Liam Hardiman *January 29, 2020 (Late)*

270B - Homework 1

Problem 1. Let X_i be i.i.d. random variables each having the Poisson distribution with mean 1, and consider $S_n = X_1 + \cdots + X_n$. Let $x \in \mathbb{R}$. Show that if $k = k(n)$ is such that $(k-n)/\sqrt{n} \to x$ as $n \to \infty$, we have

$$\sqrt{2\pi n}\mathbb{P}[S_n = k] \to \exp(-x^2/2).$$

Proof. First we claim that S_n has Poisson distribution with mean n. To see this, observe that by independence we have

$$\varphi_{S_n}(t) = \mathbb{E}\left[e^{it(X_1+\cdots+X_n)}\right] = \mathbb{E}\left[e^{itX_1}\right]\cdots\mathbb{E}\left[e^{itX_n}\right] = \varphi_1(t)\cdots\varphi_n(t), \tag{1}$$

where φ_j is the characteristic function of X_j. Now if the random variable X has Poisson distribution with intensity λ, its characteristic function is given by

$$\mathbb{E}\left[e^{itX}\right] = \sum_{k=0}^{\infty} e^{itk} \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \exp(\lambda(e^{it}-1)).$$

Using this, we see that

$$\varphi_{S_n}(t) = \exp(e^{it}-1)^n = \exp(n(e^{it}-1)),$$

which is the characteristic function of the Poisson with intensity λ. Since a distribution is determined by its characteristic function, we conclude that S_n has Poisson distribution with intensity λ.

We use Stirling's approximation.

$$\sqrt{2\pi n}\mathbb{P}[S_n = k] = \sqrt{2\pi n}\frac{n^k e^{-n}}{k!} \sim \sqrt{\frac{n}{k}} \cdot \left(\frac{n}{k}\right)^k \cdot e^{k-n}.$$

Taking the logarithm of both sides gives

$$\log(\sqrt{2\pi n}\mathbb{P}[S_n = k]) \sim \frac{1}{2}\log\frac{n}{k} - k\log\frac{k}{n} + (k-n). \tag{2}$$

We'll need to look at the asymptotic behavior of a few quantities to deal with this expression. Since $\frac{k-n}{\sqrt{n}} \to x$, we must have that $k \sim n$, so the first term in (2) vanishes as $n \to \infty$. Since $k \sim n$, we can Taylor expand the logarithm in the middle term.

$$k\log\frac{k}{n} = k\log\left(1 + \frac{k-n}{n}\right) = k\left(\frac{k-n}{n} - \frac{1}{2}\left(\frac{k-n}{n}\right)^2 + O\left(\left(\frac{k-n}{n}\right)^3\right)\right). \tag{3}$$

We have that $\frac{k-n}{\sqrt{n}} = x + o(1)$. From this it follows that

$$k = n + \sqrt{n}(x + o(1)), \quad \frac{k-n}{n} = \frac{1}{\sqrt{n}}(x + o(1)).$$

Substituting these above gives

$$k\log\frac{k}{n} = (n+\sqrt{n}(x+o(1)))\left(\frac{x+o(1)}{\sqrt{n}} - \frac{1}{2}\frac{(x+o(1))^2}{n} + O(n^{-3/2})\right)$$

$$= \sqrt{n}(x+o(1)) + \frac{1}{2}(x+o(1))^2 - \frac{1}{2}\frac{(x+o(1))^3}{\sqrt{n}} + O(n^{-1/2}).$$

Finally, plugging this all back into (2) gives

$$\log(\sqrt{2\pi n}\mathbb{P}[S_n = k]) \sim -\frac{1}{2}(x+o(1))^2 + O(n^{-1/2}),$$

which limits to $\exp(-x^2/2)$ as desired. □

Problem 2. Find an example of random variables X_n with densities f_n so that X_n converges weakly to the uniform distribution on $[0,1]$ but $f_n(x)$ dos not converge to 1 for any $x \in [0,1]$.

