270B - Homework 2

Problem 1.

(a) Prove that if the probability density functions of X_n converge pointwise to the probability density function of X, then X_n converges to X weakly.

Proof. Let f_n and f be the density functions of X_n and X, respectively, and let $\varphi : \mathbb{R} \to \mathbb{R}$ be bounded and continuous. We then have

$$|\mathbb{E}_n[\varphi] - \mathbb{E}[\varphi]| = \left| \int_{\mathbb{R}} f_n(x)\varphi(x) \, dx - \int_{\mathbb{R}} f(x)\varphi(x) \, dx \right| \tag{1}$$
$$\leq \|\varphi\|_\infty \int_{\mathbb{R}} |f_n(x) - f(x)| \, dx.$$

Now we claim that $\|f_n - f\|_{L^1} \to 0$ (argument taken from 210 notes). Since $\int f_n = \int f = 1$ we have by Fatou

$$2 \int_{\mathbb{R}} f \, dx = \int_{\mathbb{R}} \liminf_{n\to\infty} (f + f_n - |f_n - f|) \, dx \leq 2 \int_{\mathbb{R}} f \, dx - \limsup_{n\to\infty} |f_n - f| \, dx.$$

In particular, we have $\limsup_{n\to\infty} \int |f_n - f| = 0$, so $\|f_n - f\|_{L^1} \to 0$. Consequently, the right-hand side of (1) goes to zero as $n \to \infty$. We then have that $\mathbb{E}_n[\varphi] \to \mathbb{E}[\varphi]$ for all φ bounded continuous, so $X_n \to X$ weakly. $\qquad\square$

(b) Prove that if the probability mass functions of X_n converge to the probability mass function of X pointwise then X_n converges to X weakly.

Proof. We essentially mirror our proof of part (a) but with the counting measure. Let f_n and f be the mass functions of X_n and X respectively and suppose they take values in the discrete set $S \subset \mathbb{R}$. Suppose φ is a bounded function on S (note that any function from a discrete space is continuous). We then have

$$|\mathbb{E}_n[\varphi] - \mathbb{E}[\varphi]| = \left| \sum_{x\in S} f_n(x)\varphi(x) - \sum_{x\in S} f(x)\varphi(x) \right| \tag{2}$$
$$\leq \|\varphi\|_\infty \sum_{x\in S} |f_n(x) - f(x)|.$$

Fatou still works with the counting measure, so we have

$$2 \sum_{x\in S} f(x) = \sum_{x\in S} \liminf_{n\to\infty} (f(x) + f_n(x) - |f_n(x) - f(x)|) \leq 2 \sum_{x\in S} f(x) - \limsup_{n\to\infty} \sum_{x\in S} |f_n(x) - f(x)|.$$

We conclude that $\sum_{x\in S} |f_n(x) - f(x)| \to 0$ as $n \to \infty$. The right-hand side of (2) vanishes as $n \to \infty$, so $X_n \to X$ weakly. $\qquad\square$

(c) In general, weak convergence does not imply pointwise convergence of probability density functions. Show this by example.

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

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

so X_n converges weakly to the uniform distribution on $[0,1]$ even though the densities don't converge pointwise at all. $\qquad\square$

Problem 2. Consider normal random variables $X_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$. Assume X_n converges weakly to some random variable X. Prove that $X \sim \mathcal{N}(\mu, \sigma^2)$ where $\mu = \lim_{n\to\infty} \mu_n$ and $\sigma^2 = \lim_{n\to\infty} \sigma_n^2$ and both limits exits.

Proof. Since $x \mapsto e^{itx}$ is bounded and continuous, the weak convergence of X_n to X implies that the characteristic functions ϕ_n of X_n converge pointwise to the characteristic function ϕ of X. After completing the square in the exponent, ϕ_n is given by

$$\phi_n(t) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_\mathbb{R} e^{-\frac{1}{2}\left(\frac{x-\mu_n}{\sigma_n}\right)^2} \cdot e^{itx} \ dx$$

$$= e^{i\mu_n t - \sigma_n^2 t^2/2}.$$

As n goes to infinity we have

$$\lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} e^{i\mu_n t - \sigma_n^2 t^2/2} = e^{i\mu t - \sigma^2 t^2/2},$$

which is the characteristic function for the random variable $Y \sim \mathcal{N}(\mu, \sigma^2)$. Since the characteristic function determines the distribution of the random variable, we conclude that $X \sim \mathcal{N}(\mu, \sigma^2)$. $\qquad\square$

Problem 3. Let $X_1, X_2, \ldots$ be independent Rademacher random variables. Let $S_n = X_1 + \cdots + X_n$.

