STATISTICS 205A - FALL 2014 FINAL

Hoang Duong

December 8, 2014

Lemma 1. *For $p_i \in [0,1], \forall i \in \mathbb{N}$, prove that:*

$$\prod_{m=1}^{\infty} (1 - p_m) = 0 \Leftrightarrow \sum_{m=1}^{\infty} p_m = \infty$$

Proof. As proved in Homework 11, problem 2.
`https://www.dropbox.com/sh/lpohr53eycs7ayo/AACfCHL-bvuW37iROcFLLxTsa/SolHW11.pdf?dl=0` □

Lemma 2. *For a sequence of sets $A_n, n \in \mathbb{N}$, and set A, we have*

$$C \coloneqq \lim_{m \to \infty} [B_m \cap A] = [\lim B_m] \cap A \coloneqq D$$

Proof. It is an obvious lemma. Let $\omega \in C$ then ω is in infinitely many $B_m \cap A$. So $\omega \in A$, and ω is in infinitely many B_m. Thus $\omega \in D$.

On the other hand, if $\omega \in C, \Rightarrow \omega \in A$, and $\omega$is in infinitely many B_m. Thus ω is in infinitely many $B_m \cap A$. So $\omega \in C$. □

Problem 1. Let $(A_n, n \geq 1)$ be a sequence of events. Prove that $\mathbb{P}[A_n i.o.] = 1$ iff:

$$\sum_n \mathbb{P}[A \cap A_n] = \infty, \forall A : \mathbb{P}[A] > 0$$

Proof. Firstly, we will prove the " $\Leftarrow$ " direction. Fix $m \in \mathbb{N}$. We observe that:

$$\sum_{n=1}^{\infty} \mathbb{P}[A \cap A_n] = \infty, \forall A : \mathbb{P}[A] > 0 \tag{1}$$

$$\Leftrightarrow \sum_{n=m}^{\infty} \mathbb{P}[A \cap A_n] = \infty, \forall A : \mathbb{P}[A] > 0 \tag{2}$$

$$\Leftrightarrow \sum_{n=m}^{\infty} \mathbb{P}[A_n \mid A] \mathbb{P}[A] = \infty, \forall A : \mathbb{P}[A] > 0 \tag{3}$$

$$\Leftrightarrow \sum_{n=m}^{\infty} \mathbb{P}[A_n \mid A] = \infty, \forall A : \mathbb{P}[A] > 0 \tag{4}$$

$$\Leftrightarrow \prod_{n=m}^{\infty} (1 - \mathbb{P}[A_n \mid A]) = 0, \forall A : \mathbb{P}[A] > 0 \tag{5}$$

$$\Leftrightarrow \prod_{n=m}^{\infty} \mathbb{P}[A_n^c \mid A] = 0, \forall A : \mathbb{P}[A] > 0 \tag{6}$$

For (2) $\Leftrightarrow$ (3) because of Bayes's theorem for $\mathbb{P}[A] > 0$.
(4) $\Leftrightarrow$ (5) because of Lemma 1.
On the other hand we have:

$$\mathbb{P}\left[\bigcup_{n=m}^{\infty} A_n\right] =1 \tag{7}$$

$$\Leftrightarrow \mathbb{P}\left[\bigcap_{n=m}^{\infty} A_n^c\right] =0 \tag{8}$$

Now by contradiction, assuming that $\mathbb{P}\left[\bigcap_{n=m}^{\infty} A_n^c\right] > 0$. Let

$$\begin{aligned}
A^* &= \bigcap_{n=m}^{\infty} A_n^c \\
\Rightarrow A^* &\subset A_n^c, \forall n \geq m \\
\Rightarrow \mathbb{P}\left[A_n^c \mid A^*\right] &= \frac{\mathbb{P}\left[A_n^c \cap A^*\right]}{\mathbb{P}\left[A^*\right]} = \frac{\mathbb{P}\left[A^*\right]}{\mathbb{P}\left[A^*\right]} = 1, \forall n \geq m \\
\Rightarrow \prod_n \mathbb{P}\left[A_n^c \mid A^*\right] &= 1, \forall n \geq m \\
\Rightarrow \prod_{n=m}^{\infty} \mathbb{P}\left[A_n^c \mid A^*\right] &= 1
\end{aligned}$$

