Lecture III: Probability Theory on Coin Toss Space

Dangxing Chen

Duke Kunshan University

Finite Probability Spaces

Random Variables, Distributions, and Expectations

Conditional Expectations

Martingales

Markov Processes

Example

- Toss the coin three times
- The set of all possible outcomes

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

- Suppose each toss the probability of head is p
- Suppose the tosses are independent
- Probabilities:

$$\mathbb{P}(HHH) = p^3, \mathbb{P}(HHT) = \mathbb{P}(HTH) = \mathbb{P}(THH) = p^2(1-p),$$
$$\mathbb{P}(HTT) = \mathbb{P}(THT) = \mathbb{P}(TTH) = p(1-p)^2, \mathbb{P}(TTT) = (1-p)^3.$$

- Events: subsets of Ω

$$\begin{aligned} \text{“The first toss is a head”} &= \{\omega \in \Omega; \omega_1 = H\} \\ &= \{HHH, HHT, HTH, HTT\}. \end{aligned}$$

- $\mathbb{P}(\text{First toss is a head}) = \mathbb{P}(HHH) + \mathbb{P}(HHT) + \mathbb{P}(HTH) + \mathbb{P}(HTT) = p$

Probability space

- **Def:** A finite probability space consists of a sample space Ω and a probability measure $\mathbb{P}$. The sample space Ω is a nonempty finite set and the probability measure $\mathbb{P}$ is a function that assigns to each element ω of Ω a number in $[0, 1]$ s.t.

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1.$$

An event is a subset of Ω, and we define the probability of an event A to be

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega).$$

Random variables

- Let $(\Omega, \mathbb{P})$ be a finite probability space. A random variable is a real-valued function defined on Ω.
- E.g. Stock prices with $S_0 = 4, u = 2, d = \frac{1}{2}$, three-periods

$$S_0(\omega_1\omega_2\omega_3) = 4 \text{ for all } \omega_1\omega_2\omega_3 \in \Omega,$$

$$S_1(\omega_1\omega_2\omega_3) = \begin{cases} 8, & \text{if } \omega_1 = H, \\ 2, & \text{if } \omega_1 = T, \end{cases}$$

$$S_2(\omega_1\omega_2\omega_3) = \begin{cases} 16, & \text{If } \omega_1 = \omega_2 = H, \\ 4, & \text{if } \omega_1 \neq \omega_2, \\ 1, & \text{if } \omega_1 = \omega_2 = T, \end{cases}$$

$$S_3(\omega_1\omega_2\omega_3) = \begin{cases} 32, & \text{if } \omega_1 = \omega_2 = \omega_3 = H, \\ 8, & \text{if there are two heads and one tail,} \\ 2, & \text{if there is one head and two tails,} \\ 0.5, & \text{if } \omega_1 = \omega_2 = \omega_3 = T. \end{cases}$$

Distribution

- **Def:** The distribution of a random variable is a specification of the probabilities that the random variable takes various values.
- **Remark:** random variable $\neq$ distribution!
- E.g. of different random variables:
  $X =$ Total number of heads, $Y =$ Total number of tails
- E.g. of different distributions: $\widetilde{P}(X = 0) = \frac{1}{8}$, $\mathbb{P}(X = 0) = \frac{1}{4}$

Expectation

- **Def:** Let X be a random variable defined on a probability space $(\Omega, \mathbb{P})$. The **expectation (or expected value)** of X is defined to be

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega).$$

- Linearity: $\mathbb{E}[c_1 X_1 + c_2 X_2] = c_1 \mathbb{E}[X_1] + c_2 \mathbb{E}[X_2]$
- Suppose $f(x) = ax + b$, we have $\mathbb{E}[f(x)] = f(\mathbb{E}[X])$
- **Question**: What about a general f?
- E.g., $f(S) = \max(S - K, 0)$

