Solution Set 14

Problem 1: Convergence and inequalities

Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. non-negative random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and such that there exists $0 < a < b < +\infty$ with $a < X_n(\omega) \leq b$ for all $n \geq 1$ and $\omega \in \Omega$. Let also $(Y_n,\, n \geq 1)$ be the sequence defined as

$$Y_n = \left(\prod_{j=1}^{n} X_j \right)^{1/n}, \quad n \geq 1$$

a) Show that there exists a constant $\mu > 0$ such that $Y_n \underset{n\to\infty}{\to} \mu$ almost surely.

b) Compute the value of μ in the case where $\mathbb{P}(\{X_1 = a\}) = \mathbb{P}(\{X_1 = b\}) = \frac{1}{2}$ and $a, b > 0$.

c) In this case, look for a good upper bound on $\mathbb{P}(\{Y_n > t\})$ for $n \geq 1$ fixed and $t > \mu$.

Solution a) Consider

$$\log(Y_n) = \frac{1}{n} \sum_{j=1}^{n} \log(X_j)$$

As $\log(X_j)$ are i.i.d. bounded random variables, the strong law of large numbers applies, so

$$\frac{1}{n} \sum_{j=1}^{n} \log(X_j) \underset{n\to\infty}{\to} \mathbb{E}(\log(X_1)) \quad \text{almost surely}$$

so $\mu = \exp(\mathbb{E}(\log(X_1)))$.

b) In this case $\mathbb{E}(\log(X_1)) = \frac{\log(a)+\log(b)}{2}$, so $\mu = \sqrt{ab}$.

c) Observing that $\log(X_j) \in [\log(a), \log(b)]$ and using (the generalized version of) Hoeffding's inequality, we obtain

$$\mathbb{P}(\{Y_n \geq t\}) = \mathbb{P}(\{\log(Y_n) - \log(\mu) \geq \log(t) - \log(\mu)\}) \leq \exp\left(-\frac{2n\,(\log(t) - \log(\mu))^2}{(\log(b) - \log(a))^2} \right)$$

so $\mathbb{P}(\{Y_n \geq t\}) \leq C^n$ for every $t > \mu = \sqrt{ab}$, and a possible value for C is $\exp(-\frac{2\,(\log(t)-\log(\mu))^2}{(\log(b)-\log(a))^2})$ (note that the same result may be obtained by a direct computation and the use of the inequality $\cosh(x) \leq \exp(x^2/2)$).

And a weaker result can be obtained also via the inequality

$$\mathbb{P}(\{Y_n \geq t\}) \leq \frac{\mathbb{E}(Y_n^n)}{t^n} = \left(\frac{\mathbb{E}(X_1)}{t} \right)^n = \left(\frac{a+b}{2t} \right)^n$$

which shows only that concentration holds for every $t > \frac{a+b}{2}$ (and not $t > \mu$).

Problem 2: Large deviations principle

Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. $\mathcal{E}(\lambda)$ random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e., X_1 admits the following pdf:

$$p_{X_1}(x) = \begin{cases} \lambda \exp(-\lambda x) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Let also $S_n = X_1 + \ldots + X_n$. Using the large deviations principle, find a tight upper bound on

$$\mathbb{P}\left(\{S_n \geq nt\}\right) \quad \text{for } t > \mathbb{E}(X_1) = \frac{1}{\lambda}$$

Solution We use the large deviations principle to find a tight upper bound. Before this, we need to check that the moment generating function $\mathbb{E}(e^{sX_1})$ is finite in a proper neighborhood of $s = 0$:

$$\mathbb{E}(e^{sX_1}) = \int_0^\infty e^{sx}\, \lambda e^{-\lambda x}\, dx = \frac{\lambda}{\lambda - s}, \quad \text{for } s < \lambda$$

Therefore, by applying the large deviations principle, we obtain for $t > 1/\lambda$:

$$\mathbb{P}(\{S_n > nt\}) \leq \exp(-n\, \Lambda^*(t)) \quad \text{where} \quad \Lambda^*(t) = \max_{s \in \mathbb{R}} \left\{ st - \log\left(\frac{\lambda}{\lambda - s}\right) \right\}$$