Solution. Let $f_{n,k}(x)$ be a typewriter sequence weighted to have integral 1, that is

$$f_{n,k}(x) = \frac{1}{1-2^{-n}}\mathbb{1}_{[0,1]\setminus[k\cdot 2^{-n},\ (k+1)\cdot 2^{-n}]}(x), \quad n = 1,2,\ldots,\ k = 0,1,\ldots,2^n-1.$$

Intuitively, $f_{n,k}$ is a flat line of height $\frac{1}{1-2^{-n}}$ over $[0,1]$ except for a gap at $[k\cdot 2^{-n}, (k+1)\cdot 2^{-n}]$, where it is zero. Since this gap slides along the unit interval indefinitely, $f_{n,k}$ does not converge pointwise anywhere. Now for any φ bounded and continuous we have

$$|\mathbb{E}_{n,k}[\varphi] - \mathbb{E}[\varphi]| = \int_0^1 \varphi(x)(f_{n,k}(x)-1)dx \leq \|\varphi\|_\infty \cdot 2^{-n} \to 0,$$

so X_n converges weakly to the uniform distribution on $[0,1]$. □

Problem 3. Let X_i be i.i.d. random variables each having exponential distribution with mean 1, and consider $M_n = \max_{i\leq n} X_i$. Show that $M_n - \log n$ converges weakly to the standard Gumbel distribution, i.e. the distribution with cumulative distribution function $F(x) = \exp(-e^{-x})$.

Proof. For any n and t we have

$$\mathbb{P}[M_n - \log n \leq t] = \mathbb{P}[M_n \leq \log n + t] = \mathbb{P}[X_i \leq \log n + t,\ 1 \leq i \leq n].$$

Since the X_i are i.i.d. we can split this into a product

$$\begin{aligned}\mathbb{P}[M_n - \log n \leq t] &= (\mathbb{P}[X_i \leq \log_n + t])^n \\ &= (1 - \exp(-\log n - t))^n \\ &= \left(1 - \frac{e^{-t}}{n}\right)^n \\ &\to \exp(-e^{-t}), \quad \text{as } n \to \infty.\end{aligned}$$

□

Problem 4. Let X_n be random variables and c be a constant. Prove that weak convergence of X_n to c is equivalent to convergence of X_n to c in probability.

Proof. Convergence in probability always implies weak convergence to the same limit, so we only need to show that weak convergence to c implies convergence to c in probability. The distribution function $F(t)$ of the random variable that is constantly c is the shifted Heaviside function $F(t) = H(t - c)$. The only point of discontinuity of F is at $t = a$. By the definition of weak convergence, if F_n is the distribution function of X_n, we have that $F_n(x) \to F(x)$ for all $x \neq c$. For any $\epsilon > 0$ we have

$$\mathbb{P}[|X_n - c| < \epsilon] = F_n(c + \epsilon) - F_n(c - \epsilon) \to F(c + \epsilon) - F(c - \epsilon) = 1,$$

so $X_n \to c$ in probability. $\square$

Problem 5. Consider the following statement:

if $X_n \to X$ weakly and $Y_n \to Y$ weakly then $X_n + Y_n \to X + Y$ weakly.

(a) Find an example showing that this statement if false in general.

Solution. (I talked to Thomas Beardsley and Xiaowen Zhu about this one.) Let X_n be a sequence of i.i.d. $\mathcal{N}(0, 1)$ random variables. Set $Y_{2k} = X_{2k}$, but set $Y_{2k+1} = -X_{2k+1}$. Since the sequence of X_n's are i.i.d., X_n converges weakly to a standard normal random variable. Similarly, since the negative of a standard normal is a standard normal, Y_n also converges weakly to a standard normal. However, the sequence $X_n + Y_n$ doesn't converge weakly at all since every odd term is identically zero and every even term has distribution $\mathcal{N}(0, 4)$. $\square$

(b) Prove that if Y is a constant, then the statement is true.

