

Dr. J. Hermon

Dominic Klukas, SN 64343878

**This assignment is due in Canvas at 23:59 p.m. on Monday, November 17.**

**Late assignments are not accepted.** This assignment is longer than other assignments and will have a double weight. If this is your lowest graded assignment it would still be omitted from your HW assignment average (despite having a double weight).

1. A collection of random variables  $(X_n : n \in \mathcal{I})$  (defined on the same probability space) is called *uniformly integrable* if for all  $\varepsilon$  there exists some  $M = M(\varepsilon)$  such that  $\mathbb{E}[|X_n| \mathbb{1}_{|X_n| > M}] < \varepsilon$  for all  $n \in \mathcal{I}$ . Prove that  $(X_n : n \in \mathcal{I})$  is uniformly integrable if and only if  $\sup_{n \in \mathcal{I}} \mathbb{E}[|X_n|] < \infty$  and for any  $\varepsilon > 0$ , there exists a  $\delta = \delta(\varepsilon) > 0$  such that  $E[|X_n| \mathbb{1}_A] < \varepsilon$  for  $n \in \mathcal{I}$ , for all  $A \in \mathcal{F}$  with  $\mathbb{P}(A) < \delta$ .

### Solution.

( $\rightarrow$ ) Suppose  $(X_n : n \in \mathcal{I})$  is uniformly integrable. Pick any  $\varepsilon > 0$ , so that we choose  $M$  such that  $\mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq M}] < \varepsilon$  for all  $n$ . But then, for all  $n$ ,

$$\sup_{n \in \mathcal{I}} \mathbb{E}[|X_n|] = \sup_{n \in \mathcal{I}} (\mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq M}] + \mathbb{E}[|X_n| \mathbb{1}_{|X_n| < M}]) \leq \varepsilon + M < \infty.$$

We saw this technique in the proof of proposition 9.1.5 in the text.

Now, let  $\varepsilon > 0$ . Then, there exists some  $M$  such that  $\forall n \in \mathcal{I}$ ,  $\mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq M}] < \varepsilon/2$  by assumption. Let  $\delta = \varepsilon/2M$ . Let  $n \in \mathcal{I}$ . Let  $A \in \mathcal{F}$  such that  $P(A) < \delta$ . Then,

$$\begin{aligned} \mathbb{E}[|X_n| \mathbb{1}_A] &= \mathbb{E}[|X_n| \mathbb{1}_{A \cap \{|X_n| \geq M\}}] + \mathbb{E}[|X_n| \mathbb{1}_{A \cap \{|X_n| < M\}}] \\ &\leq \varepsilon/2 + MP(A \cap \{|X_n| < M\}) \\ &\leq \varepsilon/2 + M\delta \leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Therefore, the conclusions are satisfied:  $\sup_{n \in \mathcal{I}} \mathbb{E}[|X_n|] < \infty$ , and for any  $\varepsilon > 0$  there exists  $\delta > 0$  such that  $E[|X_n| \mathbb{1}_A] < \varepsilon$  for all  $n \in \mathcal{I}$ , and  $A \in \mathcal{F}$  with  $P(A) < \delta$ .

( $\leftarrow$ ) Suppose  $\sup_{n \in \mathcal{I}} \mathbb{E}[|X_n|] < \infty$ , and for any  $\varepsilon > 0$  there exists  $\delta > 0$  such that  $\mathbb{E}[|X_n| \mathbb{1}_A] < \varepsilon$  for all  $n \in \mathcal{I}$ , and  $A \in \mathcal{F}$  with  $P(A) < \delta$ . Let  $\varepsilon > 0$ . Let  $C = \sup_{n \in \mathcal{I}} \mathbb{E}[|X_n|]$ . Now, by assumption there exists some  $\delta > 0$  such that for all  $A \in \mathcal{F}$ , and for all  $n \in \mathcal{I}$ ,  $\mathbb{E}[|X_n| \mathbb{1}_A] < \varepsilon$ . Let  $M = 2C/\delta$ .

