Homework 7

Math 561: Theory of Probability I

Due date: March 14, 2018

Each problem is worth 10 points and only five randomly chosen problems will be graded if there are more than 5 problems. Please indicate whom you worked with, it will not affect your grade in any way.

1. **Mill’s ratio Approximation.** Mill’s ratio is defined as the function

\[ \Psi(x) := \frac{1 - \Phi(x)}{\phi(x)}, \quad x > 0, \]

where \( \phi(\cdot) \) and \( \Phi(\cdot) \) are the PDF and CDF of a N(0, 1) r.v., respectively. For any integer \( k \geq 0 \), define the function

\[ h_k(x) := \frac{1 - x^{-1} + 3 \cdot x^{-5} - \cdots + (-1)^k \frac{(2k-1)!!}{x^{2k+1}}}{2} \text{ for } x > 0, \]

where \( (2k-1)!! = (2k-1)(2k-3)\cdots3\cdot1 = \frac{(2k)!}{2^k k!} \). Prove that for any \( k \), we have

\[ \Psi(x) \begin{cases} \leq h_k(x) & \text{if } k \text{ is even} \\ \geq h_k(x) & \text{if } k \text{ is odd} \end{cases} \]

You can use the fact that \( \int_{0}^{\infty} x^k e^{-x^2} dx = k! \).

**Solution:** Clearly we have

\[ \Psi(x) := \frac{1 - \Phi(x)}{\phi(x)} = \int_{x}^{\infty} e^{-u^2/2} du = \int_{0}^{\infty} e^{-y^2/2} e^{-x^2/2} dy = \frac{1}{x} \int_{0}^{\infty} e^{-y^2/2} e^{-y} dy. \]

It is easy to check that \( e^{-x} \leq \sum_{i=0}^{k} (-x)^{k}/k! \text{ when } k \text{ is even and } e^{-x} \geq \sum_{i=0}^{k} (-x)^{k}/k! \text{ when } k \text{ is odd. Thus for } k \text{ even we have} \]

\[ \Psi(x) = \frac{1}{x} \int_{0}^{\infty} e^{-y^2/2} e^{-y} dy \leq \frac{1}{x} \int_{0}^{\infty} \sum_{i=0}^{k} \frac{(-1)^k}{2^{k+1} i!} x^{2k+1} e^{-y} dy = \sum_{i=0}^{k} \frac{(-1)^k}{2^{k+1} i!} \cdot \frac{(2k)!}{2^k k!}. \]

The lower bound is similar.

2. Let \( X_1, X_2, \ldots \) be i.i.d. and have the standard normal distribution.

(i) Using Mill’s ratio approximation

\[ \frac{1}{x} - \frac{1}{x^3} \leq \frac{1 - \Phi(x)}{\phi(x)} \leq \frac{1}{x}, \quad x > 0, \]

prove that for any \( \theta \in \mathbb{R} \), we have

\[ \frac{\mathbb{P}(X_1 > x + \theta/x)}{\mathbb{P}(X_1 > x)} \to e^{-\theta} \text{ as } x \to \infty. \]

(ii) Define \( b_n \) such that \( \mathbb{P}(X_1 > b_n) = 1/n \). Show that for \( M_n = \max_{1 \leq i \leq n} X_i \) we have, for any \( x \in \mathbb{R} \),

\[ \mathbb{P}(b_n(M_n - b_n) \leq x) \to \exp(-e^{-x}) \text{ as } n \to \infty. \]
(iii) Show that \( b_n \) satisfies
\[
-\frac{\varepsilon}{2\sqrt{2\log n}} \leq b_n - \left[ \sqrt{2\log n} - \frac{\log(4\pi \log n)}{2\sqrt{2\log n}} \right] \leq \frac{\varepsilon}{2\sqrt{2\log n}}
\]
for any \( \varepsilon > 0 \) and \( n \) large.

**Solution:**
(i) Use the approximation that
\[
\frac{\mathbb{P}(X_1 > x)}{\phi(x)/x} \rightarrow 1 \text{ as } x \rightarrow \infty
\]
and
\[
\frac{\phi(x + \theta/x)}{\phi(x)} \rightarrow e^{-\theta} \text{ as } x \rightarrow \infty.
\]
(ii) Note that by part (i) we have
\[
n \mathbb{P}(X_1 > b_n + x/b_n) \rightarrow e^{-x} \text{ as } n \rightarrow \infty.
\]
Thus
\[
\mathbb{P}(b_n(M_n - b_n) \leq x) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq b_n + x/b_n) = (1 - \mathbb{P}(X_1 > b_n + x/b_n))^n \rightarrow e^{-e^{-x}}.
\]
(iii) It is enough to show that
\[
1 - \Phi(u_n) \leq \frac{1}{n} \text{ and } 1 - \Phi(l_n) \geq \frac{1}{n}
\]
where
\[
u_n = \sqrt{2\log n} - \frac{\log(4\pi \log n) - \varepsilon}{2\sqrt{2\log n}}
\]
\[
l_n = \sqrt{2\log n} - \frac{\log(4\pi \log n) + \varepsilon}{2\sqrt{2\log n}}.
\]
Here we use the approximation that
\[
\frac{1 - \Phi(x)}{\phi(x)/x} \rightarrow 1 \text{ as } x \rightarrow \infty
\]
to show that \( n(1 - \Phi(u_n)) \approx n\phi(u_n)/u_n < 1 \) and \( n(1 - \Phi(l_n)) \approx n\phi(l_n)/l_n > 1 \) for large \( n \).