, which is a contradiction with (6). Thus we have:

$$\mathbb{P}\left[\bigcap_{n=m}^{\infty} A_n^c\right] =0$$

,for $m \in \mathbb{N}$ arbitrary. So it is true for all $m \in \mathbb{N}$. As such:

$$\mathbb{P}\left[\lim_{m\to\infty} \bigcup_{n=m}^{\infty} A_m\right] =1$$

Secondly, we will proceed to proving the " $\Rightarrow$ " direction also by contradiction. Assuming the opposite that $\exists A : \mathbb{P}[A] > 0$, such that:

$$\sum_n \mathbb{P}\left[A \cap A_n\right] <\infty$$

Then, by the Borel-Centelli Lemma 1, we have:

$$\mathbb{P}\left[\lim_{m\to\infty} \bigcup_{n=m}^{\infty} \left(A_n \bigcap A\right)\right] =0 \tag{9}$$

$$\Rightarrow \mathbb{P}\left[\lim_{m\to\infty} \left(\left(\bigcup_{n=m}^{\infty} A_m\right) \bigcap A\right)\right] =0 \tag{10}$$

$$\Rightarrow \mathbb{P}\left[\left(\lim_{m\to\infty} \left(\bigcup_{n=m}^{\infty} A_m\right)\right) \bigcap A\right] =0 \tag{11}$$

For (9) $\Leftrightarrow$ (10) because $(B \cap A) \cup (C \cap A) = (B \cup C) \cap A$ (and use induction we get the general case). (10) $\Leftrightarrow$ (11) because of Lemma 2.

Denote:

$$
\begin{aligned}
B = &\left(\lim_{m\to\infty}\left(\bigcup_{n=m}^{\infty} A_m\right)\right)\\
\Rightarrow \mathbb{P}\left[B\right] =&1\\
\Rightarrow \mathbb{P}\left[B^c\right] =&0\\
\Rightarrow \mathbb{P}\left[A\backslash B\right] \leq&\mathbb{P}\left[B^c\right] = 0\\
\Rightarrow \mathbb{P}\left[B\cap A\right] =&\mathbb{P}\left[A\right] - \mathbb{P}\left[A\backslash B\right] = \mathbb{P}\left[A\right] > 0
\end{aligned}
$$

, which is a contradiction with (11). So we also have the proof of the " $\Rightarrow$ " direction.

$\square$

Problem 2. Let X_1, X_2, X_3 be i.i.d taking values in a finite set, and not constant. Is it necessarily true that $\mathbb{P}[X_3 = X_2 \mid X_2 \neq X_1] \leq \mathbb{P}[X_3 = X_2]$? Give proof or a counter-example.

Proof. We will prove that the inequality is true. Assuming that X_i takes values in the finite set $\{a_1, a_2, ..., a_n\}$ with corresponding probability $\{p_1, p_2, ..., p_n\}$, for $p_1 + p_2 + ... + p_n = 1$.We have:

$$
\begin{aligned}
\mathbb{P}[X_3 = X_2 \mid X_2 \neq X_1] =& \frac{\mathbb{P}[X_3 = X_2 \wedge X_2 \neq X_1]}{\mathbb{P}[X_2 \neq X_1]} \\
\mathbb{P}[X_3 = X_2 \wedge X_2 \neq X_1] =& p_1^2(1-p_1) + p_2^2(1-p_2) + ... + p_n^2(1-p_n) \\
\mathbb{P}[X_3 = X_2] =& p_1^2 + p_2^2 + ... + p_n^2 \\
\mathbb{P}[X_2 \neq X_1] =& 1 - \mathbb{P}[X_2 = X_1] \\
=& 1 - p_1^2 - p_2^2 - ... - p_n^2
\end{aligned}
$$