Let  $n \in \mathcal{I}$ . Consider  $\|X_n\| \mathbb{1}_{\|X_n\| > M}$ . If  $P(\|X_n\| > M) < \delta$ , then  $\mathbb{E}[\|X_n\| \mathbb{1}_{\|X_n\| > M}] < \varepsilon$  and we are done. Suppose  $P(\|X_n\| > M) \geq \delta$ . Then, we compute, for this fixed  $n$ ,

$$\begin{aligned} \mathbb{E}[|X_n|] &= \mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq M}] + \mathbb{E}[|X_n| \mathbb{1}_{|X_n| < M}] \\ &\geq \mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq M}] \\ &\geq M \cdot P(\|X_n\| > M) \geq M \cdot \delta = 2 \cdot C > \sup_{n \in \mathcal{I}} \mathbb{E}[|X_n|]. \end{aligned}$$

This is a contradiction. Therefore, we must have  $P(\|X_n\| > M) < \delta$ , and so that  $\mathbb{E}[\|X_n\| \mathbb{1}_{\|X_n\| > M}] < \varepsilon$  and our conclusion holds.

2. Suppose that  $X_1, X_2, \dots$  are random variables with  $X_1 \geq X_2 \geq X_3 \geq \dots \geq 0$ , with  $X_n(\omega) \rightarrow X(\omega)$  for all  $\omega \in \Omega$ , and with  $\mathbb{E}X_1 < \infty$ .
  - (a) Prove that  $\lim_{n \rightarrow \infty} \mathbb{E}X_n = \mathbb{E}X$ .
  - (b) Give an example to show that the conclusion in part (a) need not hold if  $\mathbb{E}X_1 = \infty$ .

**Solution.**

- (a) This is a simple application of the Dominated Convergence Theorem: Since  $|X_n| \leq X_1$ ,  $\mathbb{E}[X_1] < \infty$ ,  $X_n \rightarrow X$  w.p. 1 (indeed, pointwise!), we have that the conditions of the theorem are satisfied and  $\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = E[X]$ .
- (b) Let  $X_1 = \sum_{n=1}^{\infty} 2^{n-1} \mathbb{1}_{[0,2^{-n}]}$ . Then, let  $X_n = X_1 \mathbb{1}_{[0,2^{-n}]}$ . We have that for any  $m \geq n$ , in the intervals of type  $[2^{-(m+1)}, 2^{-m}]$ , we can compute

$$X_n(\omega) = \sum_{s=n}^m 2^{s-1} = 2^m - 2^n.$$

But then, we have

$$\mathbb{E}[X_n] = \sum_{m=n}^{\infty} \mathbb{E}[X \mathbb{1}_{\omega \in [2^{-(n+1)}, 2^{-n}]})] = \sum_{m=n}^{\infty} (2^m - 2^n)(2^{-m-1}) = \sum_{m=n}^{\infty} 2 - \frac{1}{2^{m-n}} = \infty.$$

However, since  $\mathbb{1}_{[0,2^{-n}]} \rightarrow 0$  w.p. 1, we have that  $X_n \rightarrow 0$ , and  $E[0] = 0$ , so  $E[\lim_{n \rightarrow \infty} X_n] \neq \lim_{n \rightarrow \infty} E[X_n]$ .

- 3. Let  $X$  have a Bernoulli( $\frac{1}{2}$ ) distribution and define the rate function  $I(z) = \sup_{s \in \mathbb{R}} [sz - \log M_X(s)]$ . Prove that

$$I(z) = \begin{cases} \log 2 + z \log z + (1-z) \log(1-z) & (z \in [0, 1]) \\ \infty & \text{otherwise.} \end{cases}$$

(For the case  $z > \frac{1}{2}$  where we used this rate function in class for the coin flipping example, note that the supremum over  $s \in \mathbb{R}$  turns out to be the same as the supremum over  $s > 0$  as appears in Theorem 9.3.4.)

**Solution.**

First we compute

$$M_X(s) = \mathbb{E}[e^{sX}] = \frac{1}{2}(1 + e^s).$$

The function  $f(z, s) = sz - \log M_X(s)$ , so that  $I(z) = \sup_{s \in \mathbb{R}} f(z, s)$  is in  $C^\infty$  in  $s$ , so we can use the first and second derivatives to determine its optimality. Also, since continuous, the supremum is attained at some maximum.

