DISCRETE STOCHASTIC PROCESSES

Lecture 25

Review: Martingales and Stopped Martingales
Submartingales and Supermartingales
Jensen's Inequality
Kolmogorov Submartingale Inequality

Corollary 1: Kolmogorov Martingale Inequality

Corollary 2: Kolmogorov's Random Walk Inequality

Strong Law of Large Numbers

Martingale Convergence Theorem

MARTINGALES

Definition: A **martingale** $\{Z_n; n \geq 1\}$ is a stochastic process with the properties that $E[|Z_n|] < \infty$ for all n and

$$\boxed{E[Z_n \mid Z_{n-1} = z_{n-1}, Z_{n-2} = z_{n-2}, \ldots, Z_1 = z_1] = z_{n-1}}$$

for all $n > 1$ and all $z_1, z_2, \ldots, z_{n-1}$.

Definition: A stochastic process $\{Z_n; n \geq 1\}$ is a **submartingale** if

$E[|Z_n|] < \infty$ for all $n > 1$ and if $E\left[Z_n \middle| Z_{n-1}, Z_{n-2}, \ldots, Z_1\right] \geq Z_{n-1}$.

<u>**Examples**</u> If $\{S_n\}$ is a random walk, $\{(S_n - n\overline{X})\}$ is a martingale. IID multiplicative gambling process with $\overline{X} = 1$ is a martingale, even if X_n is no longer nonnegative. A branching process with $\overline{Y}$ = expected number of offspring =1 is a martingale, and for a branching process with $\overline{Y}$ = m, X_n / m^n is a martingale.

Jensen's Inequality: If $h(x)$ is convex and if X is a rv with $E[X] < \infty$, then

$$h(E[X]) \le E[h(X)].$$

Theorem: If $\{Z_n\,;\, n \ge 1\}$ is a martingale and h is convex, then $\{h(Z_n);\, n \ge 1\}$ is a submartingale.

Examples: If $\{Z_n\,;\, n \ge 1\}$ is a martingale, then $\{|Z_n|;\, n \ge 1\}$ is a submartingale,

$\{Z_n^2\,;\, n \ge 1\}$ is a submartingale, and $\{\exp(rZ_n);\, n \ge 1\}$ is a submartingale.

Similarly, $\{\ln(Z_n), n \ge 1\}$ is a supermartingale if Z_n is a positive martingale.

STOPPING RULE AND STOPPED PROCESSES

N is a **possibly defective stopping rule** for a process $\{Z_n;\, n \geq 1\}$ if event $\{N \geq i\}$ is determined by $\{Z_n;\, n < i\}$. N is a **stopping rule** if also a rv $\left(\lim_{i\to\infty} P(N \geq i) = 0\right)$.

The **stopped process** $\{Z_n^*;\, n \geq 1\}$ for a process $\{Z_n;\, n \geq 1\}$ with a possibly defective stopping rule is given by $Z_n^* = Z_n$ for $n \leq N$, $Z_n^* = Z_N$ for $n > N$.

Z_n^* is a rv for all $n \geq 1$. **After the observations stop, Z_n^* remains constant.**

Theorem: If $\{Z_n; n \geq 1\}$ is a martingale and N is a possibly defective stopping rule, then the stopped process $\{Z_n^*; n \geq 1\}$ is a martingale. $E[Z_n^*] = E[Z_1]$

If $\{Z_n; n \geq 1\}$ is a submartingale (supermartingale) and N is a possibly defective stopping rule, then the stopped process $\{Z_n^*; n \geq 1\}$ is a submartingale (supermartingale). $E[Z_n^*] \geq E[Z_1]$ $(E[Z_n^*] \leq E[Z_1])$

We have that $\lim_{n\to\infty} Z_n^* = Z_N \; WP1$ for a nondefective stopping rule N. For martingale with stopping rule, it is important to understand relation of $E[Z_N]$ to $\lim_{n\to\infty} E[Z_n^*]$, since
$\lim_{n\to\infty} Z_n^* = Z_N$ wp1 $\not\Rightarrow$ $\lim_{n\to\infty} E\{Z_n^*\} = E\{Z_N\}$.

