STAT 205A - Problem Set 07

Walid Krichene (23265217)

October 22, 2013

(7.1) Suppose S and T are stopping times. Are the following necessarily stopping times? Give proof or counter-example.

(a) $\min(S, T)$
Assume stopping times are with respect to $\mathcal{G}_n = \sigma(X_0, X_1, \dots, X_n)$.
We have $\{\min(S, T) \le n\} = \{S \le n\} \cup \{T \le n\}$, and since $\{S \le n\} \in \mathcal{G}_n$ and $\{T \le n\} \in \mathcal{G}_n$, their union is also in $\mathcal{G}_n$. Therefore $\min(S, T)$ is a stopping time.

(b) $\max(S, T)$
We have $\{\max(S, T) \le n\} = \{S \le n\} \cap \{T \le n\} \in \mathcal{G}_n$ since $\mathcal{G}$ is closed under intersection. Therefore $\max(S, T)$ is a stopping time.

(c) $S + T$
We have $\{S + T \le n\} = \cup_{s=0}^{n} \{S = s\} \cap \{T \le n - s\}$. Since S and T are stopping times, we have for all $0 \le s \le n$, $\{S = s\} \in \mathcal{G}_s \subseteq \mathcal{G}_n$, and $\{T \le n - s\} \in \mathcal{G}_{n-s} \subseteq \mathcal{G}_n$. And since $\mathcal{G}_n$ is closed under finite union and intersection, $\{S + T \le n\} \in \mathcal{G}_n$. Therefore $S + T$ is a stopping time.

(7.2) Let (X_i) be i.i.d. with $\mathrm{E}\, X^2 < \infty$. Let $S_n = \sum_{i=1}^{n} X_i$. Let T be a bounded stopping time. Is it true in general that

$$\mathrm{var}(S_T) = \mathrm{var}(X_1)\, \mathrm{E}\, T$$

Is it true in the special case $\mathrm{E}\, X = 0$?

answer Let $\sigma^2 = \mathrm{E}\, X^2$ and $\mu = \mathrm{E}\, X$. Suppose T is (almost surely) bounded by the constant M. We have $S_T = \sum_{i=0}^{M} X_i 1_{i \le T}$. Then we have

$$\begin{aligned}
\mathrm{E}\left(\sum_{i=0}^{M} X_i 1_{i \le T}\right) &= \sum_{i=0}^{M} \mathrm{E}[X_i 1_{i \le T}] \\
&= \sum_{i=0}^{M} \mathrm{E}[X_i] P(i \le T) && \text{by independence of } X_i \text{ and } \{T \ge i\} \\
&= \mu \sum_{i=0}^{M} P(i \le T) \\
&= \mu\, \mathrm{E}\, T
\end{aligned}$$

here independence of X_i and $\{T \geq i\}$ follows from the observation that $\{T \geq i\} = \{T \leq i-1\}^c \in \mathcal{F}_{i-1}$, and the fact that X_i is independent of $\mathcal{F}_{i-1}$. Similarly, we can compute

$$\begin{aligned} \mathrm{E}\left(\sum_{i=0}^{M} X_i 1_{i\leq T}\right)^2 &= \mathrm{E}\left(\sum_{i=0}^{M} X_i^2 1_{i\leq T} + 2\sum_{0\leq i<j\leq n} X_i X_j 1_{i\leq T} 1_{j\leq T}\right) \\ &= \sum_{i=0}^{M} \mathrm{E}[X_i^2 1_{i\leq T}] + 2\sum_{0\leq i<j\leq n} \mathrm{E}[X_i X_j 1_{j\leq T}] \qquad \text{since } i<j \text{ and } j\leq T \text{ implies } i\leq T \end{aligned}$$

then we have

- $\{T \geq i\}^c = \{T \leq i-1\} \in \mathcal{F}_{i-1}$, and X_i is independent from $\mathcal{F}_{i-1}$ therefore X_i and $\{T \geq i\}$ are independent
- $\{T \geq j\}^c = \{T \leq j-1\} \in \mathcal{F}_{j-1}$. Since X_j is independent from $\mathcal{F}_{j-1}$ and X_i $(i < j)$, then X_j and $X_i 1_{i\leq T}$ are independent.