Convex function

- **Def:** f is called **convex** if $\forall 0 \le t \le 1$ and $\forall x_1, x_2$:

$$f(tx_1 + (1-t)x_2) \le tf(x_1) + (1-t)f(x_2).$$

- If f'' exists, then convexity $\Rightarrow f'' \ge 0$

Convex function

- **Def:** f is called **convex** if $\forall 0 \leq t \leq 1$ and $\forall x_1, x_2$:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

- If f'' exists, then convexity $\Rightarrow f'' \geq 0$
  - $f''(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h} = \lim_{h \to 0} \frac{\frac{f(x)-f(x+h)}{-h} - \frac{f(x-h)-f(x)}{-h}}{h} = \lim_{h \to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$
  - Convexity $\Rightarrow f(x) = f\left(\frac{x+h}{2} + \frac{x-h}{2}\right) \leq \frac{1}{2}f(x+h) + \frac{1}{2}f(x-h)$
  - Therefore, $f''(x) \geq 0$
- **Question**: Why not our calculus Def?

Convex function

- **Def:** f is called **convex** if $\forall 0 \le t \le 1$ and $\forall x_1, x_2$:

$$f(tx_1 + (1-t)x_2) \le tf(x_1) + (1-t)f(x_2).$$

- If f'' exists, then convexity $\Rightarrow f'' \ge 0$
  - $f''(x) = \lim_{h\to 0} \frac{f'(x+h)-f'(x)}{h} = \lim_{h\to 0} \frac{\frac{f(x)-f(x+h)}{-h} - \frac{f(x-h)-f(x)}{-h}}{h} = \lim_{h\to 0} \frac{f(x+h)-2f(x)+f(x-h)}{h^2}$
  - Convexity $\Rightarrow f(x) = f\left(\frac{x+h}{2} + \frac{x-h}{2}\right) \le \frac{1}{2}f(x+h) + \frac{1}{2}f(x-h)$
  - Therefore, $f''(x) \ge 0$
- **Question**: Why not our calculus Def?
- E.g., $f(S) = \max(S-K, 0)$ is not differentiable at $S = K$
- $f(S) = \max(S-K, 0)$ is convex

Jensen's inequality

- **THM (Jensen's Ineq):** Let X be a random variable, let $f(x)$ be a convex function. Then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

Jensen's inequality

- **THM (Jensen's Ineq):** Let X be a random variable, let $f(x)$ be a convex function. Then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

- Proof
  - Induction
  - $n = 2$: consider $X = \begin{cases} x_1, & \text{with } p_1, \\ x_2, & \text{with } p_2. \end{cases}$ and $p_1 + p_2 = 1$
  - $f(\mathbb{E}[X]) = f(p_1 x_1 + p_2 x_2) \leq p_1 f(x_1) + p_2 f(x_2)$ by convexity
  - $\mathbb{E}[f(X)] = p_1 f(x_1) + p_2 f(x_2)$
- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, K = 5, n = 1, p = \frac{1}{2}, f(S_1) = \max(S_1 - K, 0)$
  - $f(\mathbb{E}[S_1]) = f\left(\frac{1}{2} \cdot 8 + \frac{1}{2} \cdot 2\right) = \max(5 - 5, 0) = 0$
  - $\mathbb{E}[f(S_1)] = \mathbb{E}[\max(S_1 - K, 0)] = \left[\frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 0\right] = \frac{3}{2}$

Jensen's inequality - continued

- Now suppose Jensen's ineq is true for n, show for $n+1$
- $f(\sum_{i=1}^{n} p_i x_i) \leq \sum_{i=1}^{n} p_i f(x_i)$

