Probabilities and statistics

Lecturer: Prof. Dr. Sara van de Geer
Prof. Dr. Martin Larsson

ETH Zürich

Midterm Exam: Solution

April 14th, 2015

Q1. Let $(S_k)_{k=0}^{2n}$ be a simple random walk. We want to calculate $\mathbb{P}(S_{2n} = -2k = \min_{i=0,\dots,2n} S_i)$.

(a) (2 points) For $-n \leq k \leq n$, compute $\mathbb{P}(S_{2n} = 2k)$. What is the value of $\mathbb{P}(S_{2n}$ is odd)?
Hint: Recall the formula of $\mathbb{P}(S_{2n} = 2k - 2n)$.

(b) (4 points) Define $T_{-1} = \inf\{k \in \mathbb{N} : S_k = -1\}$. Use the reflection principle to prove that, for $k \geq 0$:

$$\mathbb{P}\left(S_{2n} = 2k, T_{-1} > 2n\right) = 2^{-2n}\left(\binom{2n}{k+n} - \binom{2n}{k+n+1}\mathbf{1}_{\{k<n\}}\right).$$

(c) (4 points) Prove that

$$\mathbb{P}(S_{2n} = -2k = \min_{i=0,\dots,2n} S_i) = \mathbb{P}(S_{2n} = 2k, T_{-1} > 2n) \quad \forall k \geq 0.$$

Hint: It may be useful to make a Bijection between well-chosen sets.

(d) (3 points) Use (b) and (c) to compute $\mathbb{P}(S_{2n} = \min_{i=0,\dots,2n} S_i)$.

Solution:

(a) We know from the Skript p.9 that $\mathbb{P}(S_{2n} = 2j - 2n) = \binom{2n}{j}2^{-2n}$, then

$$\mathbb{P}(S_{2n} = 2k) = \binom{2n}{n+k}2^{-2n}.$$

We know that we are summing an even number of 1 and -1 so it's not possible to have an odd number, then $\mathbb{P}(S_{2n} \text{ odd}) = 0$.

(b) Using the reflection principle we have that

$$\begin{aligned}\mathbb{P}(S_{2n} = 2k, T_{-1} > 2n) &= \mathbb{P}(S_{2n} = 2k) - \mathbb{P}(S_{2n} = 2k, T_{-1} < 2n)\\ &= \binom{2n}{k+n}2^{-2n} - \mathbb{P}(S_{2n} = -2k-2)\\ &= 2^{-2n}\left(\binom{2n}{k+n} - \binom{2n}{k+n+1}\mathbf{1}_{\{k<n\}}\right).\end{aligned}$$

(c) Define:

$$\begin{aligned}&\Omega = \{(x_k)_{k=0}^{2n} : x_0 = 0, (\forall k \in \{0,1,..,2n\})|x_k - x_{k+1}| = 1\},\\ &A := \{x \in \Omega : x_{2n} = 2k, x_k \geq 0\},\\ &B := \left\{x \in \Omega : x_{2n} = -2k = \min_{i=0,1,\dots,2n} x_i\right\}.\end{aligned}$$

The probability measure is a Laplace model, so it's enough to show that $|A| = |B|$. Define $f : A \to B$ with $f(x)_k = x_{2n-k} - 2k$. It's clear if $x \in A$, $f(x) \in B$ because $f(x)_0 = 0$, $f(x)_{2n} = -2k$ and due to $x_k \geq 0$, we have that

$$f(x)_k \geq -2k = f(x)_{2n} = \min_{i=0,1,\ldots,2n} x_i.$$

f it's clearly invertible so $|A| = |B|$.

(d) We combine (b) and (c) to get

$$\begin{aligned}\mathbb{P}(S_{2n} = \min_{i=0,\ldots,2n} S_i) &= \sum_{k=0}^{n} \mathbb{P}(S_{2n} = \min_{i=0,\ldots,2n} S_i = -2k) \\ &= \sum_{k=0}^{n} 2^{-2n} \left(\binom{2n}{k+n} - \binom{2n}{k+n+1} \mathbf{1}_{\{k<n\}} \right) \\ &= \binom{2n}{n} 2^{-2n}.\end{aligned}$$

Q2. An interviewer stops a person in the street. He asks some question about his/her family. We assume a Laplace Model on this question.

(a) (3 points) Suppose we know that the person has two children and at least one of them is a boy. What is the probability that the other child is also a boy?
Hint: First write down an appropriate probabilistic model.

(b) (3 points) Suppose we know that he/she has two children which are not twins, and at least one of them is a boy born on Monday. Compute the probability that the other one is a boy. Assume that the day of birth is uniformly distributed over the week and that it is independent of gender.
Hint: First write down an appropriate probabilistic model.

(c) (4 points)Suppose we know that he/she has n children, and at least one of them is a boy. Compute the probability that she/he has at least 2 sons.
Hint: First write down an appropriate probabilistic model.