(a) Prove that the sequence $S_n/\sqrt{n}$ is unbounded almost surely.

Proof. (Durrett gives this hint to use Kolmogorov's zero-one). Fix some $K > 0$ and let E_n be the event that $S_n/\sqrt{n} > K$. By continuity of measure we have

$$\mathbb{P}[\limsup_{n\to\infty} E_n] = \mathbb{P}\left[\bigcap_{n\ge 1}\bigcup_{k\ge n} E_k\right] = \lim_{n\to\infty} \mathbb{P}\left[\bigcup_{k\ge n} E_k\right] \ge \limsup_{n\to\infty} \mathbb{P}[E_n] > 0.$$

2

The last inequality follows from the central limit theorem, which says that $\mathbb{P}[S_n/\sqrt{n} > K]$ approaches the probability that a standard normal random variable takes a value greater than K, which is positive. $\limsup_{n\to\infty} E_n$ is a tail event, so by Kolmogorov's zero-one law, it happens with probability zero or one. Since we've shown that its probability is positive, we conclude that $\mathbb{P}[S_n/\sqrt{n} > K \text{ i.o.}] = 1$. Since this holds for any K, $S_n/\sqrt{n}$ is unbounded a.s. □

(b) Prove that $S_n/\sqrt{n}$ does not converge in probability.

Proof. (Durrett gives this hint to set $n = m!$) Suppose $S_n/\sqrt{n}$ did converge in probability. Consider the subsequence $n = m!$. By convergence in probability, we must have for any $\epsilon > 0$

$$\mathbb{P}\left[\left|\frac{S_{(m+1)!}}{\sqrt{(m+1)!}} - \frac{S_{m!}}{\sqrt{m!}}\right| > \epsilon\right] \to 0.$$

Now $S_{m!}$ and $(S_{(m+1)!} - S_{m!})$ are independent since they represent sums over disjoint sets. The plan is to exploit this independence along with the central limit theorem to contradict the above assumption.

By independence and the central limit theorem we have

$$\mathbb{P}\left[\left|\frac{S_{m!}}{\sqrt{m!}}\right| < 1, \ \frac{S_{(m+1)!} - S_{m!}}{\sqrt{(m+1)!}} > 5\right] = \mathbb{P}\left[\left|\frac{S_{m!}}{\sqrt{m!}}\right| < 1\right]\mathbb{P}\left[\frac{S_{(m+1)!} - S_{m!}}{\sqrt{(m+1)!}} > 5\right]$$

$$\to \mathbb{P}[|g| < 1] \cdot \mathbb{P}[g > 5].$$

Note that this limiting quantity is positive. I want to use this to say something like

$$\liminf_{n\to\infty} \mathbb{P}[S_{m+1}/\sqrt{(m+1)!} \text{ and } S_{m!}/\sqrt{m!} \text{ are far apart}] > 0,$$

which contradicts convergence in probability. □

Problem 4. Let $X_1, X_2, \ldots$ be independent random variables such that there exists $M > 0$ so that $|X_i| \leq M$ almost surely for all i. Show that if $\sum_i \text{Var}[X_i] = \infty$ then the sum $S_n = X_1 + \cdots + X_n$ satisfies

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}[S_n]}} \to \mathcal{N}(0, 1) \text{ weakly.}$$

Proof. The plan is to apply the Lindeberg-Feller theorem. To this end, define the sequence $X_{n,m} = \frac{X_m - \mathbb{E}[X_m]}{\sqrt{\text{Var}[S_n]}}$, for $1 \leq m \leq n$. This sequence is centered and we have for all n

$$\sum_{m=1}^{n} \text{Var}[X_{n,m}] = \frac{1}{\text{Var}[S_n]} \sum_{m=1}^{n} \text{Var}[X_m] = 1,$$

so the first condition of Lindeberg-Feller is satisfied. As for the second, we need the following to hold for any $\epsilon > 0$.