Jensen's inequality - continued

- Now suppose Jensen's ineq is true for n, show for $n+1$
- $f(\sum_{i=1}^{n} p_i x_i) \leq \sum_{i=1}^{n} p_i f(x_i)$
- By convexity,

$$
\begin{aligned}
f\left(\sum_{i=1}^{n+1} p_i x_i\right) &= f\left((1-p_{n+1})\sum_{i=1}^{n} \frac{p_i}{1-p_{n+1}} x_i + p_{n+1}x_{n+1}\right) \\
&\leq (1-p_{n+1}) f\left(\sum_{i=1}^{n} \frac{p_i}{1-p_{n+1}} x_i\right) + p_{n+1} f(x_{n+1}) \\
&\leq \sum_{i=1}^{n} p_i f(x_i) + p_{n+1} f(x_{n+1}) \\
&= \sum_{i=1}^{n+1} p_i f(x_i).
\end{aligned}
$$

Conditional expectations

- Recall risk-neutral probability $\widetilde{p}, \widetilde{q}$
- $S_n(\omega_1 \dots \omega_n) = \frac{[\widetilde{p} S_{n+1}(\omega_1 \dots \omega_n H) + \widetilde{q} S_{n+1}(\omega_1 \dots \omega_n T)]}{1+r}$
- Define conditional expectation

$$\mathbb{E}_n^{\mathbb{Q}}[S_{n+1}](\omega_1 \dots \omega_n) = \widetilde{p} S_{n+1}(\omega_1 \dots \omega_n H) + \widetilde{q} S_{n+1}(\omega_1 \dots \omega_n T)$$

- Intuition behind: expectation based on

Conditional expectations

- Recall risk-neutral probability $\widetilde{p}, \widetilde{q}$
- $S_n(\omega_1 \ldots \omega_n) = \frac{[\widetilde{p} S_{n+1}(\omega_1 \ldots \omega_n H) + \widetilde{q} S_{n+1}(\omega_1 \ldots \omega_n T)]}{1+r}$
- Define conditional expectation

$$\mathbb{E}_n^{\mathbb{Q}}[S_{n+1}](\omega_1 \ldots \omega_n) = \widetilde{p} S_{n+1}(\omega_1 \ldots \omega_n H) + \widetilde{q} S_{n+1}(\omega_1 \ldots \omega_n T)$$

- Intuition behind: expectation based on the information at time n, i.e., knowledge of first n coin tosses
- **Def**: Let $1 \leq n \leq N$. Let $\omega_1 \ldots \omega_n$ be given and fixed. There are 2^{N-n} possible continuations of $\omega_{n+1} \ldots \omega_N$ of the sequence fixed $\omega_1 \ldots \omega_n$. Denote by $\#H(\omega_{n+1} \ldots \omega_N)$ and $\#T(\omega_{n+1} \ldots \omega_N)$ the number of heads and tails in the continuation $\omega_{n+1} \ldots \omega_N$. Define

$$\begin{aligned} &\mathbb{E}_n[X](\omega_1 \ldots \omega_n) \\ &= \sum_{\omega_{n+1} \ldots \omega_N} p^{\#H(\omega_{n+1} \ldots \omega_N)} q^{\#T(\omega_{n+1} \ldots \omega_N)} X(\omega_1 \ldots \omega_n \omega_{n+1} \ldots \omega_N). \end{aligned}$$

Examples

- For $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, \widetilde{p} = \widetilde{q} = \frac{1}{2}$

Examples

- For $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, \widetilde{p} = \widetilde{q} = \frac{1}{2}$
- $\widetilde{\mathbb{E}}_1[S_3](H) = 12.5, \widetilde{\mathbb{E}}_1[S_3](T) = 3.125$
- Two extreme cases:
  - $\mathbb{E}_0[X] =$

Examples

- For $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, \widetilde{p} = \widetilde{q} = \frac{1}{2}$
- $\widetilde{\mathbb{E}}_1[S_3](H) = 12.5, \widetilde{\mathbb{E}}_1[S_3](T) = 3.125$
- Two extreme cases:
  - $\mathbb{E}_0[X] = \mathbb{E}[X]$
  - $\mathbb{E}_N[X] =$