Thus the inequality we need to prove is equivalent to:

$$
\begin{aligned}
\mathbb{P}[X_3 = X_2 \wedge X_2 \neq X_1] \leq& \mathbb{P}[X_3 = X_2]\,\mathbb{P}[X_2 \neq X_1] \\
\Leftrightarrow \sum_{i=1}^{n} p_i^2 - \sum_{i=1}^{n} p_i^3 \leq& \left(\sum_{i=1}^{n} p_i^2\right)\left(1 - \sum_{i=1}^{n} p_i^2\right) \\
\Leftrightarrow p_1^3 + p_2^3 + ... + p_n^3 \geq& \left(p_1^2 + p_2^2 + ... + p_n^2\right)^2 \\
\Leftrightarrow (p_1 + p_2 + ... + p_n)(p_1^3 + p_2^3 + ... + p_n^3) \geq& \left(p_1^2 + p_2^2 + ... + p_n^2\right)^2
\end{aligned}
$$

The last inequality is true according to Cauchy-Schwarz inequality applied to two sequences: $\left(\sqrt{p_i}\right)_{i=1}^{n}$ and $\left(\sqrt{p_i^3}\right)_{i=1}^{n}$.

The equality holds iff $p_1 = p_2 = ... = p_n = 1/n$.

$\square$

Problem 3. Let $(X_i, i \geq 1)$ be i.i.d but not necessarily integrable. Let $S_n = \sum_{i=1}^n X_i$.

(i) Prove that $\limsup_n S_n = \infty$ a.s. iff $\exists$ a stopping time $T < \infty$ a.s. such that $\mathbb{E}S_T > 0$.

(ii) Now assume $\mathbb{E}X_1^+ = \infty$. Prove that $\limsup_n n^{-1} S_n = \infty$ a.s. iff $\exists$ a stopping time $T < \infty$ a.s. such that $\mathbb{E}S_T > \infty$.

Proof. (i) " $\Rightarrow$ " direction. If $\limsup_n S_n = \infty$ a.s., then choose $T = \inf \{n \mid S_n > 1\}$, then $T < \infty$ a.s., and $S_T > 0, \forall \omega$, thus $\mathbb{E}S_T > 0$.

" $\Leftarrow$ " direction. Assuming that $\exists T$ a stopping time, $T < \infty$ a.s., and $\mathbb{E}S_T > 0$. We need to prove $\limsup_n S_n = \infty$ a.s.

According to Theorem 4.1.2 in Durrett, for S_n a random walk on $\mathbb{R}$, there are only four possibilities, one of which has probability one:

(a) $S_n = 0, \forall n$.

(b) $S_n \to \infty$.

(c) $S_n \to -\infty$

(d) $-\infty = \liminf S_n < \limsup S_n = \infty$

Applying this theorem to our problem. First we note that case (a) can not happen, because otherwise $\lim S_T = 0$ for any stopping time, which is a contradiction. If case (b) or case (d) happens with probability one, then we have the proof.

We only need to deal with case (c), when $S_n \to -\infty$ a.s. Thus $\lim \mathbb{E}S_n = -\infty$. According to Theorem 4.1.3 in Durrett, for T is a stopping time with $\mathbb{P}[T < \infty] = 1$, condition on $\{T < \infty\}, \{X_{T+n}, n \geq 1\}$ is independent of $\mathcal{F}_T$ and has the same distribution as the original sequence. Now since for our case $T < \infty$ a.s., conditioning on this event is the same as no conditioning.

Let $\mathbb{E}S_T = \mu > 0$, and $M \in \mathbb{N}$ be arbitrary. Then $\exists k \in \mathbb{N}$ such that $k\mu > M$. Let $T_1, T_2, ..., T_k$ be independent copy of T, then applying Theorem 4.1.3 mentioned above, we have : $\mathbb{E}S_{T_1+T_2+...+T_k} = k\mu > M$. This contradicts with the fact that $\lim \mathbb{E}S_n = -\infty$. So case (c) cannot happen. As such, we complete the proof for part (i)

(ii) " $\Rightarrow$ " direction. If $\limsup n^{-1}S_n = \infty$ a.s., then choose $T = \inf \left\{n \mid n^{-1}S_n > 1\right\}$, then $T < \infty$ a.s. We also have $S_T > T > 0$, thus $\mathbb{E}S_T > -\infty$.

" $\Leftarrow$ " direction. First, if $\mathbb{E}X^- < \infty$, then according to Theorem 2.4.5 in Durrett, we have $S_n/n \to \infty$ a.s. As such, $\limsup n^{-1}S_n \to \infty$ a.s.

We consider the harder case when $\mathbb{E}X^- = \infty$.