By taking the derivative of $st - \log\left(\frac{\lambda}{\lambda - s}\right)$ with respect to s and setting it equal to zero, we obtain that $\Lambda^*(t)$ is maximum at $s^* = \lambda - \frac{1}{t}$. Hence,

$$\mathbb{P}(\{S_n > nt\}) \leq \exp(-n\,(\lambda t - 1 - \log(\lambda t)))$$

Problem 3: Moment generating function

Recall that the moment-generating function of a random variable X is defined for every $t \in \mathbb{R}$ as

$$M_X(t) = \mathbb{E}\left(e^{tX}\right).$$

a) Show that if $X \sim \mathcal{N}(0, \sigma^2)$, then

$$M_X(t) = \exp\left(\frac{1}{2} t^2 \sigma^2\right).$$

We now introduce the concept of *sub-gaussianity.* A random variable X is called sub-gaussian if, for every $t > 0$,

$$M_X(t) \leq \exp\left(\frac{1}{2} t^2 \eta^2\right)$$

for some $\eta \in \mathbb{R}^+$. (Note that η^2 need not be the variance of X!).

b) Show that if $X \sim \mathcal{U}([-a, a])$ for some $a > 0$, then X is sub-gaussian with $\eta = a$.

Hint: Recall that $e^x = \sum_{n=0}^\infty \frac{x^n}{n!}$.

c) Show that if X is sub-gaussian for some $\eta \in \mathbb{R}^+$, then for every $t > 0$,

$$\mathbb{P}(|X| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\eta^2}\right).$$

d) Prove the following generalization of Hoeffding's inequality. Let $X_i, i \in \{1, 2, \ldots, n\}$ be independent random variables, where for each i, $X_i - \mathbb{E}(X_i)$ is sub-gaussian for some $\eta_i \in \mathbb{R}^+$. Let also $S_n = \sum_{i=1}^n X_i$. Show that for every $t > 0$,

$$\mathbb{P}(|S_n - \mathbb{E}(S_n)| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^n \eta_i^2}\right).$$

e) Let $X_i, i \in \{1, 2, \ldots, n\}$ be sub-gaussian random variables with the same $\eta \in \mathbb{R}^+$. Show that

$$\mathbb{E}\left(\max_i X_i\right) \leq \eta\sqrt{2 \ln n}.$$

Hint: Start by rewriting $\mathbb{E}(\max_i X_i) = \frac{1}{t}\mathbb{E}(\ln \exp(t \max_i X_i))$.

Solution

a) For $X \sim \mathcal{N}(0, \sigma^2)$ we have

$$\begin{aligned} M_X(t) = \mathbb{E}(e^{tX}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{tx} e^{-\frac{x^2}{2\sigma^2}}\, dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{t^2\sigma^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\sigma^2 t)^2}{2\sigma^2}}\, dx \\ &= \exp\left(\frac{t^2\sigma^2}{2}\right). \end{aligned}$$

b) For $X \sim \mathcal{U}([-a, a])$ we have

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{-a}^{a} \frac{1}{2a} e^{tx}\, dx = \frac{1}{2at}(e^{ta} - e^{-ta}).$$

Now note that, using the Taylor expansion of e^x given in the hint, we can write

$$\begin{aligned} e^{ta} - e^{-ta} &= \sum_{n=0}^{\infty} \frac{(ta)^n}{n!} - \sum_{n=0}^{\infty} \frac{(-ta)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(ta)^{2n+1}}{(2n+1)!} \\ &\leq ta \sum_{n=0}^{\infty} \frac{(t^2a^2)^n}{2^n n!} \\ &= ta \exp\left(\frac{t^2a^2}{2}\right) \end{aligned}$$

where the inequality is due to the fact that $(2n+1)! \geq 2^n n!$, and the last equality is due to the Taylor expansion of $\exp\left(\frac{t^2a^2}{2}\right)$. Hence, we conclude that

$$M_X(t) \leq \frac{1}{2} \exp\left(\frac{t^2a^2}{2}\right) \leq \exp\left(\frac{t^2a^2}{2}\right).$$

c) By the Chebyshev-Markov inequality with $\psi(x) = e^{sx}$, we have

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(e^{sX})}{e^{st}} \leq \exp\left(\frac{s^2\eta^2}{2} - st\right).$$