For example, in the binary product martingale where we stop the first time we get a 0, $Z_n = Z_n^*$, $P(Z_n = 2^n) = 2^{-n}$ **and** $P(Z_n = 0) = 1 - 2^{-n}$**, so** $E[Z_n] = E[Z_n^*] = 1$ for all n **but** $Z_N = 0$ **WP1, so** $E[Z_N] = 0$.

$$0 = E\{Z_N\} \neq E\{Z_n\} = E\{Z_n^*\} = E\{Z_1\} = \lim_{n\to\infty} E\{Z_n\} = \lim_{n\to\infty} E\{Z_n^*\} = 1.$$

Properties

Martingale

$$E\left\{Z_n \left| Z_{n-1},\ldots,Z_1\right.\right\} = Z_{n-1}$$

Submartingale

$$E\left\{Z_n \left| Z_{n-1},\ldots,Z_1\right.\right\} \geq Z_{n-1}$$

Supermartingale

$$E\left\{Z_n \left| Z_{n-1},\ldots,Z_1\right.\right\} \leq Z_{n-1}$$

For n > i ≥ 1

$$E\left\{Z_n \left| Z_i,\ldots,Z_1\right.\right\} = Z_i \qquad E\left\{Z_n \left| Z_i,\ldots,Z_1\right.\right\} \geq Z_i \qquad E\left\{Z_n \left| Z_i,\ldots,Z_1\right.\right\} \leq Z_i$$

$$E\left\{Z_n\right\} = E\left\{Z_i\right\} \qquad E\left\{Z_n\right\} \geq E\left\{Z_i\right\} \qquad E\left\{Z_n\right\} \leq E\left\{Z_i\right\}$$

$$E\left\{Z_n\right\} = E\left\{Z_1\right\} \qquad E\left\{Z_n\right\} \geq E\left\{Z_1\right\} \qquad E\left\{Z_n\right\} \leq E\left\{Z_1\right\}$$

For any possibly defective stopping rule N

Z_n^* is a martingale

Z_n^* is a submartingale

Z_n^* is a supermartingale

$$E\left\{Z_1\right\} = E\left\{Z_n^*\right\} = E\left\{Z_n\right\} \qquad E\left\{Z_1\right\} \leq E\left\{Z_n^*\right\} \leq E\left\{Z_n\right\} \qquad E\left\{Z_1\right\} \geq E\left\{Z_n^*\right\} \geq E\left\{Z_n\right\}$$

More Delicate Issue

For any martingale $\{Z_n, n \geq 1\}$ and any stopping rule N,

$$\lim_{n\to\infty} Z_n^* = Z_N \text{ with probability 1.}$$

And for any n ≥ 1,

$$E\left\{Z_1\right\} = E\{Z_n^*\} = \sum_{i=1}^{n} E\{Z_n^* \mid N = i\}P\{N = i\} \; + \; E\{Z_n^* \mid N > n\}P\{N > n\} =$$

$$\underbrace{\sum_{i=1}^{n} E\{Z_N \mid N = i\}P\{N = i\}}_{\underset{n\to\infty}{\rightarrow} E\{Z_N\} \text{ if } E\{|Z_N|\} < \infty} + \; E\{Z_n \mid N > n\}P\{N > n\}$$

Theorem: Let N be a stopping rule for the martingale $\{Z_n,\; n \geq 1\}$
and assume $E\{| Z_N |\} < \infty.$ Then $E\{Z_N\} = E\{Z_1\} \quad \Leftrightarrow$

$$\lim_{n\to\infty} E\{Z_n \mid N > n\}P\{N > n\} = 0.$$

KOLMOGOROV SUBMARTINGALE INEQUALITY

For **non-negative submartingale** $\{Z_n;\, n \geq 1\}, m \geq 1, a > 0$,

$$\boxed{P\left(\max_{1\leq n\leq m} Z_n \geq a\right) \leq \frac{E[Z_m]}{a}}$$

This is like the Markov inequality, but **more powerful** because of the max over Z_n. If Z_n is a martingale, $E[Z_m] = E[Z_1]$, yielding the **even more powerful** (Version I of the) Kolmogorov martingale inequality

$$P\left(\max_{1\leq n\leq m} Z_n \geq a\right) \leq \frac{E[Z_1]}{a},\ \forall m,\ i.e.,$$

$$P\left(\max_{n} Z_n \geq a\right) \leq \frac{E[Z_1]}{a}$$

Proof: Let N be the following stopping rule: for $n < m$, $N = n$ if $Z_n \geq a$, and $Z_i < a$ for $i < n$. $N = m$ if $Z_i < a$ for $i < m$. Then $\max_{i\leq m} Z_i \geq a$ iff $Z_N \geq a$. But $Z_N = Z_m^*$, so by Markov,

$$P\left(\max_{1\leq n\leq m} Z_n \geq a\right) = P\left(Z_m^* \geq a\right) \leq \frac{E\left[Z_m^*\right]}{a}$$

But $E\left[Z_m^*\right] \leq E[Z_m]$, completing the submartingale proof. (Only this last step uses the fact that $\{Z_n\}$ is a submartingale.) For a martingale, $E[Z_m] = E[Z_1]$.