Examples

- For $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, \widetilde{p} = \widetilde{q} = \frac{1}{2}$
- $\widetilde{\mathbb{E}}_1[S_3](H) = 12.5, \widetilde{\mathbb{E}}_1[S_3](T) = 3.125$
- Two extreme cases:
  - $\mathbb{E}_0[X] = \mathbb{E}[X]$
  - $\mathbb{E}_N[X] = X$

Theorem

THM (Fundamental properties of conditional expectations): Let $N \in \mathbb{Z}^+$, X and Y be random variables depending on the first N coin tosses. Let $0 \leq n \leq N$ be given.

- **Linearity:** $\forall c_1, c_2 \in \mathbb{R}$,

$$\mathbb{E}_n[c_1 X + c_2 Y] = c_1 \mathbb{E}_n[X] + c_2 \mathbb{E}_n[Y].$$

- **Taking out what is known:** If X actually depends only on the first n coin tosses, then $\mathbb{E}_n[XY] = X \cdot \mathbb{E}_n[Y]$.
- **Iterated conditioning:** If $0 \leq n \leq m \leq N$, then

$$\mathbb{E}_n[\mathbb{E}_m[X]] = \mathbb{E}_n[X].$$

- **Independence:** If X depends only on tosses $n+1$ through N, then $\mathbb{E}_n[X] = \mathbb{E}[X]$.
- **Conditional Jensen's inequality:** If $f(x)$ is a convex function, then

$$f(\mathbb{E}_n[X]) \leq \mathbb{E}_n[f(X)].$$

Examples

- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, p = q = \frac{1}{2}$

Examples

- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, p = q = \frac{1}{2}$
  - **Linearity**:

$$\mathbb{E}_1[S_2](H) = \frac{1}{2} \cdot 16 + \frac{1}{2} \cdot 4 = 10,$$
$$\mathbb{E}_1[S_3](H) = \frac{1}{4} \cdot 32 + \frac{1}{2} \cdot 8 + \frac{1}{4} 2 = 12.5,$$
$$\mathbb{E}_1[S_2 + S_3](H) = \frac{1}{4}(32 + 16) + \frac{1}{4}(8 + 16) + \frac{1}{4}(8 + 4) + \frac{1}{4}(2 + 4)$$
$$= 22.5$$

  - **Taking out what is known:**

$$\mathbb{E}_1[S_1 S_2](H) = S_1(H)\mathbb{E}_1[S_2](H) = 8 \cdot 10 = 80.$$

Example

- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, p = q = \frac{1}{2}$

Example

- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, p = q = \frac{1}{2}$
- **Iterated conditioning:**

$$\mathbb{E}_2[S_3](HH) = \frac{1}{2} \cdot 32 + \frac{1}{2} \cdot 8 = 20,$$
$$\mathbb{E}_2[S_3](HT) = \frac{1}{2} \cdot 8 + \frac{1}{2} \cdot 2 = 5,$$
$$\mathbb{E}_2[S_3](TH) = \frac{1}{2} \cdot 8 + \frac{1}{2} \cdot 2 = 5,$$
$$\mathbb{E}_2[S_3](TT) = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 0.5 = 1.25,$$
$$\mathbb{E}_1[\mathbb{E}_2[S_3]](H) = \frac{1}{2} \cdot \mathbb{E}_2[S_3](HH) + \frac{1}{2} \cdot \mathbb{E}_2[S_3](HT)$$
$$+ \frac{1}{2} \cdot 20 + \frac{1}{2} \cdot 5 = 12.5,$$
$$\mathbb{E}_1[S_3](H) = \frac{1}{4} \cdot 32 + \frac{1}{2} \cdot 8 + \frac{1}{4} \cdot 2 = 12.5$$

Example

- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, p = q = \frac{1}{2}$
- $\frac{S_2}{S_1}$ takes either 2 or $\frac{1}{2}$, not depend on the first toss