We want to apply Theorem 2.5.9 and 2.3.7 in Durrett for $a_n = n$. We check the conditions: $\mathbb{E}\left|X_1\right| \geq \mathbb{E}\left|X^+\right| = \infty$.

$$\begin{aligned}\sum_n \mathbb{P}\left[\left|X_1\right| \geq n\right] \geq &\sum_n \mathbb{P}\left[X_1^+ \geq n\right] \\ \geq &\int_0^\infty \mathbb{P}\left[X_1^+ > x\right] dx \\ = &\mathbb{E}X_1^+ = \infty\end{aligned}$$

Thus we have: $\limsup_n \left|S_n\right|/n = \infty$. So $\limsup_n S_n/n$ is either ∞ or $-\infty$. Assuming that $\limsup_n S_n/n = -\infty$, then $\liminf_n S_n/n \leq \limsup_n S_n/n = -\infty$. Thus $\lim S_n/n = -\infty$, so $\mathbb{E}(S_n/n) = -\infty$. But again using a similar argument as in part (i), let $\mathbb{E}S_T = \mu > -\infty$. Then $\forall N \in \mathbb{N}$, exists $M > N$ of the form $T_1 + T_2 + ... + T_k$ such that $\mathbb{E}S_M = k\mu \Rightarrow \mathbb{E}S_M/M = k\mu/M \geq \mu > -\infty$, which is a contradiction. Thus $\limsup S_n/n = \infty$.

□

Problem 4. Let μ be a probability measure on $[1/4, 3/4]$ with mean $1/2$. Describe a joint distribution for random variable (X, Y) such that:

(i) X has distribution μ

(ii) Y has the uniform distribution on $[0, 1]$

(iii) $\mathbb{E}[Y \mid X] = X$

Proof. Denote $S_x = [1/4, 3/4], \mathcal{S}_x = \mathcal{B}(S_x), S_y = [0, 1], \mathcal{S}_y = \mathcal{B}(S_y)$.

Let $Q(x, B)$ be the conditional distribution of Y given X. By Proposition 5 in the note of conditional distribution, we have the joint distribution ν satisfy:

$$\nu(A \times B) = \int_A Q(x, B)\mu(dx), \forall A \in \mathcal{S}_x, B \in \mathcal{S}_y$$

This joint distribution satisfies:

$$\lambda(B) = \nu(S_x \times B) = \int_{S_x} Q(x, B)\mu(dx), \forall B \in \mathcal{S}_y$$

for λ denotes the Lebesgue measure, and:

$$\int_{S_y} yQ(X(\omega), dy) = X(\omega)$$

$\square$

Proposition 1. *Proposition 5.27 page 96 Breiman (1992).*

Let $X_1, X_2, ...$ *be a submartingale,* $a > 0, T = \inf\{n : X_n \geq a\}$*. If* $\mathbb{E}\left[\sup_n (X_{n+1} - X_n)^+\right] < \infty$*, then for* $X_{n\wedge T}$*,*

$$\limsup \mathbb{E}|X_{n\wedge T}| < \infty$$

Proof. For any n, $X_T^+ \leq a + U$, where $U = \sup_n (X_{n+1} - X_n)^+$.

By Theorem 5.2.9 in Durrett (Optional Stopping Theorem), we have:

$$\begin{aligned}
\mathbb{E}X_{n\wedge T} &\geq \mathbb{E}X_{1\wedge T} = \mathbb{E}X_1 \\
\Rightarrow \mathbb{E}X_{n\wedge T}^- &\leq \mathbb{E}X_{n\wedge T}^+ - \mathbb{E}X_1 \\
\Rightarrow \mathbb{E}|X_{n\wedge T}| &\leq 2\mathbb{E}X_{n\wedge T}^+ - \mathbb{E}X_1 \leq 2a + 2\mathbb{E}U - \mathbb{E}X_1 \\
\Rightarrow \limsup \mathbb{E}|X_{n\wedge T}| &< \infty
\end{aligned}$$