Solution:

(a) Take $\Omega := \{(0,1)\}^2$ where 0 means that the kid is a boy and 1 that is a girl. Computing

$$\mathbb{P}(\text{Both are male} \mid \text{There is one male}) = \frac{\mathbb{P}(\text{Both are male})}{\mathbb{P}(\text{There is one male})} = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3}.$$

(b) Take $\Omega = (\{0,1\} \times \{1,7\})^2$ where for $((a,b)(c,d)) \in \Omega$ (a,b) represent the first child, a represent his/her sex and b the day of the week he/she was born. Computing

$$\begin{aligned}\mathbb{P}(\text{There is one male born on monday}) &= 1 - \mathbb{P}(\text{There is no male born on monday}) \\ &= 1 - \left(\frac{13}{14}\right)^2 = \frac{27}{196}.\end{aligned}$$

Additionally

$$\begin{aligned}&\mathbb{P}(\text{Both are male and there is one born on Monday})\\ =&\mathbb{P}(\text{Both are male}) - \mathbb{P}(\text{Both are male and there is none born on Monday})\\ =&\frac{1}{4} - \left(\frac{1}{2}\cdot\frac{6}{7}\right)^2 = \frac{13}{49\cdot 4}.\end{aligned}$$

Finally

$$\mathbb{P}(\text{Both are male} \mid \text{There is one male and is born on monday}) = \frac{\frac{13}{196}}{\frac{27}{196}} = \frac{13}{27}.$$

(c) Take $\Omega := \{(0,1)\}^n$ where 0 means that the kid is a boy and 1 that is a girl. Computing

$$\mathbb{P}(\text{There is at least two male}) = 1 - \mathbb{P}(\text{There less than two male}) = 1 - 2^{-n} - n2^{-n}.$$

Additionally

$$\mathbb{P}\left(\text{There is at least one male}\right) = 1 - 2^{-n}.$$

Finally

$$\mathbb{P}\left(\text{There is at least two male} \mid \text{There is at least one male}\right) = \frac{1 - 2^{-n} - n2^{-n}}{1 - 2^{-n}}.$$

Q3. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent normal random variables, with 0 mean and variance 1. Recall that X is a normal random variable with 0 mean and variance 1 if

$$\mathbb{P}(X \geq t) = \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

(a) (2 points) State the Borell-Cantelli Lemma 1 and 2.

For X a standard normal random variable and $t > 0$, assume the following inequalities for the question (b) and (c) (we are going to prove them later):

$$\frac{1}{\sqrt{2\pi}} \frac{t}{t^2+1} e^{\frac{-t^2}{2}} \leq P(X \geq t) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{t} e^{\frac{-t^2}{2}}. \tag{1}$$

(b) (4 points) Prove that $\mathbb{P}\left(\limsup \frac{X_n}{\sqrt{\log(n)}} \leq \sqrt{2}\right) = 1.$

Hint: For $\alpha > \sqrt{2}$, define $A_n^\alpha = \{X_n \geq \alpha\sqrt{\log(n)}\}$ compute $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n^\alpha)$.

(c) (4 points) Prove that $\mathbb{P}\left(\limsup \frac{X_n}{\sqrt{\log(n)}} = \sqrt{2}\right) = 1.$

Now we are interested to prove the inequalities (1).

(d) (3 points)Prove that if $X \sim N(0,1)$ and $t > 0$

$$P(X \geq t) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{t} e^{\frac{-t^2}{2}}.$$

Hint: Note that $\mathbb{P}(X \geq t) \leq \mathbb{E}\left(\frac{X}{t}\mathbf{1}_{\{X\geq t\}}\right)$.

(e) (4 points)Prove that if $X \sim N(0,1)$ and $t > 0$

$$\frac{1}{\sqrt{2\pi}}\frac{t}{t^2+1}e^{\frac{-t^2}{2}} \leq P(X \geq t).$$

Hint: Define $f(t) = P(X \geq t) - \frac{1}{\sqrt{2\pi}}\frac{t}{t^2+1}e^{\frac{-t^2}{2}}$. Prove that $f(t)$ is decreasing and $\lim_{t\to\infty} f(t) = 0$.

Solution

(a) Let $(\Omega, \mathcal{A}.\mathbb{P})$ be a probability space and $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$. Recall that

$$\limsup A_n = \bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n.$$

Then we have that:

i. If $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$ then $\mathbb{P}(\limsup A_n) = 0$.

ii. If $(A_n)_{n\in\mathbb{N}}$ are independent and $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$ then $\mathbb{P}(\limsup A_n) = 1$.