therefore we can write

$$\begin{aligned} \mathrm{E}\left(\sum_{i=0}^{n} X_i 1_{i\leq T}\right)^2 &= \sum_{i=0}^{M} \mathrm{E}[X_i^2]\,\mathrm{E}[1_{T\geq i}] + 2\sum_{0\leq i<j\leq n} \mathrm{E}[X_j]\,\mathrm{E}[X_i 1_{T\geq j}] \\ &= \sigma^2 \sum_{i=0}^{M} P(T\geq i) + 2\mu \sum_{0\leq i<j\leq n} \mathrm{E}[X_i 1_{T\geq j}] \\ &= \sigma^2\,\mathrm{E}[T] + 2\mu \sum_{0\leq i<j\leq n} \mathrm{E}[X_i 1_{T\geq j}] \end{aligned}$$

Combining the previous results, we have

$$\begin{aligned} \mathrm{var}(S_T) &= \sigma^2\,\mathrm{E}[T] + 2\mu \sum_{0\leq i<j\leq n} \mathrm{E}[X_i 1_{T\geq j}] - \mu^2 (\mathrm{E}\,T)^2 \\ &= \mathrm{E}[T]\left(\sigma^2 - \mu^2\,\mathrm{E}\,T\right) + 2\mu \sum_{0\leq i<j\leq n} \mathrm{E}[X_i 1_{T\geq j}] \qquad (1) \end{aligned}$$

This is not equal to $\mathrm{E}[T]\,\mathrm{var}(X_1)$ in general. Counter-example: let $X_0, X_1, \ldots$ be i.i.d. Bernoulli(1/2) variables, and let $T = X_0$. We have

$$\mathrm{E}(T) = \mathrm{E}[X] = \frac{1}{2}$$

thus by equation (1), we have

$$\begin{aligned} \mathrm{var}(S_T) &= \mathrm{E}[T]\left(\sigma^2 - \mu^2\,\mathrm{E}\,T\right) + 2\mu \sum_{0\leq i<j\leq n} \mathrm{E}[X_i 1_{T\geq j}] \\ &\geq \mathrm{E}[T]\left(\sigma^2 - \mathrm{E}\,T\mu^2\right) && \text{since every term in the sum is non-negative} \\ &> \mathrm{E}[T]\left(\sigma^2 - \mu^2\right) && \text{since } \mathrm{E}\,T = \frac{1}{2} \text{ and } \mu > 0 \\ &= \mathrm{E}[T]\,\mathrm{var}[X] \end{aligned}$$

If $\mathrm{E}\,X = \mu = 0$, then (1) simply becomes

$$\mathrm{var}(S_T) = \sigma^2\,\mathrm{E}[T] = \mathrm{E}[T]\,\mathrm{var}(X)$$

(7.3) Let (X_i) be a sequence of random variables, and let $\mathcal{T}$ be its tail $\sigma-$field. Let $S_n = \sum_{i=1}^n X_i$. Let $b_n \uparrow \infty$ be constants. Which of the following events must be in $\mathcal{T}$? Give proof or counter-example.

(i) $A = \{X_n \to 0\}$
Yes, $B \in \mathcal{T}$.
Intuitively, this follows from the fact that a finite number of terms does not affect convergence to 0. More precisely, we have for all $k \in \mathbb{N}$, $X_n \to 0$ if and only if $(X_n)_{n\geq k} \to 0$, if and only if for all $\epsilon \in \mathbb{Q}_+$, there exists $N \geq k$ such that for all $n \geq N$, $|X_n| \leq \epsilon$. Thus

$$\{X_n \to 0\} = \{(X_n)_{k\geq n} \to 0\} \qquad = \cap_{\epsilon\in\mathbb{Q}_+} \cup_{N\geq k} \cap_{n\geq N} |X_n|^{-1}([0,\epsilon])$$

