

## 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,  $\mathbf{E}[X^n] = (n-1)(n-3) \cdots 1$ . Now using the Markov Inequality we have

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

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

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

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

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

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

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

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

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

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

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

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  $X_1, \dots, X_n$  are independent and identically distributed. Then

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

■

## Exercise 4

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

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

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

■

### Exercise 5

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

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

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

Is the converse true? Prove your assertion.

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

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

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

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

First recall that the Markov Inequality says that

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

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

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

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

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

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

The converse is not true: convergence in probability does not imply convergence in  $L_2$ . Recall that

■

### Exercise 6

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

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

Show that  $Y_1, Y_2, \dots$  converges in probability. Express the limit in terms of  $EX$  and  $\text{var}(X)$ .

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

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

Let  $\epsilon > 0$ . To show that  $Y_1, Y_2, \dots$  converges in probability we need to show that  $\lim_{n \rightarrow \infty} \mathbf{P}(|Y_n - Y| > \epsilon) = 0$ . I'll show that  $Y$ , the limit of  $Y_n$  as  $n \rightarrow \infty$ , 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:

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

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