3. Let \( (X_n, n \geq 1) \) be a sequence of i.i.d. Exponential(1) r.v.s with density function
\[
f(x) = e^{-x}1_{x>0}.
\]
Define \( S_n = \sum_{i=1}^{n} X_i \). Clearly \( \mathbb{E}(S_n) = n \). Define the function
\[
h(a) = \log a + 1 - a, \ a > 0.
\]
Prove that
\[
n^{-1} \log \mathbb{P}(S_n \geq an) \leq h(a) \text{ when } a > 1 \text{ and } n^{-1} \log \mathbb{P}(S_n \leq an) \leq h(a) \text{ when } 0 < a < 1.
\]

**Solution:** The MGF of Exponential(1) r.v. is \( m(t) = 1/(1-t), t < 1 \). Using Exponential Markov’s inequality we have for any \( a > 1, t \in (0,1) \)
\[
\mathbb{P}(S_n \geq an) \leq e^{-ant} \mathbb{E}(e^{tS_n}) = e^{-ant}m(t)^n = \exp(-n(at - \log m(t))).
\]
Optimizing over \( t \in (0,1) \) to maximize \( at - \log m(t) = at + \log(1-t) \), we get the maximizer \( t = 1 - 1/a \in (0,1) \) (Here we need \( a > 1 \)) and
\[
n^{-1} \log \mathbb{P}(S_n \geq an) \leq -a + 1 + \log a \text{ for } a > 1.
\]
Using the fact \( \mathbb{P}(S_n \leq an) = \mathbb{P}(-S_n \geq -an) \) and similar argument as before we have
\[
n^{-1} \log \mathbb{P}(S_n \leq an) \leq -a + 1 + \log a \text{ for } 0 < a < 1.
\]
4. **Oriented first passage percolation.** Consider the lattice quadrant \{(i, j) : i, j \geq 0\} with directed edges \((i, j) \rightarrow (i + 1, j)\) and \((i, j) \rightarrow (i, j + 1)\). Associate to each edge \(e\) an exponential(1) distributed r.v. \(X_e\), independent for different edges. For each directed path \(\Pi\) of length \(n\) started at the origin \((0, 0)\), let

\[
S_\Pi = \sum_{\text{edges } e \text{ in path } \Pi} X_e.
\]

Let \(H_n\) be the minimum of \(S_\Pi\) over all such paths \(\Pi\) of length \(n\). It can be shown that \(n^{-1}H_n \to c\) a.s., for some constant \(c\). Clearly, \(c \geq 0\). Give explicit nontrivial upper and lower bounds on \(c\).

**Hint:** Use result of previous question for lower bound.

**Solution:** Number of directed paths on length \(n\) starting from the origin is bounded by \(2^n\). Thus we have for \(a < 1\),

\[
P(H_n \leq an) = P(\cup \Pi \{S_\Pi \leq an\}) \leq \sum_\Pi P(S_\Pi \leq an) \leq 2^n P(X_1 + X_2 + \cdots + X_n \leq an) \leq e^{n \log 2 + na(a)}
\]

where \(X_i\)'s are i.i.d. Exponential(1). The function

\[
\log 2 + \log a + 1 - a, a \in (0, 1)
\]

has a unique root at \(c_0 \approx 0.232\) and is negative for \(a < c_0\). For any \(a < c_0\), we have

\[
P(H_n \leq an) \to 0 \text{ as } n \to \infty.
\]

Thus \(c \geq c_0 \approx 0.232\).

For the upper bound an easy bound can be obtained by looking at the horizontal path \(P_n\) of length \(n\) which is sum of \(n\) iid Exponential(1) r.v.s. Clearly, \(H_n \leq S_{P_n}\). By SLLN,

\[
n^{-1}S_{P_n} \to E(X_e) = 1 \text{ a.s.}
\]

Thus, \(c \leq 1\). A better upper bound can be obtained by considering the path which chooses the edge (between the up and right edges) with minimum \(X\)-value at every step starting from the origin. Then, we have

\[
H_n \leq \sum_{i=1}^n \min\{X_{i, \text{up}}, X_{i, \text{right}}\}
\]

where all \(x\)'s are iid exponential(1). One can check that \(E(\min\{X_{i, \text{up}}, X_{i, \text{right}}\}) = 1/2\). Thus by SLLN, we have

\[
c \leq 1/2.
\]

5. **(Lyapunov's Theorem)** Let \(X_1, X_2, \ldots\) be independent r.v.s with \(E(|X_i|^p) < \infty\) for all \(i \geq 1\) for some \(p > 2\). Define \(S_n = X_1 + X_2 + \cdots + X_n, s_n^2 = \text{Var}(S_n)\). Assume that

\[
\frac{1}{s_n^p} \sum_{i=1}^n E|X_i - E X_i|^p \to 0 \text{ as } n \to \infty.
\]

Show that

\[
\frac{S_n - E S_n}{s_n} \Rightarrow N(0, 1) \text{ as } n \to \infty.
\]

**Hint:** Use Lindeberg’s condition.

**Solution:** Use the fact that for \(p > 2, t > 0\) we have

\[
E(|X|^2 I_{|X| \geq t}) = t^{-(p-2)} E(t^{p-2}|X|^2 I_{|X| \geq t}) \leq t^{-(p-2)} E(|X|^p).
\]

Thus

\[
\sum_{i=1}^n E\left(\frac{|X_i - E X_i|^2}{s_n^2} \cdot I\{|X_i - E X_i| \geq \varepsilon s_n\}\right) \leq \sum_{i=1}^n t^{-(p-2)} E\left(\frac{|X_i - E X_i|^p}{s_n^p}\right)\]

\[
= \varepsilon^{2-p} \sum_{i=1}^n E|X_i - E X_i|^p = o(1).
\]