Closer Examination of Kolmogorov Submartingale Inequality and Its Proof

Result is strongest when $\{Z_n, n \geq 1\}$ is a nonnegative martingale. Then $E\{Z_m\} = E\{Z_1\}$ and theorem says

$$P\left\{\max_{1 \leq n \leq m} Z_n \geq a\right\} \leq \frac{E\{Z_1\}}{a},$$

where the rhs bound is independent of m. Thus

$$P\left\{\sup_{n \geq 1} Z_n > a\right\} \leq \frac{E\{Z_1\}}{a},$$

Compare this with the random process $\{Y_n, n \geq 1\}$, where $\{Y_n\}$ is a set of nonnegative iid random variables with $E\{Y_n\} = E\{Y_1\} < \infty$.

- Is $\{Y_n, n \geq 1\}$ a nonnegative martingale?
- Does the Kolmogorov inequality hold for any such process $\{Y_n\}$?
- If not, precisely where does the proof fail?

- $\{Y_n\}$ is not a martingale in general, since $E\{Y_n | Y_{n-1},\ldots,Y_1\} = E\{Y_1\} \neq Y_{n-1}$.
- The Kolmogorov inequality generally does not hold for $\{Y_n\}$. Suppose for some $a > E\{Y\}$, $P\{Y \geq a\} = \varepsilon > 0$. Then

$$P\left\{\max_{1\leq n\leq m} Y_n < a\right\} = (1-\varepsilon)^m$$

$$P\left\{\max_{1\leq n\leq m} Y_n \geq a\right\} = 1-(1-\varepsilon)^m \underset{m\to\infty}{\longrightarrow} 1 \not\leq \frac{E\{Y_1\}}{a}$$

- The proof fails for $\{Y_n\}$ only in the last step, which asserts $E\{Z_m^*\} \leq E\{Z_m\}$ or, in the special case of a martingale $E\{Z_m^*\} = E\{Z_m\} = E\{Z_1\}$, since a stopped martingale is a martingale. For $\{Y_n\}$, if $a > E\{Y\}$ and m sufficiently large,

$$E\{Y_m^*\} > E\{Y_m\} = E\{Y_1\}$$

and the proof fails: unlike a martingale, stopping $\{Y_n\}$ (with $a > \bar{Y}$) can raise the mean.

Using the example of the nonnegative iid sequence Y_n with $P(Y \geq a > E[Y]) = \varepsilon > 0$,

$$E\left\{Y_m^*\right\} = \sum_{k=1}^{m-1} E\left\{Y_m^* \middle| N=k\right\} P\left\{N=k\right\} + E\left\{Y_m^* \middle| N=m\right\} P\left\{N=m\right\} >$$

$$\sum_{k=1}^{m-1} E\left\{Y_m^* \middle| N=k\right\} P\left\{N=k\right\} > \sum_{k=1}^{m-1} a\left(1-\varepsilon\right)^{(k-1)} \varepsilon \underset{m\to\infty}{\longrightarrow} a,$$

i.e., you can improve your expected return with $\left\{Y_n\right\}$ by waiting for an unusually good outcome and stopping then.

This emphasizes a key feature buried in the definition of martingales. Unlike $\left\{Y_n\right\}$, a stopped martingale still has $E\left\{Z_m^*\right\} = E\left\{Z_m\right\} = E\left\{Z_1^* = Z_1\right\}$. **Stopping** a martingale **cannot** affect its expectation.

Hopefully this adds some understanding of the significance of the

Theorem: A stopped martingale remains a martingale

Kolmogorov Martingale Inequality

For a **martingale** $\{Z_n ; n \geq 1\}$ **with** $E\left[Z_n^2\right] < \infty$ for all n, $\left\{Z_n^2 ; n \geq 1\right\}$ is a nonnegative submartingale (by Jensen's theorem). Therefore

$$P\left(\max_{1 \leq n \leq m} |Z_n|^2 \geq b^2\right) \leq \frac{E\left[Z_m^2\right]}{b^2}, \text{ i.e.,}$$

$$\boxed{P\left(\max_{1 \leq n \leq m} |Z_n| \geq b\right) \leq \frac{E\left[Z_m^2\right]}{b^2}}$$

This has the same relationship to the submartingale inequality as the Chebyshev inequality does to the Markov inequality.