where for all ϵ and $N \geq k$, for all $n \geq N$, $|X_n|^{-1}([0,\epsilon]) \in \mathcal{G}_n \subseteq \mathcal{G}_k$. Therefore $\cap_{\epsilon\in\mathbb{Q}_+} \cup_{N\geq k} \cap_{n\geq N} |X_n|^{-1}([0,\epsilon])$ is also in $\mathcal{G}_k$ (closed under countable unions and intersections). Therefore

$$\{X_n \to 0\} \in \mathcal{G}_k$$

since this holds for all $k \in \mathbb{N}$, the event is also in $\cap_{k\in\mathbb{N}}\mathcal{G}_k = \mathcal{T}$.

(ii) $B = \{S_n \text{ converges}\}$
Yes, $B \in \mathcal{T}$.
Let $k \in \mathbb{N}$. We have S_n converges if and only if $(S_n)_{n\geq k}$ converges, if and only if $(S_n)_{n\geq k}$ is Cauchy, if and only if for all rational $\epsilon > 0$, there exists $N \geq k$ such that for all $m > n \geq N$, $|S_n - S_m| \leq \epsilon$. Therefore

$$\{S_n \text{ converges}\} = \cap_{\epsilon\in\mathbb{Q}_+} \cup_{N\geq k} \cap_{n>m\geq N} |S_m - S_n|^{-1}([0,\epsilon])$$

where for all ϵ and $N \geq k$, for all $n > m \geq N$, $|S_n - S_m|^{-1}([0,\epsilon]) \in \mathcal{G}_N \subseteq \mathcal{G}_k$, since $S_n - S_m = X_{m+1} + \cdots + X_n$. Therefore

$$\{S_n \text{ converges}\} \in \mathcal{G}_k$$

and since this holds for any $k \in \mathbb{N}$, the event is also in $\mathcal{T}$.

(iii) $C = \{X_n > b_n \text{ infinitely often}\}$
Yes, it is in $\mathcal{T}$.
Let $k \in \mathcal{N}$. Then we have $\{X_n > b_n \text{ infinitely often}\}$ if and only if $\{X_n > b_n, n \geq k \text{ infinitely often}\}$. Thus

$$\{X_n > b_n \text{ infinitely often}\} = \cap_{N\geq k} \cup_{n\geq N} \{X_n > b_n\}$$

where for all $N \geq k$ and any $n \geq N$, the event $\{X_n > b_n\}$ is in $\mathcal{G}_n \subseteq \mathcal{G}_k$. Therefore

$$\{X_n > b_n \text{ infinitely often}\} \in \mathcal{G}_k$$

since this holds for all $k \in \mathbb{N}$, the event is in $\mathcal{T}$.

(iv) $D = \{S_n > b_n \text{infinitely often}\}$ No, D is not in $\mathcal{T}$ in general.
Counter example: let $b_n = n$ for all n, and define X_0 to be Bernoulli(1/2), and $X_i = 1$ (deterministic) for all $i \geq 1$. Then we have $S_n > b_n$ if and only $X_0 > 0$ if and only if $X_0 = 1$, thus

$$\{S_n > b_n \text{infinitely often}\} = \{X_0 = 1\}$$

which is in $\mathcal{G}_0$, but not in any $\mathcal{G}_i$, $i \geq 1$ (these are trivial sub-σ-fields since X_i is deterministic), thus $\{S_n > b_n \text{ infinitely often }\}$ is not in $\mathcal{T}$.

(v) $\{\frac{\sqrt{\sum_{i=1}^n X_i^2}}{S_n} \to 0\}$
Yes, it is in the tail σ-field.

Let $k \in \mathbb{N}$. We first show that $\frac{\sqrt{\sum_{i=1}^n X_i^2}}{S_n} \to 0$ if and only if $\frac{\sqrt{\sum_{i=k+1}^n X_i^2}}{S_n - S_k} \to 0$.