$\square$

Theorem 1. *(Theorem 5.28 page 96 Breiman (1992))*

Let $\{(X_n, \mathcal{F}_n)\}_{n\geq 1}$ *be a martingale such that*

$$\mathbb{E}\left[\sup_{n\geq 1}|X_{n+1} - X_n|\right] < \infty$$

If

$$\begin{aligned}
A_1 &= \{\omega : \lim X_n \text{ exists and is finite}\} \\
A_2 &= \{\omega : \limsup X_n = \infty, \liminf X_n = -\infty\}
\end{aligned}$$

then $\mathbb{P}[A_1 \cup A_2] = 1$ *a.s.*

Proof. Consider $T = \inf\{n : X_n \geq K\}$, then by Proposition 1, and Martingale Convergence Theorem, we have $X_{n\wedge T} \to X$ a.s. On the set $F_K = \{\omega : \sup_n X_n < K\}, X_n = X_{n\wedge T}, \forall n$. Hence on F_K, $\lim_n X_n$ exists and is finite a.s. Thus this limit exists and is finite a.s. on the set $\bigcup_{K=1}^{\infty} F_K$, but this set is exactly the set $\{\limsup X_n < \infty\}$.

Similarly, using the MG sequence $-X_1, -X_2, ...$, we conclude that $\lim_n X_n$ exists and is finite a.s. on the set $\{\liminf X_n > -\infty\}$. Hence $\lim X_n$ exists and is finite for almost all ω in the set $\{\liminf X_n > -\infty\} \cup \{\limsup X_n < \infty\}$. Thus the theorem is proved. $\square$

Problem 5. Let $(X_n, n \geq 0)$ be a martingale w.r.t. a filtration $(\mathcal{F}_n)$. Write $\Delta_n = X_n - X_{n-1}$. Suppose $\mathbb{E}\sup_{n\geq 1}|\Delta_n| < \infty$. Consider the events

$$\begin{aligned}
A &:= \left\{\sum_n \Delta_n^2 < \infty\right\} \\
B &:= \{X_n \text{ converges to a finite limit}\}
\end{aligned}$$

Prove that $A \subseteq B$ a.s.

Proof. From Theorem 1, with similar notation A_1, A_2, we have $\mathbb{P}[A_1 \cup A_2] = 1$. We just need to prove that A_2 happens with probability 0.

$\square$

Problem 6. Let $(\xi_m, 1 \leq m < \infty)$ be i.i.d with exponential(1) distribution. Consider the "alternating signs random walk"

$$S_n = \sum_{m=1}^{n} (-1)^{m-1} \xi_m$$

Let $T^* = \inf\{n \geq 1 : S_n < 0\}$ and for $x > 0$, let $U_x = \inf\{n : S_n \geq x\}$

(a) Show that

$$\mathbb{E} S_{\min(U_x+1, T^*)} = 0$$

(b) Find the distribution of $M \coloneqq \sup_{n \leq T^*} S_n$.

[Hint: for (b) use the memoryless property of the exponential to analyze overshoots.]

Proof. (a) Define a sequence $(e_n)_{n=1}^{\infty}$ as $e_{2k} = 0, \forall k \in \mathbb{N}, e_{2k+1} = 1, \forall k \in \mathbb{N}$, then it is obvious that $S_n - e_n = \sum_{m=1}^{n} (-1)(\xi_m - 1)$ is a martingale since $\xi_m - 1$ are i.i.d with mean zero. We have

$$U_x + 1 = \inf\{n : S_{n-1} \geq x\}$$

is a stopping time. T^* is a stopping time. Thus $N = \min\{U_x + 1, T^*\}$ is also a stopping time.

First we will prove that $N < \infty$. Indeed, consider the even subsequence of S_n, which is $S_{2k} = (\xi_1 - \xi_2) + (\xi_3 - \xi_4) + ... + (\xi_{2k-1} - \xi_{2k})$ is a random walk (of Laplace random variable). Thus by Theorem 4.1.2 in Durrett, either $\liminf S_{2k} = -\infty$ or $\limsup S_{2k} = \infty$ with probability 1 (since we can rule out the case of $S_{2k} = 0, \forall k$). So $N < \infty$.

Second, we will prove that $e_N = 1$. Indeed, observe that T^* must be even because if it was even, then S_{T^*-1} is smaller, as such $S_{T^*-1} < 0$, which contradicts with the definition of T^*. Similarly, U_x must be odd, since if it was even, S_{U_x-1} is bigger, as such $S_{U_x-1} > x$, which contradicts with the definition of U_x. As such $U_x + 1$ is even. So N is even, which implies $e_N = 0$.

