Math 170B **Winter 2017**

Problem Set 5

Lecturer: Steven Heilman *Kyle Barron*

Exercise 1

Let X be a standard Gaussian random variable. Let $t > 0$ and let n be a positive even integer. Show that

$$\mathbf{P}(X > t) \leq \frac{(n-1)(n-3)\cdots(3)(1)}{t^n}.$$

That is, the function $t \mapsto \mathbf{P}(X > t)$ decays faster than any monomial.

Proof: Let X be a standard Gaussian random variable and let $t > 0$. Let n be a positive even integer. Recall that from problem set 3, question 2, we found that for n even, $\mathrm{E}[X^n] = (n-1)(n-3)\cdots 1$. Now using the Markov Inequality we have

$$\mathbf{P}(X > t) \leq \mathbf{P}(|X| > t) = \mathbf{P}(X^n > t^n) \leq \frac{\mathrm{E}[X^n]}{t^n} = \frac{(n-1)(n-3)\cdots 1}{t^n},$$

where the second equality is since n is even, and the inequality is using the Markov Inequality, since X^n is nonnegative. ■

Exercise 2

Let X be a random variable. Let $t > 0$. Show that

$$\mathbf{P}(|X| > t) \leq \frac{\mathrm{E}X^4}{t^4}.$$

Proof: Let X be a random variable and let $t > 0$.

$$\begin{aligned}
\mathbf{P}(|X| > t) &= \mathbf{P}(|X|^4 > t^4) \\
&= \mathbf{P}(X^4 > t^4) \\
&\leq \frac{\mathrm{E}[X^4]}{t^4}.
\end{aligned}$$

where the first equality is true because x^4 is an increasing function when $x \geq 0$ (and here both $|X|$ and t are nonnegative), the second equality is true because $x^4 \geq 0$ always, and the third line uses the Markov Inequality. ■

Exercise 3

(The Chernoff Bound.) Let X be a random variable and let $r > 0$. Show that, for any $t > 0$,

$$\mathbf{P}(X > r) \leq e^{-tr} M_X(t).$$

Consequently, if $X_1, \ldots, X_n$ are independent random variables with the same CDF, and if $r, t > 0$,

$$\mathbf{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i > r\right) \leq e^{-trn}(M_{X_1}(t))^n.$$

For example, if $X_1, \ldots, X_n$ are independent Bernoulli random variables with parameter $0 < p < 1$, and if $r, t > 0$,

$$\mathbf{P}\left(\frac{X_1 + \cdots + X_n}{n} - p > r\right) \leq e^{-trn}(e^{-tp}[pe^t + (1-p)])^n.$$

And if we choose t appropriately, then the quantity $\mathbf{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - p) > r\right)$ becomes exponentially small as either n or r become large. That is, $\frac{1}{n}\sum_{i=1}^{n} X_i$ becomes very close to its mean. Importantly, the Chernoff bound is much stronger than either Markov's or Cheyshev's inequality, since they only respectively imply that

$$\mathbf{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - p\right| > r\right) \leq \frac{2p(1-p)}{nr}, \quad \mathbf{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - p\right| > r\right) \leq \frac{p(1-p)}{nr^2}.$$

Proof: Let X be a random variable and let $r > 0$. Then we have that

$$\mathbf{P}(X \geq r) = \mathbf{P}(e^{tX} \geq e^{tr}) \leq \frac{\mathrm{E}[e^{tX}]}{e^{tr}} = e^{-tr}M_X(t),$$

where the first equality uses that $t > 0$ so the exponential function is increasing, and the inequality uses the Markov inequality (which is valid since $e^x \geq 0$ for all x).

Now suppose that $X_1, ...X_n$ are independent and identically distributed. Then

$$\begin{aligned}\mathbf{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i > r\right) &= \mathbf{P}\left(\sum_{i=1}^{n} X_i > rn\right) \\ &\leq e^{-trn} M_{\sum_{i=1}^{n} X_i}(t) \\ &= e^{-trn}\left(M_{X_1}(t)\right)^n\end{aligned}$$

■

Exercise 4

Let $X_1, X_2, \ldots$ be independent random variables, each with exponential distribution with parameter $\lambda = 1$. For any $n \geq 1$, let $Y_n \coloneqq \max(X_1, \ldots, X_n)$. Let $0 < a < 1 < b$. Show that $\mathbf{P}(Y_n \leq a \log n) \to 0$ as $n \to \infty$, and $\mathbf{P}(Y_n \leq b \log n) \to 1$ as $n \to \infty$. Conclude that $Y_n / \log n$ converges to 1 in probability as $n \to \infty$.