Example

- E.g., $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, p = q = \frac{1}{2}$
- $\frac{S_2}{S_1}$ takes either 2 or $\frac{1}{2}$, not depend on the first toss
- **Independence:**

$$\begin{aligned}
\mathbb{E}_1\left[\frac{S_2}{S_1}\right](H) &= \frac{1}{2}\cdot\frac{S_2(HH)}{S_1(H)} + \frac{1}{2}\cdot\frac{S_2(HT)}{S_1(H)} \\
&= \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot\frac{1}{2} = \frac{5}{4}. \\
\mathbb{E}_1\left[\frac{S_2}{S_1}\right](T) &= \frac{1}{2}\cdot\frac{S_2(TH)}{S_1(T)} + \frac{1}{2}\cdot\frac{S_2(TT)}{S_1(T)} \\
&= \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot\frac{1}{2} = \frac{5}{4}, \\
\mathbb{E}\left[\frac{S_2}{S_1}\right] &= \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot\frac{1}{2} = \frac{5}{4}.
\end{aligned}$$

Motivation

- Recall risk-neural probability:

$$\widetilde{p} = \frac{1 + r - d}{u - d}, \ \widetilde{q} = \frac{u - (1 + r)}{u - d}.$$

- Under risk-neural probability, we have

Motivation

- Recall risk-neural probability:

$$\widetilde{p} = \frac{1+r-d}{u-d}, \ \widetilde{q} = \frac{u-(1+r)}{u-d}.$$

- Under risk-neural probability, we have

$$\widetilde{\mathbb{E}}_n\left[\frac{S_{n+1}}{(1+r)^{n+1}}\right] = \frac{\widetilde{p}S_n(H) + \widetilde{q}S_n(T)}{(1+r)^{n+1}} = \frac{\frac{u(1+r)-ud+ud-(1+r)d}{u-d}S_n}{(1+r)^{n+1}}$$
$$= \frac{S_n}{(1+r)^n}.$$

- Let $M_n = \frac{S_n}{(1+r)^n}$, then $M_n = \widetilde{\mathbb{E}}_n[M_{n+1}]$
- Interpretation: For risk-neutral probability, the best estimate based on the information at the current time is the discounted future stock price

Martingales

- **Def:** Consider the binomial asset-pricing model. Let $M_0, M_1, \ldots, M_N$ be a sequence of random variables, with each M_n depending only on the first n coin tosses (and M_0 constant). Such a sequence of random variables is called an adapted stochastic process.
  - **Martingale:** $M_n = \mathbb{E}_n[M_{n+1}]$
  - **Submartingale:** $M_n \leq \mathbb{E}_n[M_{n+1}]$, have a tendency to increase
  - **Supermartingale:** $M_n \geq \mathbb{E}_n[M_{n+1}]$, have a tendency to decrease
- **Remark:** The expectation of a martingale is constant over time

$$\mathbb{E}[M_0] = \mathbb{E}[M_1] = \cdots = \mathbb{E}[M_N].$$

Examples

- $S_0 = 4, u = 2, d = \frac{1}{2}, r = \frac{1}{4}$
- Martingale:
  - $\widetilde{p} = \frac{1+r-d}{u-d}, \ \widetilde{q} = \frac{u-(1+r)}{u-d}$.
  - $M_n = \frac{S_n}{(1+r)^n}$, then $M_n = \widetilde{\mathbb{E}}_n[M_{n+1}]$
- Submartingale:
  - $p = \frac{2}{3}, \ q = \frac{1}{3}$