The optimal s (which can be found by taking the derivative of the right-hand side and putting it equal to 0) is $s = \frac{t}{\eta^2}$, which we can substitute into the equation to get

$$\mathbb{P}(X \geq t) \leq \exp\left(\frac{t^2}{2\eta^2}\right).$$

The same upper-bound can be obtained similarly for $\mathbb{P}(X \leq -t)$, proving the result.

d) Note that, if Y_1 and Y_2 are two independent sub-gaussian random variables for some η_1 and η_2 , then $Y_1 + Y_2$ is sub-gaussian with $\eta^2 = \eta_1^2 + \eta_2^2$. In fact,

$$M_{Y_1+Y_2}(t) = \mathbb{E}(e^{t(Y_1+Y_2)}) = \mathbb{E}(e^{tY_1})\mathbb{E}(e^{tY_2}) \leq \exp\left(\frac{t^2(\eta_1^2 + \eta_2^2)}{2}\right).$$

One can apply this result recursively to prove the same property for the sum of n independent random variables. Then, the required result follows directly from part 3 with $X = \sum_{i=1}^n (X_i - \mathbb{E}(X_i))$.

e) Using the hint, we have

$$\begin{aligned}
\mathbb{E}\left(\max_i X_i\right) &= \frac{1}{t}\mathbb{E}\left(\ln\exp\left(t\max_i X_i\right)\right) \\
&\leq \frac{1}{t}\ln\mathbb{E}\left(\exp\left(t\max_i X_i\right)\right) \\
&= \frac{1}{t}\ln\mathbb{E}\left(\max_i\exp\left(tX_i\right)\right) \\
&\leq \frac{1}{t}\ln\mathbb{E}\left(\sum_{i=1}^n\exp\left(tX_i\right)\right) \\
&= \frac{1}{t}\ln\left(\sum_{i=1}^n\mathbb{E}(\exp(tX_i))\right) \\
&\leq \frac{\ln n}{t} + \frac{\eta^2 t}{2}
\end{aligned}$$

where the first inequality follows from Jensen's inequality, and the last one is due to the fact that the n random variables are sub-gaussian with the same η . The optimal t (obtained once again by putting the derivative equal to 0) is $t = \frac{\sqrt{2\ln(n)}}{\eta}$. Substituing this value into the last equation gives

$$\mathbb{E}\left(\max_i X_i\right) \leq 2\eta\sqrt{\frac{\ln n}{2}} = \eta\sqrt{2\ln n}.$$

Problem 4: Inequalities

Part I. Let $(X_n, n \geq 1)$ be a sequence of i.i.d. random variables such that $\mathbb{P}(\{X_n = +1\}) = p$ and $\mathbb{P}(\{X_n = -1\}) = 1 - p$ for some fixed $0 < p < 1/2$.

Let $S_0 = 0$ and $S_n = X_1 + \ldots + X_n$, $n \geq 1$. Let also $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, $n \geq 1$.

Preliminary question. Deduce from Hoeffding's inequality that for any $0 < p < 1/2$,

$$\mathbb{P}(\{|S_n - n\,(2p-1)| \geq nt\}) \leq 2\exp\left(-\frac{nt^2}{2}\right) \qquad \forall t > 0,\, n \geq 1.$$

This inequality will be useful at some point in this exercise.

Let now $(Y_n,\, n \in \mathbb{N})$ be the process defined as $Y_n = \lambda^{S_n}$ for some $\lambda > 0$ and $n \in \mathbb{N}$.

a) *Using Jensen's inequality only*, for what values of λ can you conclude that the process Y is a submartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$?

b) Identify now the values of $\lambda > 0$ for which it holds that the process $(Y_n = \lambda^{S_n},\, n \in \mathbb{N})$ is a martingale / submartingale / supermartingale with respect to $(\mathcal{F}_n,\, n \in \mathbb{N})$.

c) Compute $\mathbb{E}(|Y_n|)$ and $\mathbb{E}(Y_n^2)$ for every $n \in \mathbb{N}$ (and every $\lambda > 0$).

d) For what values of $\lambda > 0$ does it hold that $\sup_{n\in\mathbb{N}} \mathbb{E}(|Y_n|) < +\infty$? $\quad \sup_{n\in\mathbb{N}} \mathbb{E}(Y_n^2) < +\infty$?

e) Run the process Y numerically. For what values of $\lambda > 0$ do you observe that there exists a random variable Y_∞ such that $Y_n \underset{n\to\infty}{\to} Y_\infty$ a.s.?