Proof: Let $X_1, X_2, ...$ be independent random variables, each exponentially distributed with $\lambda = 1$. For any $n \geq 1$, define $Y_n \coloneqq \max(X_1, ..., X_n)$. First let $c > 0$. Then

$$\begin{aligned}\mathbf{P}(Y_n \leq c\log n) &= \mathbf{P}(\max(X_1, ..., X_n) \leq c \log n) \\ &= \mathbf{P}(X_1 \leq c\log n \cap X_2 \leq c \log n \cap \cdots \cap X_n \leq c\log n) \\ &= \mathbf{P}(X_1 \leq c\log n) \cdots \mathbf{P}(X_n \leq c \log n) \\ &= \mathbf{P}(X_1 \leq c\log n)^n \\ &= (1 - e^{-c\log n})^n.\end{aligned}$$

Now define $f(u) = e^u$ and let $g(n) = n\log(1-\frac{1}{n^c})$.

$$\begin{aligned}\lim_{n\to\infty}\mathbf{P}(Y_n \leq c\log n) &= \lim_{n\to\infty}(1-e^{-c\log n})^n \\ &= \lim_{n\to\infty} e^{n\log(1-\frac{1}{n^c})}\end{aligned}$$

Now just consider $n\log(1-\frac{1}{n^c})$, which we can rewrite as $\frac{\log(1-\frac{1}{n^c})}{\frac{1}{n}}$. This limit as $n\to\infty$ is an indeterminate firm, so we use l'hopital's rule

$$\begin{aligned}\frac{d}{dn}(\log(1-\frac{1}{n^c})) &= \frac{c}{n^{c+1}-n} \\ \frac{d}{dn}(\frac{1}{n}) &= \frac{-1}{n^2}.\end{aligned}$$

Now note that again $\lim_{n\to\infty}\frac{-cn^2}{n^{c+1}-n} = \lim_{n\to\infty}\frac{-cn}{n^c-1}$ which is again an indeterminate form. Using l'hopital's rule again, we get:

$$\begin{aligned}\frac{d}{dn}(-cn) &= -c \\ \frac{d}{dn}(n^c-1) &= cn^{c-1}.\end{aligned}$$

So this limit becomes $\lim_{n\to\infty}\frac{-1}{n^{c-1}}$. Now consider the case when $c\in(0,1)$ and when $c>1$.

- Let $c\in(0,1)$. Then $c-1\in(-1,0)$ and $\lim_{n\to\infty}\frac{-1}{n^{c-1}} = -\infty$. So we have $\lim_{n\to\infty} g(n) = -\infty$. Then by the limit chain rule, we have $\lim_{u\to-\infty} f(u) = \lim_{u\to-\infty} e^u = 0$.
- Let $c>1$. Then $c-1>0$, so $\lim_{n\to\infty}\frac{-1}{n^{c-1}} = 0$. So $\lim_{n\to\infty} g(n) = 0$. Then by the limit chain rule, $\lim_{u\to 0} f(u) = e^0 = 1$.

Therefore, we have that $\lim_{n\to\infty}\mathbf{P}(Y_n\leq c\log n) = \begin{cases}0, & c\in(0,1)\\ 1, & c>1.\end{cases}$ Now we need to conclude that $\frac{Y_n}{\log n}\to 1$ as $n\to\infty$ in probability. First, note that $x\leq c, |x-1|>\epsilon$ is true if and only if $x>1+\epsilon$ or $x<1-\epsilon$.

$$\mathbf{P}(|\frac{Y_n}{\ln n}-1|>\epsilon) = \mathbf{P}(\frac{Y_n}{\ln(n)}-1>\epsilon)+\mathbf{P}(\frac{Y_n}{\ln n}-1<-\epsilon).$$

Now let $c = 1+\epsilon$. Note that $\mathbf{P}(\frac{Y_n}{\ln n}-1<-\epsilon)\to 0$ corresponds to case 1 above and $\mathbf{P}(\frac{Y_n}{\ln(n)}-1>\epsilon)\to 1$ corresponds to case 2 above. Then note that the complement of $(\frac{Y_n}{\ln n}\leq 1+\epsilon)$ is $(\frac{Y_n}{\ln n}-1>\epsilon)$. So $\mathbf{P}(\frac{Y_n}{\ln n}-1>\epsilon) = 1-\mathbf{P}(\frac{Y_n}{\ln n}\leq 1+\epsilon)$, and in the limit this converges to zero. Therefore, $\mathbf{P}(|\frac{Y_n}{\ln n}-1|)$ converges in probability to 1. ■

Exercise 5

We say that random variables $X_1, X_2, \ldots$ converge to a random variable X in L_2 if

$$\lim_{n\to\infty}\mathrm{E}|X_n-X|^2 = 0.$$

Show that, if $X_1, X_2, \ldots$ converge to X in L_2, then $X_1, X_2, \ldots$ converges to X in probability.

Is the converse true? Prove your assertion.