- suppose $\frac{\sqrt{\sum_{i=1}^n X_i^2}}{S_n} \to 0$. Then we necessarily have $\sum |X_n| \to \infty$, otherwise, $|S_n|$ would be bounded, and this contradicts the assumption that $\frac{\sqrt{\sum_{i=1}^n X_i^2}}{S_n}$ converges 0. We have

$$\sqrt{\sum_{i=k+1}^{n} X_i^2} \leq \sqrt{\sum_{i=1}^{n} X_i^2}$$

and

$$\frac{|S_n - S_k|}{|S_n|} \geq \frac{|S_n| - |S_k|}{|S_n|} = 1 - \frac{|S_k|}{|S_n|}$$

which converges to 1 as $n \to \infty$, since $|S_k|$ is finite and $|S_n| \to \infty$. Therefore for n large enough, $|S_n - S_k| \geq \frac{1}{2}|S_n|$, therefore for n large enough,

$$\frac{\sqrt{\sum_{i=k+1}^n X_i^2}}{|S_n - S_k|} \leq 2\frac{\sqrt{\sum_{i=1}^n X_i^2}}{|S_n|}$$

and the right hand side converges to 0 by assumption. This proves that the modified sequence converges to 0.

- Conversely, assume that the modified sequence $\frac{\sqrt{\sum_{i=k+1}^n X_i^2}}{|S_n - S_k|} \to 0$. Then $|S_n| \to \infty$, and

$$\frac{\sqrt{\sum_{i=1}^n X_i^2}}{|S_n|} \leq \frac{\sqrt{\sum_{i=k+1}^n X_i^2}}{|S_n - S_k|}\frac{|S_n - S_k|}{|S_n|}$$

where $\frac{\sqrt{\sum_{i=k+1}^n X_i^2}}{|S_n - S_k|} \to 0$ by assumption, and $\frac{|S_n - S_k|}{|S_n|} \to 1$. Thus $\frac{\sqrt{\sum_{i=1}^n X_i^2}}{|S_n|} \to 0$

This proves the claim. Now fix k. We have

$$\begin{aligned}\{\frac{\sqrt{\sum_{i=1}^n X_i^2}}{S_n} \to 0\} &= \{\frac{\sqrt{\sum_{i=k+1}^n X_i^2}}{S_n - S_k} \to 0\} \\ &= \cap_{\epsilon \in \mathbb{Q}_+} \cup_{N \geq k} \cap_{n \geq N} \{\sum_{i=n+1}^{\infty} X_i^2 \leq \epsilon(\sum_{i=n+1}^{\infty} X_i^2)\}\end{aligned}$$

where for all $n \geq N \geq k$, $\{\sum_{i=n+1}^{\infty} X_i^2 \leq \epsilon(\sum_{i=n+1}^{\infty} X_i^2)\} \in \mathcal{G}_n \subseteq \mathcal{G}_k$. Therefore $\{\frac{\sqrt{\sum_{i=1}^n X_i^2}}{S_n} \to 0\} \in \mathcal{G}_k$. Since this holds for all $k \in \mathbb{N}$, it follows that it is also in $\mathcal{T}$.

(7.4) Let $S_n = \sum_{i=1}^n X_i$, where (X_i) are i.i.d. with exponential(1) distribution. Use the large deviation theorem to get explicit limits for

- $\frac{\log P(\frac{S_n}{n} > a)}{n}$, $a > 1$

  Consider the function, defined for $\theta \in (0,1)$, $\phi(\theta) = \mathrm{E}(e^{\theta X_i}) = \int_{\mathbb{R}_+} e^{\theta x} e^{-x} dx = \int_{\mathbb{R}_+} e^{x(\theta-1)} dx = \frac{1}{1-\theta}$. $\phi(\theta) > 0$ and finite for all $\theta \in (0,1)$. Then since $a > \mathrm{E}[X_i] = 1$, we have by the theorem of Large Deviations