Kolmogorov Random Walk Inequality

For a **random walk** $S_n = X_1 + \cdots + X_n$, with X's iid with variance σ_X^2 ,
Sn – nE[X] is a martingale with variance $n\sigma_X^2$. Applying the martingale inequality yields

$$P(\max_{1 \le n \le m} |S_n - n\overline{X}| \ \ge\ b) \ \le\ \frac{\mathrm{E}[|S_m - m\overline{X}|^2]}{b^2} \ =\ \frac{m\sigma_X^2}{b^2}$$

and, in particular, for $\varepsilon > 0$

$$P(\max_{1 \le n \le m} |S_n - n\overline{X}| \ \ge\ \mathrm{m}\varepsilon) \ \le\ \frac{\sigma_X^2}{m\varepsilon^2}$$

STRONG LAW OF LARGE NUMBERS (ASSUMING FINITE VARIANCE)

Let $\{X_i; i \geq 1\}$ be IID with mean $\overline{X}$ and variance σ^2. Then for any $\varepsilon > 0$,

$$\lim_{m\to\infty} P\left(\sup_{n\geq m} \left| \frac{S_n}{n} - \overline{X} \right| \geq \varepsilon \right) = 0$$

This is the same as the weak law of large numbers except for the sup.

The weak law says that $S_m/\ m$ is close to $\overline{X}$ with high probability, whereas the strong law says that *the infinite sequence* $S_m/m,\ S_{m+1}/(m+1),\ldots$ *stay*s close to $\overline{X}$ with high probability. The sequence does not have any rare large excursions from the mean once m is sufficiently large.

Strategy for Proof

The strong law,

$$\lim_{m\to\infty} P\left(\sup_{n\geq m} \left| \frac{S_n}{n} - \overline{X} \right| \geq \varepsilon \right) = 0,$$

follows from the random walk inequality,

$$P\left(\max_{1\leq n\leq m} \left| S_n - n\overline{X} \right| \geq m\varepsilon \right) = P\left(\max_{1\leq n\leq m} \left| \frac{S_n}{n} - \overline{X} \right| \geq \frac{m\varepsilon}{n} \right) \leq \frac{\mathrm{Var}(X)}{m\varepsilon^2}.$$

While one can perhaps imagine some vague relationship between the two, it is far from clear how we could ever derive the first from the second. **First** of all, the random walk inequality applies only to the first m terms of the series, while the strong law applies to the *infinite* set of terms beginning with the m-th term. **Second,** a straightforward approach of letting m $\to\infty$ in the random walk inequality won't work because the deviation bound, $m\varepsilon \;\; (or \;\; m\varepsilon/n)$, goes to infinity if we do that.

A **third** and slightly more subtle approach of choosing ε in the random walk inequality equal to δ/m keeps the deviation $m\varepsilon \;\; (or \;\; m\varepsilon/n)$ bounded, but at the unacceptable cost that the probability bound $(Var(X)/m\varepsilon^2)$ becomes $(mVar(X)/\delta^2)$, which again goes to infinity as m Increases.

A **fourth** idea, which turns out to be of some use, is to break up the event

$$E_m = \{\sup_{n \geq m} \left| \frac{S_n}{n} - \overline{X} \right| > \varepsilon\} = \bigcup_{n=m}^{\infty} \{\left| \frac{S_n}{n} - \overline{X} \right| > \varepsilon\},$$

relevant to the strong law, into the countable union of events of the form

$$E^k = \{\left| \frac{S_n}{n} - \overline{X} \right| > \varepsilon, L^k \leq n < U^k\},$$

(where the upper and lower bounds U^i and L^k will be chosen later,)

in such a way that ,

$$\{n \mid n \geq m\} = \bigcup_{i=i_0(m)}^{\infty} \{n \mid L^i \leq n < U^i\}, and \quad therefore \;\; E_m \subseteq \bigcup_{i=i_0(m)}^{\infty} E^i$$

which implies (from the union bound)

$$P\{E_m\} \leq \sum_{i=i_0(m)}^{\infty} P\{E^i\}$$

Sufficiently tight bounds on $P\{E^i\}$ will therefore suffice to prove the strong law.