(b) Let us compute for $\alpha > \sqrt{2}$

$$\mathbb{P}(A_n^\alpha) \leq \frac{1}{\sqrt{2a\pi}}\frac{1}{\sqrt{\log(n)}}\frac{1}{n^{\frac{\alpha^2}{2}}}$$
$$\Rightarrow \sum_{n\in\mathbb{N}} \mathbb{P}(A_n) \leq \frac{1}{\alpha\sqrt{2a\pi}}\sum_{n\in\mathbb{N}}\frac{1}{\sqrt{\log(n)}}\frac{1}{n^{\frac{\alpha^2}{2}}} < \infty.$$

Then thanks to Borel-Cantelli's Lemma (i) we have that $\mathbb{P}(\limsup A_n^\alpha) = 0$, then

$$\mathbb{P}\left(\bigcup_{\alpha>\sqrt{2},\alpha\in\mathbb{Q}}\bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n\right) = 0.$$

The it's sufficient to show that:

$$L := \left\{\omega \in \Omega : \limsup_n \frac{X_n(\omega)}{\sqrt{\log n}} > \sqrt{2}\right\} \subseteq \bigcup_{\alpha>\sqrt{2},\alpha\in\mathbb{Q}}\bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n.$$

Take $\omega \in L$, there exists $\sqrt{2} < \alpha \in \mathbb{Q}$ such that $\limsup \frac{X_n(\omega)}{\sqrt{\log n}} > \alpha$, thus, for all $k \in \mathbb{N}$ there exists $n \geq k$ such that $\frac{X_n(\omega)}{\sqrt{\log n}} < \alpha$. Then $\omega \in \bigcup_{\alpha>\sqrt{2},\alpha\in\mathbb{Q}}\bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n$.

(c) It's enough to prove that $\mathbb{P}(\limsup \frac{X_n}{\sqrt{\log n}} \geq \sqrt{2}) = 1$. Let's compute

$$\mathbb{P}(A_n^{\sqrt{2}}) \geq \frac{1}{\sqrt{\pi}}\frac{\sqrt{\log n}}{2\log(n)+1}\frac{1}{n}$$
$$\Rightarrow \sum_{n\in\mathbb{N}} \mathbb{P}(A_n) \geq \frac{1}{\sqrt{\pi}}\frac{\sqrt{\log n}}{2\log(n)+1}\frac{1}{n} = \infty.$$

Then, thanks to Borel-Cantelli's Lemma (ii) we have that $\mathbb{P}(\limsup A_n^{\sqrt{2}}) = 1$. So it's sufficient to show that

$$\bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n \subseteq \left\{\omega\in\Omega : \limsup_n \frac{X_n(\omega)}{\sqrt{\log n}} \geq \sqrt{2}\right\} =: I$$

Let $\omega \in \bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n$, then for all $k\in\mathbb{N}$ there exists $n(\omega)\geq k$ such that $\frac{X_n(\omega)}{\sqrt{\log n}} \geq \sqrt{2}$, so $\limsup \frac{X_n(\omega)}{\sqrt{\log n}} \geq \sqrt{2}$, then $\omega \in L$. We have then that

$$\mathbb{P}\left(\limsup \frac{X_n(\omega)}{\sqrt{\log n}} = \sqrt{2}\right) = \mathbb{P}\left(\limsup \frac{X_n(\omega)}{\sqrt{\log n}} \leq \sqrt{2}, \limsup \frac{X_n(\omega)}{\sqrt{\log n}} \geq \sqrt{2}\right) = 1.$$

(d) Just computing

$$\begin{aligned}\mathbb{P}(X\geq t) &\leq \mathbb{E}\left(\mathbf{1}_{\{X\geq t\}}\frac{X}{t}\right)\\ &= \frac{1}{\sqrt{2\pi}t}\int_t^\infty x e^{\frac{-x^2}{2}}dt\\ &= \frac{1}{\sqrt{2\pi}t}\left(-e^{\frac{-x^2}{2}}\right)\Big|_t^\infty\\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{t}e^{-\frac{t^2}{2}}.\end{aligned}$$

(e) We are going to follow the hint, we just have to prove that $f(t)\geq 0$. It's clear that $\lim_{t\to\infty} f(t) = 0$. Let's prove that f is decreasing, we just have to compute its derivative:

$$\begin{aligned}f'(t) &= \left(\frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-\frac{x^2}{2}}dx - \frac{1}{\sqrt{2\pi}}\frac{t}{t^2+1}e^{-\frac{t^2}{2}}\right)'\\ &= -\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}} + \frac{1}{\sqrt{2\pi}}\frac{t^2}{t^2+1}e^{-\frac{t^2}{2}} - \frac{1}{\sqrt{2\pi}}\frac{1}{t^2+1}e^{-\frac{t^2}{2}} + \frac{1}{\sqrt{2\pi}}\frac{2t^2}{(t^2+1)^2}e^{-\frac{t^2}{2}}\\ &= -\frac{2e^{-\frac{t^2}{2}}}{\sqrt{2\pi}(t^2+1)^2} < 0.\end{aligned}$$

Then f is decreasing so $\lim_{t\to\infty} f(t) = \inf_{t\in(0,\infty)} f(t) = 0$, thus, $f(t)\geq 0$.