Prove it then theoretically and compute the random variable Y_∞ when it exists (this computation might depend on λ, of course).

f) For what values of $\lambda > 0$ does it hold that $Y_n \overset{L^2}{\underset{n\to\infty}{\to}} Y_\infty$?

g) Finally, for what values of $\lambda > 0$ does it hold that $\mathbb{E}(Y_\infty|\mathcal{F}_n) = Y_n$, $\forall n \in \mathbb{N}$?

Part II. Consider now the (interesting) value λ for which the process Y is a martingale.
(Spoiler: there is a unique such value of λ, and it is greater than 1.)

Let $a \geq 1$ be an integer and consider the stopping time $T_a = \inf\{n \in \mathbb{N} : Y_n \geq \lambda^a \quad \text{or} \quad Y_n \leq \lambda^{-a}\}$.

a) Estimate numerically $\mathbb{P}(\{Y_{T_a} = \lambda^a\})$ for some values of a. Explain your method.

b) Is it true that $\mathbb{E}(Y_{T_a}) = \mathbb{E}(Y_0)$? Justify your answer.

c) If possible, use the previous statement to compute $P = \mathbb{P}(\{Y_{T_a} = \lambda^a\})$ theoretically. How fast does this probability decay with a?

Consider finally the other stopping time $T'_a = \inf\{n \in \mathbb{N} : Y_n \geq \lambda^a\}$.

d) Estimate numerically $\mathbb{P}(\{Y_{T'_a} = \lambda^a\})$ for some values of a. Explain your method.

e) Is it true that $\mathbb{E}(Y_{T'_a}) = \mathbb{E}(Y_0)$? Justify your answer.

f) If possible, use the above statement to compute $P' = \mathbb{P}(\{Y_{T'_a} = \lambda^a\})$ theoretically. Is this probability P' greater or smaller than P?

Solution a) As $0 < p < 1/2$, S is a supermartingale. Also, $\lambda^x = e^{x\log(\lambda)}$ is a convex function $\forall \lambda > 0$, but is increasing for $\lambda \geq 1$ and decreasing for $\lambda \leq 1$. Therefore, applying Jensen's inequality gives

$$\mathbb{E}(\lambda^{S_{n+1}} \,|\, \mathcal{F}_n) \geq \lambda^{\mathbb{E}(S_{n+1} \,|\, \mathcal{F}_n)} \geq \lambda^{S_n}$$

only when $\lambda \leq 1$.

b) We have

$$\mathbb{E}(\lambda^{S_{n+1}} \,|\, \mathcal{F}_n) = \lambda^{S_n}\, \mathbb{E}(\lambda^{X_{n+1}}) = \lambda^{S_n}\, (\lambda p + \frac{1}{\lambda}(1-p))$$

which says that Y is a submartingale if and only if $E = \lambda p + \frac{1}{\lambda}(1-p) \geq 1$. Solving this (quadratic) inequation, we obtain the condition: $\lambda \in]-\infty, 1] \cup [\frac{1-p}{p}, +\infty[$. For the martingale condition to hold ($E = 1$), it must be that either $\lambda = 1$ (trivial case) or $\lambda = \frac{1-p}{p}$. Finally, Y is a supermartingale ($E \leq 1$) if and only if $\lambda \in [1, \frac{1-p}{p}]$.

c) By independence, we have $\mathbb{E}(|Y_n|) = \mathbb{E}(Y_n) = \prod_{j=1}^n \mathbb{E}(\lambda^{X_j}) = (\lambda p + \frac{1}{\lambda}(1-p))^n$ and

$$\mathbb{E}(Y_n^2) = \prod_{j=1}^n \mathbb{E}(\lambda^{2X_j}) = (\lambda^2 p + \frac{1}{\lambda^2}(1-p))^n$$

d) Using the analysis done in question b), we obtain that

$$\sup_{n\in\mathbb{N}} \mathbb{E}(|Y_n|) < +\infty \quad \text{if and only if} \quad \lambda p + \frac{1}{\lambda}(1-p) \leq 1 \quad \text{if and only if} \quad \lambda \in [1, \frac{1-p}{p}]$$

Likewise,

$$\sup_{n\in\mathbb{N}} \mathbb{E}(Y_n^2) < +\infty \quad \text{if and only if} \quad \lambda^2 p + \frac{1}{\lambda^2}(1-p) \leq 1 \quad \text{if and only if} \quad \lambda \in [1, \sqrt{\frac{1-p}{p}}]$$

as we simply need to replace λ by λ^2 here.

