case, we assume that $\{S_n\}$ are IID and exponentially distributed with parameter μ, whereas the $\{\tau_n\}$ are IID and exponentially distributed with parameter λ.
  - Consider the process

$$X(t) = \text{Number of jobs in system at time } t \geq 0$$

$\{X(t) : t \geq 0\}$ is a continuous-time Markov Chain with countable state space. In fact, it is a birth-and-death process:



We now proceed to analyze this CTMC (note that $f(h) = o(h) \Rightarrow \lim_{h\to\infty} \frac{f(h)}{h} = 0$)

Consider $t > 0$ and a small $h > 0$. We have is continuous, every up-crossing must be followed by a down-crossing, and so the difference in numbers will be at most 1.

2. Using the ergodic theorem of Markov Chains, we have

$$\lim_{t\to\infty} \frac{D_j(t)}{D(t)} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{X_k=j\}} = \pi(j) \quad \text{a.s.}$$

(The first equality follows from the fact that X_k is precisely the state of the system after the k^{th} departure).

3. Using the bound in (1), we have that $D_j(t) - 1 \le A_j(t) \le D_j(t) + 1$, so

$$\overbrace{\frac{D(t)}{A(t)}}^{\to 1} \overbrace{\frac{D_j(t)-1}{D(t)}}^{\to \pi(j) \text{ by (2)}} \le \frac{A_j(t)}{A(t)} \le \overbrace{\frac{D_j(t)+1}{D(t)}}^{\to \pi(j) \text{ by (2)}} \cdot \overbrace{\frac{D(t)}{A(t)}}^{\to 1}$$

And so

$$\frac{A_j(t)}{A(t)} \to \pi(j) \text{ a.s}$$

4. Letting t_k be the time of the k^{th} arrival, we have

$$\lim_{t\to\infty} \frac{A_j(t)}{A(t)} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{X(t_k)=j\}}$$

5. A well known property of queues with Poisson arrivals is PASTA (Poisson arrivals see time averages – see addendum at the end of this lecture).

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{1}_{\{X(s)=j\}}\, \mathrm{d}s = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{X(t_k)=j\}}$$

where t_k is the time of the k^{th} arrival to the system. Effectively, this states that in working out the average work in the queue, we don't need to sample at *every* time-step; it is enough to sample at arrivals.

6. By the ergodic theorem for Markov Chains

$$\hat{\pi}(j) =_{\text{a.s}} \lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{1}_{\{X(s)=j\}}\, \mathrm{d}s$$

7. Finally, combine all the above

Where the last step follows by Fubini, since h is bounded. Now, since h is *any* bounded function, the above implies that $X(t) \Rightarrow X(\infty)$, and finally, taking $h_x(\cdot) = \mathbb{1}_{\{\cdot = x\}}$, we recover the last required statement.

- ***Example***: Let $\left(X(t) : t \geq 0\right)$ be a positive recurrent regenerative process. Fix a set $A \in S$, and let

$$T(A) = \inf\left\{t \geq 0 : X(t) \in A\right\}$$

Again, for simplicity, assume $\tau(0) = 0$. We are now interested in the expression $\mathbb{P}\left(T(A) > t\right) = ?$, especially for t large. For $t > 0$, note that

$$\begin{aligned} a(t) &= \mathbb{P}\left(T(A) > t\right) \\ &= \mathbb{P}\left(T(A) > t, \tau_1 > t\right) + \mathbb{P}\left(T(A) > t, \tau_1 \leq t\right) \\ &= \mathbb{P}\left(\left[T(A) \wedge \tau_1\right] > t\right) + \int_0^t \mathbb{P}\left(T(A) > t, \tau_1 = s\right) \, \mathrm{d}s \\ &= \mathbb{P}\left(\left[T(A) \wedge \tau_1\right] > t\right) + \int_0^t \mathbb{P}\left(T(A) > t, T(A) > \tau_1 \mid \tau_1 = s\right) \, \mathrm{dF}s \\ &= \mathbb{P}\left(\left[T(A) \wedge \tau_1\right] > t\right) + \int_0^t \mathbb{P}\left(\tilde{T}(A) > t - s, T(A) > \tau_1 \mid \tau_1 = s\right) \, \mathrm{dF}s \end{aligned}$$

Where $\tilde{T}(A) = \inf\left\{t > \tau(1), X(t) \in A\right\} = \inf\left\{t \geq 0, X(t + \tau(1)) \in A\right\}$. Now

$$\begin{aligned} a(t) &= \mathbb{P}\left(\left[T(A) \wedge \tau_1\right] > t\right) + \int_0^t \mathbb{P}\left(T(A) > t - s\right)\mathbb{P}\left(T(A) > \tau_1 \mid \tau_1 = s\right) \, \mathrm{dF}s \\ &= \mathbb{P}\left(\left[T(A) \wedge \tau_1\right] > t\right) \\ &\qquad + \int_0^t \mathbb{P}\left(T(A) > t - s\right)\mathbb{P}\left(\tau_1 = s \mid T(A) > \tau_1\right)\mathbb{P}\left(T(A) > \tau_1\right) \, \mathrm{dF}s \\ &= b(t) + \int_0^t a(t-s)\beta \, \mathrm{d}\tilde{\mathrm{F}}s \end{aligned}$$

Where

$$\tilde{F}(s) = \mathbb{P}\left(\tau_1 \leq s \mid T(A) > \tau_1\right)$$

So

$$a = b + \beta\left(a * F\right)$$

For fixed $\lambda \in \mathbb{R}$,

$$\begin{aligned} a(t)e^{\lambda t} &= b(t)e^{\lambda t} + \beta e^{\lambda t} \int_0^t a(t-s) \, \mathrm{d}\tilde{F}(s) \\ a_\lambda(t) &= b_\lambda(t) + \int_0^t \beta e^{\lambda(t-s)} a(t-s) \beta e^{\lambda s} \, \mathrm{d}\tilde{F}(s) \\ a_\lambda(t) &= b_\lambda(t) + \int_0^t a_\lambda(t-s) \, \mathrm{d}F_\lambda(s) \qquad \mathrm{d}F_\lambda(s) = \beta e^{\lambda s} \mathrm{d}\tilde{F}(s) \end{aligned}$$

Suppose $\exists \lambda$ s.t. F_λ is a bona-fide distribution. Then

$$F_\lambda(\infty) = 1 = \beta \int_0^\infty e^{\lambda s} \, \mathrm{d}\tilde{F}(s)$$

$$\beta^{-1} = \int_0^\infty e^{\lambda s} \, \mathrm{d}\tilde{F}(s) = \tilde{\mathbb{E}}\left(e^{\lambda Z}\right)$$

Assume there is such a solution λ^* [can show using hyeristic argument]. By appealing to the key renewal theorem

$$a_\lambda(t) \to \frac{1}{\mathbb{E}_\lambda\left(\tau_1\right)} \int_0^\infty e^{\lambda s} \mathbb{P}\left(\left[T(A) \wedge \tau_1\right] > s\right) \, \mathrm{d}s = \eta \qquad t \to \infty$$

$$\mathbb{E}_\lambda \tau_1 = \int_0^\infty x \, \mathrm{d}F_\lambda(x)$$

And so

$$a(t) = \mathbb{P}\left\{T(A) > t\right\} \sim \eta e^{-\lambda t}$$

So the probability has an exponential type tail. □