Examples

- $S_0 = 4, u = 2, d = \frac{1}{2}, r = \frac{1}{4}$
- Martingale:
  - $\widetilde{p} = \frac{1+r-d}{u-d},\ \widetilde{q} = \frac{u-(1+r)}{u-d}.$
  - $M_n = \frac{S_n}{(1+r)^n}$, then $M_n = \widetilde{\mathbb{E}}_n[M_{n+1}]$
- Submartingale:
  - $p = \frac{2}{3},\ q = \frac{1}{3}$
  - $\mathbb{E}_n[S_{n+1}] = \frac{2}{3} \cdot 2S_n + \frac{1}{3} \cdot \frac{1}{2} S_n = \frac{3}{2} S_n$
  - Let $M_n = \frac{S_n}{(1+r)^n}$
  - $\mathbb{E}_n[M_{n+1}] = \left(\frac{4}{5}\right)^{n+1} \cdot \frac{3}{2} S_n = \left(\frac{4}{5}\right)^n \cdot \frac{4}{5} \cdot \frac{3}{2} \cdot S_n \geq \left(\frac{4}{5}\right)^n S_n$
  - In real markets, we expect submartingale

Hedging

- Consider the replicating strategy in the binomial model with N coin tosses
- Position of Δ_n shares of stock and holds until time $n+1$
- Δ_n only depend on the first n coin tosses, i.e., adapted
- Wealth equation:

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n).$$

- **THM:** Under risk-neutral measure, $\frac{X_n}{(1+r)^n}$

Hedging

- Consider the replicating strategy in the binomial model with N coin tosses
- Position of Δ_n shares of stock and holds until time $n+1$
- Δ_n only depend on the first n coin tosses, i.e., adapted
- Wealth equation:

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n).$$

- **THM:** Under risk-neutral measure, $\frac{X_n}{(1+r)^n}$ is a martingale

$$\begin{aligned}
\widetilde{\mathbb{E}}_n\left[\frac{X_{n+1}}{(1+r)^{n+1}}\right] &= \widetilde{\mathbb{E}}_n\left[\frac{\Delta_n S_{n+1}}{(1+r)^{n+1}} + \frac{X_n - \Delta_n S_n}{(1+r)^n}\right] \\
&= \Delta_n \widetilde{\mathbb{E}}_n\left[\frac{S_{n+1}}{(1+r)^{n+1}}\right] + \frac{X_n - \Delta_n S_n}{(1+r)^n} \\
&= \Delta_n \frac{S_n}{(1+r)^n} + \frac{X_n - \Delta_n S_n}{(1+r)^n} \\
&= \frac{X_n}{(1+r)^n}.
\end{aligned}$$

Wealth equation - continued

- **Cor:** Under the same condition, we have

$$\widetilde{\mathbb{E}}\left[\frac{X_n}{(1+r)^n}\right] = X_0.$$

- Proof:

Wealth equation - continued

- **Cor:** Under the same condition, we have

$$\widetilde{\mathbb{E}}\left[\frac{X_n}{(1+r)^n}\right] = X_0.$$

- Proof: iterated conditioning
- Two consequences:
  - No-arbitrary: With $X_0 = 0$, if $X_N(\omega) \geq 0, \forall \omega$ and $X_N(\overline{\omega}) > 0$ for at least one $\overline{\omega}$, then $\widetilde{\mathbb{E}}[X_0] = 0$ and $\widetilde{\mathbb{E}}\left[\frac{X_N}{(1+r)^N}\right] > 0$
  - Pricing:
    - $\frac{V_n}{(1+r)^n} = \frac{X_n}{(1+r)^n} = \widetilde{\mathbb{E}}_n\left[\frac{X_N}{(1+r)^N}\right] = \widetilde{\mathbb{E}}_n\left[\frac{V_N}{(1+r)^N}\right]$
    - $V_0 = \widetilde{\mathbb{E}}\left[\frac{V_N}{(1+r)^N}\right]$

Finite Probability Spaces

Random Variables, Distributions, and Expectations

Conditional Expectations

Martingales

Markov Processes

Markov process

- **Def:** Consider the binomial model. Let $X_0, X_1, \dots, X_N$ be an adapted process. If $\forall n$, $\exists g(x)$ s.t.

$$\mathbb{E}_n[f(X_{n+1})] = g(X_n),$$

that we say that $X_0, X_1, \dots, X_N$ is a **Markov process**.