Denote

$$\begin{aligned} \frac{\partial f}{\partial s}(s, z) &= z - \frac{1}{1 + e^s} e^s \\ \frac{\partial^2 f}{\partial s^2}(s, z) &= -\frac{e^s + e^{2s}}{(1 + e^s)^2}. \end{aligned}$$

This second derivative is everywhere negative, so the function is concave in  $s$ . Therefore, the maximum occurs at  $\infty$  if strictly increasing,  $-\infty$  if strictly decreasing, or potentially at some local maximum.

We compute the limits:

$$\lim_{s \rightarrow \infty} zs - \log(1 + e^s) + \log 2 \approx_{s \gg 0} \lim_{s \rightarrow \infty} (z - 1)s + \log 2 \rightarrow \begin{cases} \infty & \text{if } z > 1 \\ \log 2 & \text{if } z = 1 \\ -\infty & \text{if } z < 1 \end{cases}$$

$$\lim_{s \rightarrow -\infty} zs - \log(1 + e^s) \approx_{s \ll 0} \lim_{s \rightarrow \infty} zs \rightarrow \begin{cases} -\infty & \text{if } z > 0 \\ \log 2 & \text{if } z = 0 \\ \infty & \text{if } z < 0 \end{cases}$$

Therefore,  $\sup_{s \in \mathbb{R}} f(s, z) = \infty$  when  $z \in (-\infty, 0) \cup (1, \infty)$ . If  $z \in (0, 1)$ , then both the limits to  $\pm\infty$  are  $-\infty$  so we look for some global maximum. We compute:

$$\begin{aligned} 0 &= f'(z, s) \\ &= z - \frac{e^s}{1 + e^s} \\ &\Rightarrow s = \log(z) - \log(1 - z). \end{aligned}$$

If  $z = 1$ , we have  $\sup f(1, s) = \log 2$  as  $s \rightarrow \infty$ . When  $z = 0$ , then we have  $e^s = 0$  which occurs as  $s \rightarrow -\infty$ , so we must have that  $\sup f(0, s) = \log 2$  as well, as  $s \rightarrow -\infty$ .

But then, when  $z \in (0, 1)$ , we have that the maximum occurs at  $s = \log(z) - \log(1 - z)$ , so we plug this in to get:

$$I(z) = \begin{cases} \log 2 + z \log z + (1 - z) \log(1 - z) & (z \in [0, 1]) \\ \infty & \text{otherwise} \end{cases}.$$

We observe that when  $z \in \{0, 1\}$ , the first equation satisfies the correct value,  $\log 2$ .

4. Use Theorem 9.3.4 to obtain a numerical upper bound for the probability that the proportion of Heads in  $n$  coin flips (fair coin) is 0.6 or higher, for the cases  $n = 10, 100, 1000$ .

### Solution.

If the probability of heads is 0.6 or higher, we have that  $I(0.6) = \log 2 + 0.4 \log(0.4) + 0.6 \log(0.6) \approx 0.02014$ . Then, we plug in:

$$\begin{aligned} P(S_n/n \geq 0) : n = 10 &\Rightarrow e^{-I(0.6)(10)} = e^{-6} \approx 0.818 \\ n = 100 &\Rightarrow e^{-I(0.6)(100)} \approx 0.133 \\ n = 1000 &\Rightarrow e^{-I(0.6)(1000)} \approx 1.8 \times 10^{-9}. \end{aligned}$$

5. (a) If  $X_n \Rightarrow X$  and  $Y_n \Rightarrow c$ , where  $c$  is a constant (and for all  $n$ ,  $X_n$  and  $Y_n$  are defined on the same probability space), then  $X_n + Y_n \Rightarrow X + c$ . (Consequently if  $X_n \Rightarrow X$  and  $Z_n - X_n \Rightarrow 0$  then  $Z_n \Rightarrow X$ .)  
(b) Prove or disprove that, more generally, if  $X_n \Rightarrow X$  and  $Y_n \Rightarrow Y$  then  $X_n + Y_n \Rightarrow X + Y$ .

### Solution.

- (a) Denote  $\mu_X = \mathcal{L}(X)$  for a random variable  $X$ . By theorem 10.1.1, it suffices to show that  $\mu_{X_n+Y_n}(-\infty, x] \rightarrow \mu_{X+c}(-\infty, x]$  for all  $x$  such that  $\mu_{X+c}(\{x\}) = 0$ . In particular,  $\mu_{X+c}(\{x\}) = 0 \Leftrightarrow \mu_X(\{x - c\}) = 0$ . Also, we must have that  $\mu_{X+c}(\{x\}) \neq 0$  for at most countably many  $x$ , since  $\mu_{X+c}(-\infty, x]$  is an increasing function, and so can only have countably many discontinuities, and for each  $x$  such that  $\mu_{X+c}(\{x\}) \neq 0$ , we have a discontinuity at  $x$ .