Unfortunately the most straightforward ways to bound

$P\{E^i\} = P\{\max_{L^i \le n \le U^i} \left| \frac{S_n}{n} - \overline{X} \right| \ge a\}$ can also go astray. For example,

$$P\{\max_{L \le n \le U} \left| \frac{S_n}{n} - \overline{X} \right| \ge a\} = P\{\max_{L \le n \le U} \left| S_n - n\overline{X} \right| \ge na\} \le P\{\max_{1 \le n \le U} \left| S_n - n\overline{X} \right| \ge 1a\} \le$$

(by the random walk inequality $P(\max_{1 \le n \le m} | S_n - n\overline{X} | \ \ge \ b) \ \le \ \frac{m\sigma_X^2}{b^2}$

$$E\{\left| S_U - U\overline{X} \right|^2\} / a^2 = U\delta_X^2 / a^2$$

which sadly goes to infinity as U becomes large. The following proof succeeds by cleverly choosing L^i and U^i and replacing the first inequality immediately above by a less conservative one that substitutes La for 1a.   .

Proof of Strong Law

Since $P\left(\sup_{n\geq m}\left|\frac{S_n}{n}-\bar{X}\right|>\varepsilon\right)$ is non-negative and non-increasing in m, it necessarily

converges to a limit. We can find the limit by considering only the subsequence

where $m=2^k$:

$$P\left(\sup_{n\geq 2^k}\left|\frac{S_n}{n}-\bar{X}\right|>\varepsilon\right)=P\left\{\text{for at least one } j\geq k,\ \max_{2^j\leq n<2^{j+1}}\left|\frac{S_n}{n}-\bar{X}\right|>\varepsilon\right\}=$$

$$P\left(\bigcup_{j=k}^{j=\infty}\max_{2^j\leq n<2^{j+1}}\left|\frac{S_n}{n}-\bar{X}\right|>\varepsilon\right)\leq \text{ (by the union bound) }\sum_{j=k}^{\infty}P\left(\max_{2^j\leq n<2^{j+1}}\left|\frac{S_n}{n}-\bar{X}\right|>\varepsilon\right)$$

$$P\left(\sup_{n\geq 2^k}\left|\frac{S_n}{n}-\overline{X}\right|>\varepsilon\right) \;\leq\; \sum_{j=k}^{\infty}P\left(\max_{2^j\leq n<2^{j+1}}\left|\frac{S_n}{n}-\overline{X}\right|>\varepsilon\right)$$

Each term on the right can be bounded by

$$P\left(\max_{2^j\leq n<2^{j+1}}\left|\frac{S_n}{n}-\overline{X}\right|>\varepsilon\right)\leq P\left(\max_{2^j\leq n<2^{j+1}}\left|S_n-n\overline{X}\right|>n\varepsilon\right)\leq$$

$$P\left(\max_{2^j\leq n<2^{j+1}}\left|S_n-n\overline{X}\right|>2^j\varepsilon\right)\leq\; P\left(\max_{1\leq n\leq 2^{j+1}}\left|S_n-n\overline{X}\right|\geq\frac{\varepsilon}{2}2^{j+1}\right)\leq$$

(using the random walk inequality $P(\max_{1\leq n\leq m}|S_n-n\overline{X}|\;\geq\;\mathrm{m}\varepsilon)\;\leq\;\frac{\sigma_X^2}{m\varepsilon^2}$)

$$\frac{\sigma^2}{(\varepsilon/2)^2\,2^{j+1}}=\frac{\sigma^2}{2^{j-1}\varepsilon^2}$$

Summing from j = k to j = ∞ shows that for any $\varepsilon > 0$,

$$P\left(\sup_{n\geq 2^k}\left|\frac{S_n}{n}-\overline{X}\right|>\varepsilon\right)\leq\sum_{j=k}^{\infty}\frac{\sigma^2}{2^{j-1}\varepsilon^2}=\frac{\sigma^2}{2^{k-2}\varepsilon^2}\xrightarrow[k\to\infty]{}0.$$

■

Martingale Convergence Theorem

Let $\{Z_n, n \geq 1\}$ be a martingale and assume there is some finite M such that $E\{|Z_n|\}<M$ for all n. Then there is a random variable Z such that, for a set of sample sequences with probability 1,

$$\lim_{n\to\infty} Z_n = Z$$

Example

Let Xn be a branching process with an expected number of offspring $\overline{Y}$ for each individual. Then $\{W_n = X_n/(\overline{Y})^n,\ n \geq 1\}$ is a martingale. If $\overline{Y} > 1$, the limiting random variable equals 0 with the probability <1 that the population eventually dies out and is positive otherwise. The interpretation is that, when n is sufficiently large, the population tends to grow at a constant exponential rate so that

$$X_n \underset{n\to\infty}{\longrightarrow} (\overline{Y})^n Z$$

where Z reflects the initial uncertainty in the growth rate when the population is small.