e) Numerically (see the code), we see clearly that $Y_n \equiv 1$ when $\lambda = 1$ and that $Y_n \underset{n\to\infty}{\to} 0$ a.s. for all $\lambda > 1$. Theoretically, this is due to the fact that $S_n \underset{n\to\infty}{\to} -\infty$ a.s. (one can use Hoeffding's inequality to justify this fact), so that $Y_n = \lambda^{S_n}$ converges a.s. to 0 when $\lambda > 1$ (even though Y is a submartingale for $\lambda > \frac{1-p}{p}$).

The martingale convergence theorem itself only allows to conclude that Y converges a.s. when $\lambda \in [1, \frac{1-p}{p}]$, i.e., when the first condition in question d) is satisfied.

f) For the process Y to converge also in L^2 towards its limit, we need the second condition in part d) to hold, i.e., $\lambda \in [1, \sqrt{\frac{1-p}{p}}]$ (indeed, if this condition is not met, then $\mathbb{E}(Y_n^2)$ diverges, so no L^2 convergence can take place).

g) The only value of λ for which the equality $\mathbb{E}(Y_\infty|\mathcal{F}_n) = Y_n$ is satisfied for all $n \in \mathbb{N}$ is the trivial case $\lambda = 1$. For all the other values of $\lambda \in [1, \sqrt{\frac{1-p}{p}}]$, Y is a supermartingale, which gives $\mathbb{E}(Y_\infty|\mathcal{F}_n) \leq Y_n$ for all $n \in \mathbb{N}$. But in this case, we already know that $Y_\infty = 0$, so what the above inequality is actually saying is that $0 \leq Y_n$, i.e., not much...

Part II. We consider the case $\lambda = \frac{1-p}{p}$ (> 1).

a) Method: observe first that $T_a = \inf\{n \in \mathbb{N} : S_n \geq a \quad \text{or} \quad S_n \leq -a\}$. Run then $M = 1'000$ times the process S for $N = 1'000$ time steps; count the number of times m the barrier a is reached before $-a$ (see the code for the implementation). The probability $P = \mathbb{P}(\{Y_{T_a} = \lambda^a\})$ is then approximately given by m/M (there are of course fluctuations here).

b) It is true that $\mathbb{E}(Y_{T_a}) = \mathbb{E}(Y_0)$. The third version of the optional stopping theorem allows to say this, as the martingale Y makes bounded jumps and is contained in the interval $[\lambda^{-a}, \lambda^a]$ before the stopping time T_a.

c) As the martingale convergence theorem applies, we obtain

$$1 = \mathbb{E}(Y_0) = \mathbb{E}(Y_{T_a}) = \lambda^a \, P + \lambda^{-a} \, (1 - P)$$

so

$$P = \frac{1 - \lambda^{-a}}{\lambda^a - \lambda^{-a}} = \frac{\lambda^a - 1}{\lambda^{2a} - 1} = \frac{1}{\lambda^a + 1}$$

This probability decays exponentially in a.

d) Method: Run again $M = 1'000$ times the process S for $N = 1'000$ time steps; count the number of times m the barrier a is reached at all (see the code for the implementation). The probability $P' = \mathbb{P}(\{Y_{T'_a} = \lambda^a\})$ is then approximately given by m/M (there are of course also fluctuations here).

e) It is also true that $\mathbb{E}(Y_{T'_a}) = \mathbb{E}(Y_0)$! The third version of the optional stopping theorem allows to say this, as the martingale Y makes bounded jumps and is contained in the interval $[0, \lambda^a]$ before the stopping time T'_a (even though the process S itself is unbounded).

f) As the martingale convergence theorem applies, we obtain

$$1 = \mathbb{E}(Y_0) = \mathbb{E}(Y_{T'_a}) = \lambda^a \, P' + 0 \, (1 - P')$$

[Comment: if level a is not reached, this is saying that the process S goes to $-\infty$ and never comes back to level a; of course, this can only happen with positive probability when S is a (strict) supermartingale with a negative drift.] Therefore,

$$P' = \frac{1}{\lambda^a}$$

which is strictly greater than P, but decays also exponentially in a.