All that being said, let  $x$  be such that  $\mu_{X+c}(\{x\}) = 0$ . Then, for any  $\delta > 0$ , there exists some  $\delta' > \delta' > 0$  such that  $\mu_{X+c}(\{x - \delta'\}) = \mu_{X+c}(\{x + \delta'\}) = 0$ . Let  $\varepsilon > 0$ . Since  $Y_n \Rightarrow c$ , we have by Theroem 10.1.1 that there exists some  $N$  such that for all  $n > N$ , both  $P(|Y_n - c| > \delta) < \varepsilon$ .

Now, for  $F_{X_n+Y_n}(x)$ , we observe that

$$\begin{aligned}\{X_n + Y_n \leq x\} &\subseteq \{X_n \leq x - c + \delta'\} \cup \{|Y_n - c| > \delta'\} \\ \{X_n + Y_n \geq x\} &\supseteq \{X_n \leq x - c - \delta'\} \cap \{|Y_n - c| \leq \delta'\} = \{X_n \leq x - c - \delta'\} \setminus \{|Y_n - c| > \delta'\}.\end{aligned}$$

Taking probabilities, this comes to:

$$\begin{aligned}F_{X_n}(x - c - \delta') - P(|Y_n - c| > \delta') &\leq F_{X_n+Y_n}(x) \leq F_{X_n}(x - c + \delta') + P(|Y_n - c| > \delta') \\ F_{X_n}(x - c - \delta') - \varepsilon &\leq F_{X_n+Y_n}(x) \leq F_{X_n}(x - c + \delta') + \varepsilon.\end{aligned}$$

Then, taking  $n$  to infinity, and using the fact that  $X_n \Rightarrow X$  and  $\mu_X(x - c - \delta') = \mu_X(x - c + \delta') = 0$  with 10.1.1 (3), it follows that

$$F_X(x - c - \delta') - \varepsilon \leq \lim_{n \rightarrow \infty} F_{X_n+Y_n}(x) \leq F_X(x - c + \delta') + \varepsilon.$$

Since  $\varepsilon$  was arbitrary, this gives us the inequality  $F_X(x - c - \delta') \leq F_{X+Y}(x) \leq F_X(x - c + \delta')$ . Since  $\delta$  was arbitrary, we can take  $\delta \rightarrow 0$ , so we have that

$$\mu_X(-\infty, x - c) \leq \lim_{n \rightarrow \infty} F_{X_n+Y_n}(x) \leq F_X(x - c).$$

However, since  $\mu_X(x - c) = \mu_{X-c}(x) = 0$ , it follows that  $\mu_X((-\infty, x - c)) = F_X(x - c)$ , so that the limit  $\lim_{n \rightarrow \infty} F_{X_n+Y_n}(x)$  exists and equals  $F_{X+Y}(x) = F_X(x - c) = F_{X-c}(x)$ , completing the proof.

- (b) We provide a counter example. The problem relies on the fact that distributions don't rely on the underlying probability space, on which the random variables are defined. Let  $\Omega$  be the interval  $[0, 1]$ , with  $\mu = \mathcal{L}$  the lebesgue measure and  $\mathcal{F} = \mathcal{B}$  the borel subsets of  $[0, 1]$ . Define

$$\begin{aligned}X_n &= \begin{cases} 1 & \omega < 1/2 \\ -1 & \omega \geq 1/2 \end{cases} \\ Z = Y_n &= \begin{cases} -1 & \omega < 1/2 \\ 1 & \omega \geq 1/2 \end{cases}.\end{aligned}$$

for all  $n$ . Now, consider that  $X_n \Rightarrow Z$ , and  $Y_n \Rightarrow Z$  since for any continuous function  $f$ , and for all  $n$ ,

$$\int f dX_n = \frac{1}{2}f(-1) + \frac{1}{2}f(1) = \int fdZ = \int fdY_n.$$

Then,  $X_n + Y_n = 0$  for all  $n$ , so  $X_n + Y_n \rightarrow 0$  in probability, and so also in distribution. However,

$$Z + Z = \begin{cases} -2 & \omega < 1/2 \\ 2 & \omega \geq 1/2 \end{cases}.$$

Then, for any continuous function,

$$\int fd(Z + Z) = 0.5(f(-2) + f(2)) \neq f(0) = \lim_{n \rightarrow \infty} \int fd(X_n + Y_n).$$