Proof. We'll show pointwise convergence of the characteristic functions.

$$\begin{aligned}\left|\mathbb{E}\left[e^{it(X_n+Y_n)} - e^{it(X+c)}\right]\right| &= \left|\mathbb{E}\left[e^{it(X_n+Y_n)} - e^{it(X_n+c)} + e^{it(X_n+c)} - e^{it(X+c)}\right]\right| \\ &\leq \mathbb{E}\left|e^{itY_n} - e^{itc}\right| + e^{itc}\mathbb{E}\left[e^{itX_n} - e^{itX}\right].\end{aligned} \tag{4}$$

Let's look at the first term. We split the region of integration like so

$$\mathbb{E}\left|e^{itY_n} - e^{itc}\right| \cdot \mathbb{1}_{|Y_n - c| \geq \epsilon} + \mathbb{E}\left|e^{itY_n} - e^{itc}\right| \cdot \mathbb{1}_{|Y_n - c| < \epsilon}.$$

Since weak convergence to a constant is equivalent to convergence in probability to a constant, the measure of the first region of integration goes to zero as $n \to \infty$, so the first term vanishes. Continuity of the exponential makes the second term small in ϵ. Consequently, the first term on the last line of (4) can be made smaller than ϵ for n large. The second term goes to zero since $X_n \to X$ weakly. Since the characteristic functions converge pointwise, we have weak convergence. $\square$

(c) Prove that if X_n and Y_n are independent then the statement is true.

Proof. We compute the characteristic functions and use independence to split the product.

$$\mathbb{E}\left[e^{it(X_n+Y_n)}\right] = \mathbb{E}\left[e^{itX_n}\right] \cdot \mathbb{E}\left[e^{itY_n}\right]$$

$$\to \mathbb{E}\left[e^{itX}\right] \mathbb{E}\left[e^{itY}\right].$$

Since the characteristic functions converge pointwise, we have weak convergence. ☐

Problem 6.

(a) Prove the following implication: if $X_n \to X$ weakly, $Y_n \geq 0$ and $Y_n \to c$ weakly where c is a constant, then $X_n Y_n \to cX$.

Proof. For any $0 < \epsilon < c$ we split the probability

$$\mathbb{P}[X_n Y_n \leq t] = \mathbb{P}[X_n Y_n \leq t,\ |Y_n - c| < \epsilon] + \mathbb{P}[X_n Y_n \leq t,\ |Y_n - c| \geq \epsilon].$$

Since weak convergence to a constant is equivalent to convergence to that constant in probability, the second term vanishes for n large. Now we also have

$$\mathbb{P}[X_n Y_n \leq t,\ |Y_n - c| < \epsilon] \leq \mathbb{P}[X_n \leq \frac{t}{c-\epsilon},\ |Y_n - c| < \epsilon].$$

Taking the $\liminf_{n\to\infty}$ on both sides, we have that $\liminf n \to \infty \mathbb{P}[X_n Y_n \leq t] \leq \mathbb{P}[X \leq t/(c-\epsilon)]$ at all points of continuity of the distribution of X. Similarly,

$$\mathbb{P}[X_n Y_n \leq t,\ |Y_n - c| < \epsilon] \geq \mathbb{P}[X_n \geq \frac{t}{c+\epsilon},\ |Y_n - c| < \epsilon].$$

Taking the $\limsup_{n\to\infty}$ on both sides, we have that $\limsup_{n\to\infty} \mathbb{P}[X_n Y_n \leq t] \geq \mathbb{P}[X \leq t/(c+\epsilon)]$ at all points of continuity of the distribution of X. We conclude that $X_n Y_n \to cX$ weakly. ☐

(b) Let Z_n be a random vector uniformly distributed on the unit Euclidean sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$. Prove that the distribution of the first coordinate of Z_n converges weakly to the standard normal distribution.

Proof. Let X_n be a standard normal random vector and let $Z_n = X_n \cdot \sqrt{n}/\|X_n\|_2$. $X_n/\|X_n\|_2$ has norm 1, so Z_n is definitely valued on the $\sqrt{n}$-sphere. Since X_n is rotation invariant in distribution, so is Z_n, so we conclude that Z_n is uniformly distributed on the $\sqrt{n}$-sphere.