$$\lim_{n\to\infty} \sum_{m=1}^{n} \mathbb{E}[|X_{n,m}|^2 : \ |X_{n,m}| > \epsilon] = 0.$$

3

Let's look at the set where $|X_{n,m}| > \epsilon$. This condition is equivalent to

$$|X_m - \mathbb{E}[X_m]| > \epsilon\sqrt{\operatorname{Var}[S_n]}.$$

Since $|X_m| \leq M$ for all m, the left-hand side is at most $2M$. Since $\sum \operatorname{Var}[X_m] = \infty$, the right-hand side blows up in n. In particular, we have that the measure of the set where $|X_{n,m}| > \epsilon$ becomes zero for n large enough, so () is satisfied. Since both conditions of Lindeberg-Feller are satisfied, we have that

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\operatorname{Var}[S_n]}} \to \mathcal{N}(0, 1) \text{ weakly}$$

as desired. $\square$

Problem 5. Let $X_1, X_2, \ldots$ be independent random variables with zero means and unit variances (but not necessarily iid). Assume that

$$\sup_i \mathbb{E}[|X_i|^{2+\delta}] = M < \infty$$

for some $\delta > 0$. Prove that

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \to \mathcal{N}(0, 1) \text{ weakly.}$$

Proof. Similar to the previous exercise, the plan is to use Lindeberg-Feller. Let $S_n = X_1 + \cdots + X_n$ and set $X_{n,m} = \frac{1}{\sqrt{\operatorname{Var}[S_n]}} X_m$ for $1 \leq m \leq n$. Just like in the previous problem we have that $\sum_{m=1}^{n} \operatorname{Var}[X_{m,n}] = 1$ for all n, so the first condition of Lindeberg-Feller is satisfied. As for the second, fix $\epsilon > 0$. We compute.

$$\mathbb{E}[X_{m,n}^2 \ : \ |X_{m,n}| > \epsilon] = \frac{1}{\operatorname{Var}[S_n]} \mathbb{E}[X_m^2 \ : \ |X_m| > \epsilon\operatorname{Var}[S_n]]$$

$$= \frac{1}{n} \mathbb{E}[X_m^{2+\delta} \cdot X_m^{-\delta} \ : \ |X_m| > \epsilon\sqrt{n}]$$

$$\leq \frac{1}{n(\epsilon\sqrt{n})^\delta} \mathbb{E}[X_m^{2+\delta}]$$

$$\leq \frac{M}{n(\epsilon\sqrt{n})^\delta}.$$

Summing over $m = 1, \ldots, n$ gives

$$\sum_{m=1}^{n} \mathbb{E}[X_{m,n}^2 \ : \ |X_{m,n}| > \epsilon] \leq \frac{M}{(n\epsilon)^\delta} \to 0.$$

Hence, the second condition of Lindeberg-Feller is satisfied and $\frac{S_n}{\sqrt{n}} \to \mathcal{N}(0, 1)$ weakly. $\square$

Problem 6. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and let $\mathcal{F} \subset \Sigma$ be a sub-σ-algebra. Consider two events $E \in \Sigma$ and $F \in \mathcal{F}$.

(a) Show that

$$\mathbb{P}[F \mid E] = \frac{\mathbb{E}[\mathbb{P}[E \mid \mathcal{F}] \cdot \mathbb{1}_F]}{\mathbb{E}[\mathbb{P}[E \mid \mathcal{F}]]}.$$

Proof. By the definition of conditional expectation we have

$$\mathbb{P}[F \mid E] = \frac{\mathbb{P}[E \cap F]}{\mathbb{P}[E]}$$

$$= \frac{\mathbb{E}[\mathbb{P}[E|\mathcal{F}] \cdot \mathbb{1}_F]}{\mathbb{E}[\mathbb{P}[E|\mathcal{F}]]}.$$