6. Let  $X, X_1, X_2, \dots$  be integer-valued discrete random variables. Show that  $X_n \Rightarrow X$  if and only if  $\lim_{n \rightarrow \infty} P(X_n = m) = P(X = m)$  for all integers  $m$ .

**Solution.**

We use Theorem 10.1.1, in particular condition (2):  $X_n \Rightarrow X$  if for all measurable sets  $A$ ,  $\mu_{X_n}(A) \rightarrow \mu_X(A)$ .

Let  $A$  be a measurable set. The nature of  $X_n$  (that the values it takes on don't change from the integers) mean that the boundary of  $A$ ,  $\mu_X(\partial A)$ , actually isn't information required for our proof. Let  $\varepsilon > 0$ . Since  $X_n$  and  $X$  are integer valued random variables, we have that  $\mu_X(A \cap \mathbb{Z}) = \mu_X(A)$  and  $\mu_{X_n}(A \cap \mathbb{Z}) = \mu_{X_n}(A)$ . Let  $M_N = \{m \in \mathbb{Z} : -N \leq m \leq N\}$ . By continuity of probabilities, there exists some  $N$  such that  $\mu_X(M_N) > 1 - \varepsilon$ , since  $\mu_X(\mathbb{Z}) = 1$ , since  $\lim_{N \rightarrow \infty} M_N \nearrow \mathbb{Z}$ . Then, since  $\lim_{n \rightarrow \infty} P(X_n = m) = P(X = m)$  for all  $m \in \mathbb{Z}$ , we have that since  $M_N$  contains finitely many integers, there is some  $N'$  such that for all  $n > N'$ ,  $|P(X_n = m) - P(X = m)| < \varepsilon/|M_N|$  for each  $m \in |M_N|$ . However, this also implies that  $|\mu_{X_n}(M_N) - \mu_X(M_N)| < \varepsilon$ , and so  $\mu_{X_n}(M_N) > \mu_X(M_N) - \varepsilon > 1 - 2\varepsilon$ , so  $\mu_{X_n}(\mathbb{Z} \setminus M_N) < 2\varepsilon$ . Putting these together, we have:

$$\begin{aligned} |\mu_{X_n}(A) - \mu_X(A)| &\leq |\mu_{X_n}(A \cap M_N) - \mu_X(A \cap M_N)| + |\mu_{X_n}(A \cap M_N^c)| + |\mu_X(A \cap M_N^c)| \\ &\leq \varepsilon + 2\varepsilon + \varepsilon = 4\varepsilon \end{aligned}$$

for all  $n > N'$ . Since  $\varepsilon > 0$  was arbitrary, we conclude  $\mu_{X_n}(A) \rightarrow \mu_X(A)$ . Then, applying Theorem 10.1.1 (2), we have  $X_n \Rightarrow X$ , as required.

7. Let  $\phi(t)$  be the characteristic function of the random variable  $X$ . Suppose that  $|\phi(t_0)| = 1$  for some  $t_0 \neq 0$ . Prove there exist  $a, b \in \mathbb{R}$  such that  $P(X \in a + b\mathbb{Z}) = 1$ .

**Solution.** Suppose  $|\phi(t_0)| = 1$  for some  $t_0 \neq 0$ . Suppose that for all  $a, b \in \mathbb{R}$ , we have that  $P(X \in \{a + b\mathbb{Z}\}) < 1$ . Now,  $\phi(t_0) = e^{i\theta}$  for some  $\theta \in \mathbb{R}$ . Then,  $P(X \in \{\theta/t_0 + 2\pi/t_0\mathbb{Z}\}) < 1$ . It follows from the definition of the exponential function that  $x \in \{\theta/t_0 + 2\pi/t_0\mathbb{Z}\}$  iff  $e^{it_0x}/\phi(t_0) = 1$  iff  $\Re(e^{it_0x}/\phi(t_0)) = 1$ . By continuity of probabilities, there must be some  $\alpha > 0$  such that  $P(\Re(e^{it_0X}) < 1 - \alpha) > \varepsilon > 0$  for some  $\varepsilon$ . Now, consider  $\phi(t_0) = \mathbb{E}[e^{it_0X}]$ .