Proof: Let $X, X_1, X_2, ...$ be random variables such that $X_1, X_2, ...$ converge to X in L_2. That is,

$$\lim_{n\to\infty} \mathrm{E}[|X_n - X|^2] = 0.$$

We want to show that $X_1, X_2, ...$ converges to X in probability, that is, if for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X| > \epsilon) = 0.$$

First recall that the Markov Inequality says that

$$\mathbf{P}(|Z| \geq t) \leq \frac{\mathrm{E}[|Z|^2]}{t^2}, \forall t > 0.$$

Let $\epsilon > 0$, let $t = \epsilon$, and let $Z = X_n - X$. So then we have

$$\mathbf{P}(|X_n - X| > \epsilon) \leq \frac{\mathrm{E}[(X_n - X)^2]}{\epsilon^2}.$$

Since we know that $X_1, X_2, ...$ converge to X in L_2,

$$\lim_{n\to\infty} \mathrm{E}[(X_n - X)^2] = 0 \Rightarrow \lim_{n\to\infty} \mathbf{P}(|X_n - X| > \epsilon) = 0.$$

So $X_1, X_2, ...$ converge in probability.

The converse is not true: convergence in probability does not imply convergence in L_2. For a counterexample, let X_n be a random variable such that for $n = 1, 2, ...,$

$$X_n = \begin{cases} e^n, & \text{with probability } \frac{1}{n} \\ 0, & \text{with probability } 1 - \frac{1}{n}. \end{cases}$$

Then for all $\epsilon > 0$, $\mathbf{P}(|X_n| < \epsilon) = 1 - \frac{1}{n} \to 1$ as $n \to \infty$. Therefore, X_n converges to 0 in probability since $\mathbf{P}(|X_n| > \epsilon) = 1 - \mathbf{P}(|X_n| < \epsilon)$. However $\mathrm{E}[|X_n - 0|^2] = \mathrm{E}[|X_n|^2] = \frac{e^{2n}}{n}$. Since $\frac{e^{2n}}{n} \to \infty$ as $n \to \infty$, X_n does not converge to 0 in L_2. ■

Exercise 6

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables such that $\mathrm{E}|X| < \infty$ and $\mathrm{var}(X) < \infty$. For any $n \geq 1$, define

$$Y_n := \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Show that $Y_1, Y_2, \ldots$ converges in probability. Express the limit in terms of $\mathrm{E}X$ and $\mathrm{var}(X)$.

Proof: Let $X_1, X_2, ...$ be independent and identically distributed with $\mathrm{E}[|X|] < \infty$ and $\mathrm{Var}(X) < \infty$. Define

$$Y_n := \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Let $\epsilon > 0$. To show that $Y_1, Y_2, ...$ converges in probability we need to show that $\lim_{n\to\infty} \mathbf{P}(|Y_n - Y| > \epsilon) = 0$. I'll show that Y, the limit of Y_n as $n \to \infty$, is equal to the second moment of X_1 (which is the same for all X since they're independent and identically distributed). I use Markov's inequality with $n = 2$ and then simplify:

$$\begin{aligned}
\mathbf{P}(|Y_n - Y| > \epsilon) &= P\left(\left|\frac{x_1^2 + ... + x_n^2}{n} - \mathrm{E}[X_1^2]\right| > \epsilon\right) \\
&\leq \frac{1}{\epsilon^2}\mathrm{E}\left[\left|\frac{x_1^2 + ... + x_n^2}{n} - \mathrm{E}[X_1^2]\right|^2\right] \\
&= \frac{1}{\epsilon^2}\mathrm{E}\left[\left(\frac{x_1^2 + ... + x_n^2}{n} - \frac{n\mathrm{E}[X_1^2]}{n}\right)^2\right] \\
&= \frac{1}{\epsilon^2}\mathrm{E}\left[\left(\frac{x_1^2 + ... + x_n^2}{n} - \frac{\mathrm{E}[X_1^2] + ... + \mathrm{E}[X_n^2]}{n}\right)^2\right] \\
&= \frac{1}{\epsilon^2}\mathrm{E}\left[\left(\frac{x_1^2 + ... + x_n^2}{n} - \frac{\mathrm{E}[X_1^2 + ... + X_n^2]}{n}\right)^2\right] \\
&= \frac{1}{\epsilon^2 n^2}\mathrm{E}\left[\left(x_1^2 + ... + x_n^2 - \mathrm{E}[X_1^2 + ... + X_n^2]\right)^2\right] \\
&= \frac{1}{\epsilon^2 n^2}\mathrm{Var}(X_1^2 + ...x_n^2) \\
&= \frac{1}{\epsilon^2 n}\mathrm{Var}(X_1^2).
\end{aligned}$$

Now letting $n \to \infty$, we have that $\frac{\mathrm{Var}(X_1^2)}{\epsilon^2 n} \to 0$. Since probabilities are nonnegative and we have that $\mathbf{P}(|Y_n - Y| > \epsilon) \leq \frac{\mathrm{Var}(X_1^2)}{\epsilon^2 n}$, we know that $\mathbf{P}(|Y_n - Y| > \epsilon) \to 0$. Therefore, $Y_1, Y_2, ...$ converges in probability to $\mathrm{E}[X_1^2] = \mathrm{Var}(X_1) + \mathrm{E}[X_1]^2$. (Note that this is well defined since both the variance and mean of X_1 are finite.) ■