  $$\lim_{n\to\infty} \frac{1}{n} \log P(\frac{S_n}{n} > a) = \inf_{\theta \in (0,1)} \log \phi(\theta) - \theta a$$

  let $f_a(\theta) = \log \phi(\theta) - \theta a = -\log(1-\theta) - \theta a$. Then $f_a'(\theta) = \frac{1}{1-\theta} - a$, and the derivative is zero when $\theta = 1 - \frac{1}{a} \in (0,1)$. The limit is then given by

  $$f_a(1 - \frac{1}{a}) = -\log(1 - (1 - \frac{1}{a})) - (1 - \frac{1}{a})a = \log a + 1 - a$$

- $\frac{\log P(\frac{S_n}{n} \leq a)}{n}$, $a < 1$ We start by writing

  $$P(\frac{S_n}{n} \leq a) = P(-\frac{S_n}{n} \geq -a)$$

  applying the theorem of large deviations to the sum of variables $-X_i$, we have (since $-a > -1 = \mathrm{E}[-X_i]$)

  $$\lim_{n\to\infty} \frac{1}{n} \log P(\frac{S_n}{n} \leq a) = \inf_{\theta > 0} \log \phi(\theta) - (-a)\theta$$

  where

  $$\phi(\theta) = \mathrm{E}[e^{\theta(-X)}] = \int_{x \in \mathbb{R}} e^{-\theta x} e^{-x} dx = \int_{\mathbb{R}_+} e^{-x(1+\theta)} dx = \frac{1}{1+\theta}$$

  which is finite for all $\theta > 0$. Let $f_a(\theta) = \log \phi(\theta) - (-a)\theta = -\log(1+\theta) + a\theta$, then

  $$f_a'(\theta) = \frac{-1}{1+\theta} + a$$

  and f_a is minimal at $\theta = \frac{1}{a} - 1 > 0$. The limit is then

  $$f_a(\frac{1}{a} - 1) = -\log(1 + \frac{1}{a} - 1) + a(\frac{1}{a} - 1) = \log a + 1 - a$$

(7.5. Oriented first passage percolation) Consider the lattice quadrant $\{(i,j) : i,j \geq 0\}$ with directed edges $(i,j) \to (i+1,j)$ and $(i,j) \to (i,j+1)$. Associate to each edge e an exponential(1) r.v. X_e, independent for different edges. For each directed path π of length d, started at (00), let $S_\pi = \sum_{e\in\pi} X_e$. Let H_d be the minimum of S_π over all such paths π of length d. It can be shown that $\frac{H_d}{d} \to c$ a.s., for some constant c. Given explicit upper and lower bounds on c. (use result of previous question for lower bound).

Lower bound Let Π_d be the set of paths starting at $(0,0)$ of length d. Then we have $H_d = \min_{\pi\in\Pi_d} S_\pi$. Let $a < 1$. The event $H_d \leq a$ is equivalent to the event "there exists $\pi \in \Pi_d$ such that $S_\pi \leq a$"

$$\begin{aligned} P(\frac{H_d}{d} \leq a) &= P(\cup_{\pi\in\Pi_d}\{S_\pi \leq a\}) \\ &\leq \sum_{\pi\in\Pi_d} P(S_\pi \leq a) \\ &= 2^d P(S_\pi \leq a) \end{aligned}$$

since the path weights S_π are identically distributed for $\pi \in \Pi_d$ (sum of d i.i.d. edge weights), and there are 2^d paths in Π_d. Now taking the logarithm, we have

$$\frac{1}{d}\log P(\frac{H_d}{d} \leq a) \leq \log 2 + \frac{1}{d}\log P(S_\pi \leq a)$$

where $S_\pi = \sum_{e\in\pi} X_e$ is the sum of i.i.d. Exponential(1) variables. By the previous problem, since $a < 1$, we have

$$\lim_{d\to\infty} \frac{1}{d}\log P(S_\pi \leq a) = -a + 1 + \log a$$

Now consider the function

$$\begin{aligned} g : (0,1) &\to \mathbb{R} \\ a &\mapsto g(a) = \log 2 - a + 1 + \log a \end{aligned}$$

We have $\lim_{a\uparrow 1} g(a) = \log 2 > 0$, and $\lim_{a\downarrow 0} g(a) = -\infty$, therefore $g(a) < 0$ for some $a \in (0,1)$.