Now applying the Optional Stopping Theorem to the martingale $S_n - e_n$ and the finite stopping time N, we have:

$$\begin{aligned}
\mathbb{E}[S_0 - e_0] =& \mathbb{E}[S_N - e_N] \\
\Rightarrow 0 =& \mathbb{E} S_N - 0 \\
\mathbb{E} S_N =& 0
\end{aligned}$$

(b) Since $S_1 \geq 0$, we have M is a non-negative random variable. We have:

$$\begin{aligned}
\mathbb{P}[M < x] =& \mathbb{P}\left[\sup_{n \leq T^*} S_n < x\right] \\
=& \mathbb{P}[T^* + 1 \leq U_x] \coloneqq p
\end{aligned}$$

From (a) we have:

$$\begin{aligned}
0 =& \mathbb{P}[T^* < U_x + 1] \, \mathbb{E}[S_{T^*} - e_{T^*} \mid T^* < U_x + 1] + \\
&+ \mathbb{P}[T^* \geq U_x + 1] \, \mathbb{E}[S_{U_x+1} - e_{U_x+1} \mid T^* \geq U_x + 1] \\
=& \mathbb{P}[T^* < U_x + 1] \, \mathbb{E}[S_{T^*} \mid T^* \leq U_x - 1] + \\
&+ (1 - \mathbb{P}[T^* < U_x + 1]) \, \mathbb{E}[S_{U_x+1} \mid T^* \geq U_x + 1]
\end{aligned}$$

And since T^* and $U_x + 1$ are both even, $T^* < U_x + 1 \Leftrightarrow T^* \leq U_x - 1$. So $p = \mathbb{P}[T^* < U_x + 1]$. So we have:

$$p = \frac{\mathbb{E}\left[S_{U_x+1} \mid T^* \geq U_x + 1\right]}{\mathbb{E}\left[S_{U_x+1} \mid T^* \geq U_x + 1\right] + \mathbb{E}\left[S_{T^*} \mid T^* \leq U_x - 1\right]}$$

Now consider:

$$\mathbb{E}\left[S_{U_x+1} \mid T^* \geq U_x + 1\right] =$$

□

Problem 7. (a) Consider events (A_n) adapted to a filtration $(\mathcal{F}_n)$. Suppose $\mathbb{P}[A_n \mid \mathcal{F}_{n-1}] \geq 0.6$ for all $n \geq 1$. Let $M_k = \sum_{i=1}^{k} \mathbb{I}_{A_i}$. Find a constant C such that

$$\mathbb{P}[M_k \leq 0.55k - C \text{ for some } k \geq 0] \leq 0.05$$

(b) You are planning a tennis tournament with n players. When player i plays against player $j \neq i$, player i will win with some probability $p_{ij} = 1 - p_{ji}$, independent of other matches. The number (p_{ij}) are unknown to you, but suppose there is some "best" player i (you don't know which player) such that $p_{ij} \geq 0.6$ for all $j \neq i$. You want an algorithm for scheduling matches in such a way that after a deterministic number t_n of matches you can announce a winner and know that with chance $\geq 95\%$ the announced winner is the best player. Describe how to do this using $t_n \leq Bn$ matches, for some constant B not depending on n.

Proof. We have:

$$\begin{aligned} \mathbb{P}[M_k \leq 0.55k - C \text{ for some } k \geq 0] &\leq 0.05 \\ \Leftrightarrow \mathbb{P}[M_k > 0.55k - C, \forall k \geq 0] &\geq 0.95 \end{aligned}$$

From the Conditional Borel-Cantelli Lemma, we have: $\lim M_k = \infty$, as $\sum_n \mathbb{P}[A_n \mid \mathcal{F}_{n-1}] = \infty$. Moreover, we have:

$$\frac{M_k}{\sum \mathbb{P}[A_n \mid \mathcal{F}_{n-1}]}$$

converges to 1 a.s., as a result of Dubins and Freedman (1965).

Reference:

Durrett, R (2013). Probability Theory.

Breiman, L (1992). Probability.

Dubins, L.E., Freedman, D.A., (1965). A Sharper Form of The Borel-Cantelli Lemma And The Strong Law ☐