$\square$

(b) Specialize this equation to the case where $\mathcal{F}$ is generated by a partition $\Omega = F_1 \sqcup \cdots \sqcup F_n$. Deduce Bayes formula:

$$\mathbb{P}[F_i|E] = \frac{\mathbb{P}[E|F_i] \cdot \mathbb{P}[F_i]}{\sum_j (\mathbb{P}[E|F_j] \cdot \mathbb{P}[F_j])}.$$

Proof. This immediately follows from the definition of conditional expectation.

$$\mathbb{P}[F_i \mid E] = \frac{\mathbb{E}[\mathbb{P}[E|\mathcal{F}] \cdot \mathbb{1}_F]}{\mathbb{E}[\mathbb{P}[E|\mathcal{F}]]}$$

$$= \frac{\sum_j \mathbb{P}[E|F_j] \cdot \mathbb{P}[F_i \cap F_j]}{\sum_j (\mathbb{P}[E|F_j] \cdot \mathbb{P}[F_j])}$$

$$= \frac{\mathbb{P}[E|F_i] \cdot \mathbb{P}[F_i]}{\sum_j (\mathbb{P}[E|F_j] \cdot \mathbb{P}[F_j])}.$$

$\square$

Problem 7. Show that

$$(\mathbb{E}[XY|\mathcal{F}])^2 \le \mathbb{E}[X^2|\mathcal{F}] \cdot \mathbb{E}[Y^2|\mathcal{F}]$$

almost surely.

Proof. We prove this in pretty much the same way one proves Cauchy-Schwarz for real or complex numbers. For all real t we have that $\mathbb{E}[(X + tY)^2|\mathcal{F}] \ge 0$ almost surely. Expanding and using the linearity of conditional expectation gives

$$\mathbb{E}[X^2|\mathcal{F}] + 2t\mathbb{E}[XY|\mathcal{F}] + t^2\mathbb{E}[Y^2|\mathcal{F}] \ge 0 \text{ a.s.}$$

Treating this as a real polynomial in t, we conclude that the discriminant must be non-positive, i.e.

$$4\mathbb{E}[XY|\mathcal{F}] - 4\mathbb{E}[X^2|\mathcal{F}]\mathbb{E}[Y^2|\mathcal{F}] \le 0.$$

Dividing by 4 gives the desired result. $\square$

Problem 8. Define the conditional variance of X by

$$\text{Var}[X|\mathcal{F}] = \mathbb{E}[X^2|\mathcal{F}] - (\mathbb{E}[X|\mathcal{F}])^2.$$

Show that

$$\text{Var}[X] = \mathbb{E}\big[\text{Var}[X|\mathcal{F}]\big] + \text{Var}\big[\mathbb{E}[X|\mathcal{F}]\big].$$

Proof. We attack the right-hand side using the law of total expectation.

$$\mathbb{E}[\mathrm{Var}[X|\mathcal{F}]] + \mathrm{Var}[\mathbb{E}[X|\mathcal{F}]] = \mathbb{E}[\mathbb{E}[X^2|\mathcal{F}] - \mathbb{E}[X|\mathcal{F}]^2] + \left(\mathbb{E}[\mathbb{E}[X|\mathcal{F}]^2] - \mathbb{E}[\mathbb{E}[X|\mathcal{F}]]^2\right)$$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$= \mathrm{Var}[X].$$

□

Problem 9. Let $Y = \mathbb{E}[X|\mathcal{F}]$. Show that if $\mathbb{E}[Y^2] = \mathbb{E}[X^2]$ then $X = Y$ a.s.

Proof. $\mathbb{E}[X - Y] = 0$ by definition. The idea is to show that the variance of the difference is zero as well under this assumption.

$$\mathrm{Var}[X - Y] = \mathbb{E}[(X - Y)^2] - \mathbb{E}[X - Y]^2$$

$$= \mathbb{E}[X^2] - 2\mathbb{E}[XY] + \mathbb{E}[Y^2]$$

$$= 2\mathbb{E}[X^2] - 2\mathbb{E}[\mathbb{E}[XY|\mathcal{F}]]$$

$$= 2\mathbb{E}[X^2] - 2\mathbb{E}[Y \cdot \mathbb{E}[X|\mathcal{F}]]$$

$$= 0.$$

Since $X - Y$ has mean and variance zero, $X = Y$ a.s. □