MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070 Fall 2013

Problem Set 1 due 9/30/2013

Problem 1. The following identity was used in the proof of Theorem 1 in Lecture 4: $\sup\{\theta > 0 : M(\theta) < \exp(C\theta)\} = \inf_{t>0} tI(C + \frac{1}{t})$ (see the proof for details). Establish this identity.

Hint: Establish the convexity of $\log M(\theta)$ in the region where $M(\theta)$ is finite. Letting $\theta^* = \sup\{\theta > 0 : M(\theta) < \exp(C\theta)\}$ use the convexity property above to argue that $\log M(\theta) \geq \theta^* C + \frac{d \log M(\theta)}{d\theta}\Big|_{\theta^*}(\theta - \theta^*)$. Use this property to finish the proof of the identity.

Problem 2. This problem concerns the rate of convergence to the limits for the large deviations bounds. Namely, how quickly does $n^{-1} \log \mathbb{P}(n^{-1}S_n \in A)$ converge to $-\inf_{x\in A} I(x)$, where S_n is the sum of n i.i.d. random variables? Of course the question is relevant only to the cases when this convergence takes place.

(a) Let S_n be the sum of n i.i.d. random variables $X_i, 1 \leq i \leq n$ taking values in $\mathbb{R}$. Suppose the moment generating function $M(\theta) = \mathbb{E}[\exp(\theta X)]$ is finite everywhere. Let $a \geq \mu = \mathbb{E}[X]$. Recall that we have established in class that in this case the convergence $\lim_n n^{-1} \log \mathbb{P}(n^{-1}S_n \geq a) = -I(a)$ takes place. Show that in fact there exists a constant C such that

$$\left| n^{-1} \log \mathbb{P}(n^{-1}S_n \geq a) + I(a) \right| \leq \frac{C}{n},$$

for all sufficiently large n. Namely, the rate of convergence is at least as fast as $O(1/n)$.

Hint: Examine the lower bound part of the Cramér's Theorem.

(b) Show that the rate $O(1/n)$ cannot be improved.

Hint: Consider the case $a = \mu$.

Problem 3. Let X_n be an i.i.d. sequence with a Gaussian distribution $N(0, 1)$ and let Y_n be an i.i.d. sequence with a uniform distribution on $[-1, 1]$. Prove that the limit

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\Big((\frac{1}{n}\sum_{1\le i\le n}X_i)^2+(\frac{1}{n}\sum_{1\le i\le n}Y_i)^2\ge 1\Big)$$

exist and compute it numerically. You may use MATLAB (or any other software of your choice) and an approximate numeric answer is acceptable.

Problem 4. Let Z_n be the set of all length-n sequences of $0, 1$ such that a) 1 is always followed by 0 (so for, for example $001101\cdots$ is not a legitimate sequence, but $001010\cdots$ is); b) the percentage of 0-s in the sequence is at least 70%. (Namely, if $X_n(z)$ is the number of zeros in a sequence $z\in Z_n$, then $X_n(z)/n\ge 0.7$ for every $z\in Z_n$). Assuming that the limit $\lim_{n\to\infty}\frac{1}{n}\log Z_n$ exists, compute its limit. You may use MATLAB (or any other software of your choice) to assist you with computations, and an approximate numerical answer is acceptable. You need to provide details of your reasoning.

Problem 5. Problem 1 from Lecture 6.

Problem 6. Problem 2 from Lecture 6.

15.070J / 6.265J Advanced Stochastic Processes
Fall 2013