We then compute (since complex integration is the sum of the real and imaginary integrals):

$$\begin{aligned} \Re[\mathbb{E}[e^{it_0X}/\phi(t_0)]] &= \mathbb{E}[\Re(e^{it_0X}/\phi(t_0))\mathbb{1}_{x \in \theta/t_0 + 2\pi/t_0\mathbb{Z}}] + \mathbb{E}[\Re(e^{it_0X}/\phi(t_0))\mathbb{1}_{1-\alpha \leq \Re(e^{it_0X}/\phi(t_0)) < 1}] \\ &\quad + \mathbb{E}[\Re(e^{it_0X}/\phi(t_0))\mathbb{1}_{\Re(e^{it_0X}/\phi(t_0)) < 1-\alpha}] \\ &\leq (1)P(x \in \theta/t_0 + 2\pi/t_0\mathbb{Z}) + (1)P(1 - \alpha \leq \Re(e^{it_0X}/\phi(t_0)) < 1) \\ &\quad + (1 - \alpha)P(\Re(e^{it_0X}/\phi(t_0)) < 1 - \alpha) \\ &= 1 - \alpha P(\Re(e^{it_0X}/\phi(t_0)) < 1 - \alpha) < 1. \end{aligned}$$

However, this contradicts  $\mathbb{E}[e^{it_0X}/\phi(t_0)] = 1$ . Therefore, we must have that  $P(X \in \{\theta/t_0 + 2\pi/t_0\mathbb{Z}\}) = 1$ .

8. Using characteristic functions,<sup>1</sup> prove the following:

- (a) Suppose  $X_i$  are independent with  $N(0, \sigma_i^2)$  distributions. Let  $S_n = X_1 + \dots + X_n$ . Then  $S_n$  has distribution  $N(0, \sum_{i=1}^n \sigma_i^2)$ . In particular, if  $Z_i$  has a standard normal  $N(0, 1)$  distribution then  $\frac{1}{\sqrt{n}}(Z_1 + \dots + Z_n)$  also has a standard normal distribution.
- (b) Suppose  $X_i$  are independent Cauchy random variables (density  $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$ ,  $x \in \mathbb{R}$ ). Let  $S_n = X_1 + \dots + X_n$ . Then  $\frac{1}{n}S_n$  has a Cauchy distribution.<sup>2</sup>

<sup>1</sup>You can look up the characteristic functions for the normal, Cauchy and exponential random variables, it is not necessary to perform the integrals yourself.

<sup>2</sup>Cf. the simulation depicted in Assignment 5.

- (c) Recall that a Geometric( $p$ ) random variable has p.m.f.  $g(k) = (1-p)^{k-1}p$  for  $k \in \mathbb{N}$ . Consider the following variation:  $X_n$  has p.m.f.  $P(X_n = k/n) = (1 - \lambda/n)^{k-1}(\lambda/n)$  for  $k \in \mathbb{N}$ , with  $\lambda > 0$ . Apply the Continuity Theorem<sup>3</sup> to prove that  $X_n$  converges weakly to an  $\text{Exp}(\lambda)$  random variable.<sup>4</sup>

**Solution.**

- (a) We look up the characteristic function of the normal distribution: We have that  $\phi_{X_i}(t) = \exp(-(1/2)\sigma_i^2 t^2)$ . But then,

$$\phi_{S_n}(t) = \phi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n \phi_{X_i}(t) = \prod_{i=1}^n \exp\left(-\frac{1}{2}\sigma_i^2 t^2\right) = \exp\left(-\frac{1}{2}t^2 \sum_{i=1}^n \sigma_i^2\right) = \phi_{N(0, \sum_{i=1}^n \sigma_i^2)}(t).$$

As desired. In particular, if  $Z_i$  has a standard normal distribution, then  $(Z_1 + \dots + Z_n) \sim N(0, n)$ . However, since the mean is zero, the variance is just  $\mathbb{E}[\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i\right)^2] = \frac{1}{n} \mathbb{E}[(\sum_{i=1}^n Z_i)^2] = \frac{1}{n}n = 1$ , so that  $\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \sim N(0, 1)$ .