Now consider any a such that $g(a) < 0$. Then

$$\limsup_d \frac{1}{d}\log P(\frac{H_d}{d} \leq a) \leq g(a) < 0 \tag{2}$$

and we must have $c \geq a$, otherwise, we would have $c < a$ and

$$P(\frac{H_d}{d} > a) \leq P(\frac{H_d}{d} > c + \epsilon) \text{ for some small } \epsilon > 0$$

which converges to 0 as $d \to \infty$, since $\frac{H_d}{d}$ converges to c almost surely, therefore in probability. Therefore $\lim_{d\to\infty} P(\frac{H_d}{d} > a) = 0$, or

$$\lim_{d\to\infty} P(\frac{H_d}{d} \leq a) = 1$$

then

$$\frac{1}{d}\log P(\frac{H_d}{d} \leq a) \leq \log P(\frac{H_d}{d} \leq a) \to 0$$

which contradicts (2). Therefore $c \geq a$.

Upper bound We use the following notation: for any node v and any length d,

- Let $\Pi_{v,d}$ be the set of paths of length d that start at v.
- Let $\pi^*_{v,d} = \arg\min_{\pi \in \Pi_{v,d}} S_\pi$, i.e. the path with minimal weight among paths of length d which start at V (to break ties, we pick the minimal path in the lexicographic order). $\pi^*_{v,d}$ is a random variable.
- Let V_d be the (random) end node of the path $\pi^*_{0,d}$.
- Let $H_{v,d} = S_{\pi^*_{v,d}}$, i.e. the minimum weight of paths starting at v and of length d.

Then we have for all d

$$H_{0,d} \leq H_{0,d-1} + H_{V_{d-1},1}$$

Indeed, $H_{V_{d-1},1}$ is by definition the cost of $\pi^*_{V_{d-1},1}$, and $H_{0,d-1}$ is the cost of path $\pi^*_{0,d-1}$. These two paths share the end node V_{d-1}, thus the sum is the cost of a path that starts at 0 and of length d. Therefore the sum is greater than the minimum weight $H_{0,d}$.

By induction, we have

$$H_{0,d} \leq \sum_{i=0}^{d-1} H_{V_i,1}$$

using the convention $V_0 = 0$.

Now $H(V_i, 1)$ is the minimum weight of a path which starts at V_i and of length one. There are two such paths, let their weights be $X_{V_i\rightarrow}$ (the edge which moves right) and $X_{V_i\uparrow}$ (the edge which moves up). Both are Exponential(1) variables. Therefore $H(V_i, 1)$ is the minimum of two Exponential(1) distributions, and is Exponential(2). Indeed, its cumulative function is

$$\begin{aligned} 1 - P(H(V_i,1) > a) &= 1 - P(\min(X_{V_i\rightarrow}, X_{V_i\uparrow}) > a) \\ &= 1 - P(X_{V_i\rightarrow} > a) P(X_{V_i\uparrow} > a) && \text{by independence} \\ &= 1 - e^{-a}e^{-a} \\ &= 1 - e^{-2a} \end{aligned}$$

which is the cumulative function of Exponential(2) distribution. It expectation is then

$$\mathrm{E}\, H_{V_i,1} = \frac{1}{2}$$

therefore the upper bound is the sum of d i.i.d. random variables Exponential(2). By the SLLN, we have $\frac{1}{d}\sum_{i=0}^{d-1} H_{V_i,1}$ converges to $\frac{1}{2}$ almost surely, therefore

$$\lim \frac{H_{0,d}}{d} \leq \frac{1}{2}$$

this gives us an upper bound on c.