First we claim that $\sqrt{n}/\|X_n\| \xrightarrow{w} 1$. By Chebyshev we have that (a squared standard normal has mean 1, variance 2 and the components of X are independent)

$$\mathbb{P}\left[\left|\frac{\|X_n\|^2}{n} - 1\right| > \epsilon\right] = \mathbb{P}[|\|X_n\|^2 - n\| > n\epsilon]$$

$$\leq \frac{4n}{n^2\epsilon^2} \to 0.$$

We conclude that $\|X_n\|^2/n$ converges to 1 in probability. If $f \xrightarrow{\mathbb{P}} c \neq 0$, then $1/f \xrightarrow{\mathbb{P}} 1/c$, so $n/\|X_n\|^2$ converges to 1 in probability, and therefore weakly. Since $x \mapsto \sqrt{x}$ is continuous, $\sqrt{n}/\|X_n\| \xrightarrow{w} 1$.

Now the first component of Z_n has distribution $g_n\sqrt{n}/\|X_n\|$, where g_n is a standard normal random variable. g_n converges to a standard normal weakly and we've shown that $\sqrt{n}/\|X_n\| \xrightarrow{w} 1$, so by part (a), the first component of Z_n converges weakly to a standard normal. □

Problem 8. Prov that if ϕ is a characteristic function of some random variable, then $\mathrm{Re}(\phi)$ and $|\phi|^2$ are too.

Proof. Suppose ϕ_X is the characteristic function of the random variable X and let ζ take values ± 1 with probability $1/2$ independent of X. We then have

$$\phi_{X\zeta} = \mathbb{E}[e^{itX\zeta}] = \mathbb{E}[e^{itX} \cdot \mathbb{1}_{\zeta=1} + e^{-itX} \cdot \mathbb{1}_{\zeta=-1}] = \frac{1}{2}(\phi_X + \overline{\phi_X}) = \mathrm{Re}(\phi_X).$$

We conclude that $\mathrm{Re}(\phi_X)$ is the characteristic function of $X\zeta$.

Let the random variable Y have the same distribution as $-X$ independent of X. We then have

$$\phi_{XY} = \mathbb{E}[e^{itXY}] = \mathbb{E}[e^{itX}]\mathbb{E}[e^{-itX}] = |\phi_X|^2.$$

We conclude that $|\phi_X|^2$ is the characteristic function of XY. □

Problem 9. Let X be a random variable with characteristic function ϕ. Prove that for any $a \in \mathbb{R}$ we have

$$\mathbb{P}[X = a] = \lim_{T\to\infty} \frac{1}{2T} \int_T^{-T} e^{-ita}\phi(t)\, dt.$$

Problem 11. Consider independent random variables X_k such that X_k takes values $\pm k$ with probability $k^{-2}/2$ each and values ± 1 with probability $(1-k^{-2})/2$ each. Show that although $\mathrm{Var}[S_n]/n \to 2$, $S_n/\sqrt{n}$ does not converge to $\mathcal{N}(0,1)$ weakly. Why does this not contradict the Lindeberg-Feller central limit theorem?

Proof. If $S_n/\sqrt{n}$ converged to $\mathcal{N}(0,1)$ weakly then its moments would converge to those of a standard normal. Let's look at the fourth moment of $S_n/\sqrt{n}$. Since each X_k has mean zero and they're mutually independent, the only nonvanishing terms in the sum will have the form $X_i^2 X_j^2$ for some i, j (not necessarily distinct).

$$\begin{aligned}
\mathbb{E}[(S_n/\sqrt{n})^4] &= \frac{1}{n^2}\mathbb{E}\left[\sum_{k=1}^{n} X_k^4 + \binom{n}{2}\sum_{1\le i<j\le n} X_i^2 X_j^2\right] \\
&\ge \frac{1}{n^2}\sum_{k=1}^{n}\mathbb{E}[X_k^4] \\
&= \frac{1}{n^2}\sum_{k=1}^{n}\big(k^2 + (1-1/k^2)\big) \\
&= \Theta(n) \to \infty.
\end{aligned}$$

But $\mathcal{N}(0,1)$ has finite moments of all orders. We conclude that $S_n/\sqrt{n}$ does not converge weakly to $\mathcal{N}(0,1)$.

The problem is that this sequence of random variables does not satisfy the Lindeberg condition:

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[X_k^2 \cdot \mathbb{1}_{|X_k|>\sqrt{n}\epsilon}] \to 0$$

for all ϵ. As n grows, the last term contributes $\Theta(n)$ to this sum, so when divided by n the result is $\omega(1)$, which doesn't go to zero. $\square$