- (b) We look up the characteristic function of the Cauchy distribution, and we find that it is  $\phi_{\text{Cauchy}}(t) = \exp(-|t|)$ . But then,

$$\phi_{S_n}(t) = \prod_{i=1}^n \phi_{X_i}(t) = \exp(-n|t|).$$

Next, we compute:

$$\mathbb{E}[\exp(it \frac{1}{n} S_n)] = \mathbb{E}[\exp(i(t/n) S_n)] = \exp(-n|(t/n)|) = \exp(-|t|) = \phi_{\text{Cauchy}}(t).$$

Here, we used the fact that  $n \in \mathbb{N}$ , so  $n \geq 1$  and  $|t/n| = |t|/n$ . Therefore,  $\frac{1}{n} S_n \sim \phi_{\text{Cauchy}}(t)$ .

- (c) We look up the characteristic function of  $g(k)$ , and find that it is  $\frac{p}{e^{-it} - (1-p)}$ . Now, if for  $X_n$ , we have  $P(X_n = k/n) = (1 - \lambda/n)^{k-1}(\lambda/n)$ . But then, observe that  $P(nX_n = k) = (1 - \lambda/n)^{k-1}(\lambda/n)$ , so  $nX_n$  is a geometric random variable for  $p = \lambda/n$ . (Naturally, this requires that  $n > \lambda$ , but this should not be a problem since we are fixing  $\lambda$  and considering the limit of  $n$ ). Then, we have

$$\begin{aligned} \phi_{X_n}(t) &= \mathbb{E}[\exp(itX_n)] = \mathbb{E}[\exp(i(t/n)(nX_n))] = \frac{\lambda/n}{e^{-i(t/n)} - (1 - \lambda/n)} \\ &= \frac{\lambda}{n(e^{-i(t/n)} - 1) + \lambda}. \end{aligned}$$

However, then,  $\lim_{n \rightarrow \infty} \phi_{X_n}(t) = \frac{\lambda}{\lim_{n \rightarrow \infty} n(e^{-i(t/n)} - 1) + \lambda} = \frac{\lambda}{\frac{e^{-it(0+1/n)} - e^{-it(0)}}{1/n}}$ , we can see that this is just the limit of a difference quotient of the function  $e^{-itx}$  evaluated at  $x = 0$ , along the series of points  $\{1/n\}_{n \in \mathbb{N}}$ . Since  $e^{-itx}$  is differentiable at  $x = 0$ , every such limit evaluates to the derivative, so we conclude that  $\lim_{n \rightarrow \infty} n(e^{-i(t/n)} - 1) = -it$ . Therefore,  $\lim_{n \rightarrow \infty} \phi_{X_n}(t) = \frac{\lambda}{\lambda - it}$ . We look up the characteristic function of the exponential distribution, and see that it is  $\frac{\lambda}{\lambda - it}$ . We apply the Continuity Theorem, to see that since we have  $\phi_{X_n} \rightarrow \phi_{\text{Exp}(\lambda)}$  then  $X_n \Rightarrow \text{Exp}(\lambda)$ , as desired.

<sup>3</sup>We will not finish the proof of the Continuity Theorem in class until November 17, but its statement in Theorem 11.1.14 (appearing in the lecture notes) is easy to understand and apply.

<sup>4</sup>It is straightforward to prove this via the characterization of convergence in distribution in terms of convergence of the CDFs to the CDF of the limit. This is also the case for question 9.

9. Let  $Y_1, Y_2, \dots$  be i.i.d. random variables (on the same probability space), each of which takes any value in  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$  with equal probability  $\frac{1}{10}$ . Let

$$X_n = \sum_{i=1}^n \frac{Y_i}{10^i}.$$

Apply the Continuity Theorem to prove that  $X_n$  converges in distribution to a uniform random variable on  $[0, 1]$ .