- E.g.,
  - In the binomial model, $\mathbb{E}_n[f(S_{n+1})](\omega) = pf(uS_n(\omega)) + qf(dS_n(\omega))$
  - Long-memory process is not Markovian
- **LEM (Independence):** In the binomial model. Suppose the random variables $X^1, \dots, X^K$ depend only on $1 \dots n$ and random variables $Y^1, \dots, Y^L$ depend only on $n+1 \dots N$. Consider the function $f(x^1, \dots, x^K, y^1, \dots, y^L)$ and define

$$g(x^1, \dots, x^K) = \mathbb{E}\left[f(x^1, \dots, x^K, Y^1, \dots, Y^L)\right].$$

Then

$$\mathbb{E}_n\left[f\left(X^1, \dots, X^K, Y^1, \dots, Y^L\right)\right] = g\left(X^1, \dots, X^K\right).$$

Non-Markov process

- For $S_0 = 4, u = 2, d = \frac{1}{2}, N = 3, \widetilde{p} = \widetilde{q} = \frac{1}{2}$
- In the binomial model, consider $\{(S_n, M_n)\}_{n=1}^N$, where $M_n = \max_{0 \le k \le n} S_k$
- M_n is non-Markovian
- Suppose $\widetilde{p} = \widetilde{q} = 0.5$

$$\begin{aligned} M_2(TH) &= M_2(TT) = 4, \\ \widetilde{\mathbb{E}}_2[M_3](TH) &= \frac{1}{2}M_3(THH) + \frac{1}{2}M_3(THT) = 6, \\ \widetilde{\mathbb{E}}_2[M_3](TT) &= \frac{1}{2}M_3(TTH) + \frac{1}{2}M_3(TTT) = 4. \end{aligned}$$

- Sometimes it is possible to make it Markovian

Multi-dimensional Markov process

- **Def:** Consider the binomial model. Let $\{(X_n^1, \ldots, X_n^K)\}_{n=0}^N$ be a K-dimensional adapted process. If $\forall n$, $\exists g$ s.t.

$$E_n\left[f\left(X_{n+1}^1, \ldots, X_{n+1}^K\right)\right] = g\left(X_n^1, \ldots, X_n^K\right),$$

  we say that $\{(X_n^1, \ldots, X_n^K)\}_{n=0}^N$ is a K-dimensional Markov process.
- E.g., $\{(S_n, M_n)\}_{n=1}^N$, where $M_n = \max_{0\leq k\leq n} S_k$

Multi-dimensional Markov process

- **Def:** Consider the binomial model. Let $\{(X_n^1, \ldots, X_n^K)\}_{n=0}^N$ be a K-dimensional adapted process. If $\forall n$, $\exists g$ s.t.

$$E_n \left[f \left(X_{n+1}^1, \ldots, X_{n+1}^K \right) \right] = g \left(X_n^1, \ldots, X_n^K \right),$$

  we say that $\{(X_n^1, \ldots, X_n^K)\}_{n=0}^N$ is a K-dimensional Markov process.
- E.g., $\{(S_n, M_n)\}_{n=1}^N$, where $M_n = \max_{0 \le k \le n} S_k$
  - Let $S_{n+1} = Y S_n$ and $M_{n+1} = \max(Y S_n, M_n)$, $Y = \begin{cases} u, & H, \\ d, & T \end{cases}$
  - Let $g(s, m) = \mathbb{E}[f(sY, \max(m, sY))] =$ $p \cdot f(us, \max(m, us)) + q \cdot f(ds, \max(m, ds))$
  - $\mathbb{E}_n[f(S_{n+1}, M_{n+1})] = g(S_n, M_n)$
- **Remark:** The pricing can be done with the same way!