**Solution.**

We compute, using  $j$  instead of  $i$  to avoid confusion with  $i = \sqrt{-1}$ :

$$\phi_{\frac{Y_j}{10^j}}(t) = \mathbb{E} [\exp(itY_j/10^j)] = \sum_{m=0}^9 \exp(itm/10^j) = \frac{1}{10} \frac{1 - e^{it/10^{j-1}}}{1 - e^{it/10^j}}.$$

Then,

$$\begin{aligned} \phi_{X_n}(t) &= \prod_{j=1}^n \phi_{\frac{Y_j}{10^j}}(t) = \frac{1}{10^n} \prod_{j=1}^n \left( \frac{1 - e^{it/10^{m-1}}}{1 - e^{it/10^m}} \right) = \frac{1}{10^n} (1 - e^{it}) \prod_{j=1}^{n-1} \left( \frac{1 - e^{it/10^j}}{1 - e^{it/10^j}} \right) (1/(1 - e^{it/10^n})) \\ &= \frac{1}{10^n} \frac{1 - e^{it}}{1 - e^{it/10^n}}. \end{aligned}$$

Similar to last time, we sniff out a difference quotient to compute the limit:

$$\lim_{n \rightarrow \infty} \frac{1}{10^n} \frac{1 - e^{it}}{1 - e^{it/10^n}} = \frac{1 - e^{it}}{\lim_{n \rightarrow \infty} \frac{e^{it(0+1/10^n)} - e^{it(0)}}{1/10^n}} = \frac{1 - e^{it}}{-it} = \frac{i}{t} (1 - e^{it}).$$

We compute the characteristic formula of the uniform random variable on  $[0, 1]$ :

$$\mathbb{E}[\exp(itU_{[0,1]})] = \int_0^1 e^{itx} dx = \frac{1}{it} e^{itx} \Big|_{x=0}^1 = \frac{i}{t} (1 - e^{it}).$$

Therefore,  $\lim_{n \rightarrow \infty} \phi_{X_n} = \phi_{U_{[0,1]}}$ , so by the Continuity Theorem, we have that  $X_n \Rightarrow U_{[0,1]}$ , as desired.

**Recommended problems.** The following problems from Rosenthal are recommended but are not to be handed in:

9.5.4, 9.5.10, 9.5.15, 9.5.16, 10.3.1, 10.3.2, 10.3.3, 10.3.4, 10.3.6, 10.3.8, 10.3.9, 10.3.10.

For solutions to even-numbered problems see: <http://www.probability.ca/jeff/grprobbook.html>.

More recommended problems (not to be handed in):

- A) Let  $X$  be a Poisson random variable with parameter  $\lambda > 0$ . Compute the moment generating function of  $X$  and differentiate it to compute the mean and variance of  $X$ .
- B) Let  $\alpha, \lambda > 0$ . A random variable  $X$  has a Gamma( $\alpha, \lambda$ ) distribution if its density function is given by

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0). \end{cases}$$

Here  $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ . Recall that  $\Gamma(n) = (n-1)!$  for  $n \in \mathbb{N}$ . Note that an  $\text{Exp}(\lambda)$  random variable has the same distribution as a  $\text{Gamma}(1, \lambda)$  random variable.

(a) Prove that the sum of  $n$  i.i.d.  $\text{Exp}(\lambda)$  random variables has a  $\text{Gamma}(n, \lambda)$  distribution in two ways: by induction on  $n$  using a calculation of the density function using the convolution formula and by comparison of moment generating functions or of characteristic functions.<sup>5</sup>

(b) Prove in two ways (as above) that the sum of independent  $\text{Gamma}(\alpha, \lambda)$  and  $\text{Gamma}(\beta, \lambda)$  random variables is  $\text{Gamma}(\alpha + \beta, \lambda)$ .

$$\text{A useful fact: } \int_0^x t^{r-1}(x-t)^{s-1} dt = x^{r+s-1} \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} \text{ for } r, s, x > 0.$$

Quote of the week: "*The other girls study mathematics,*" Ai-ming said, trying again.

*"That's what we need!" the vendor said, smacking her chopsticks against the metal pot. "Real numbers. Without real numbers, how can we fix our economy, make plans, understand what we need? Young lady, I don't mean to be rude but you should really think about studying mathematics, too."*

*"I will."*

Madeleine Thien in *Do not say we have nothing*

Second quote of the week: *Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different.*

J.W. Goethe

---

<sup>5</sup>We have not proved it, but it is the case that if two random variables have the same moment generating function, which is defined on a neighbourhood of the origin, then the two random variables have the same distribution (see Theorem 11.4.3). A related statement that we will see sooner in class is Corollary 11.1.7.