Chapter 7

The Central Limit Theorem

7.25. Let $X_1, \dots, X_n$ be n i.i.d. Poisson variables with mean 1 each. Then $X_1 + \cdots + X_n$ is Poisson with mean n, and so

$$\begin{aligned} e^{-n}\left(1 + n + \frac{n^2}{2} + \frac{n^3}{3!} + \cdots + \frac{n^n}{n!}\right) &= P\{X_1 + \cdots + X_n \leqslant n\} \\ &= P\left\{\sum_{i=1}^n (X_i - EX_i) \leqslant 0\right\} \\ &= P\left\{n^{-1/2}\sum_{i=1}^n (X_i - EX_i) \leqslant 0\right\} \\ &\to P\{N(0\,,1) \leqslant 0\} = \frac{1}{2}, \end{aligned}$$

as $n \to \infty$, thanks to the central limit theorem.

7.30. First let us check that if $Y_n \Rightarrow c$ for a non-random c then $Y_n \to c$ in probability. But this is easy to see, since for all but possibly a countable number of $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{-\varepsilon < Y_n - c \leqslant \varepsilon\} = P\{-\varepsilon \leqslant c - c \leqslant \varepsilon\} = 1.$$

1. Let us prove that if $X_n \Rightarrow X$ and $Y_n \to Y$ in probability then $(X_n\,, Y_n) \Rightarrow (X\,, Y)$. Let f be a bounded, uniformly continuous function of two variables. Then, we can write

$$\begin{aligned} Ef(X_n\,, Y_n) &= E\left[f(X_n\,, Y_n);\ |Y_n - Y| \geqslant \varepsilon\right] + E\left[f(X_n\,, Y_n);\ |Y_n - Y| < \varepsilon\right] \\ &:= T_1 + T_2. \end{aligned}$$

Evidently, $|T_1| \leqslant \sup|f| \times P\{|Y_n - Y| \geqslant \varepsilon\} \to 0$. Since f is uniformly continuous, for all $\eta > 0$ we can choose $\varepsilon > 0$ such that $|f(x\,,y) - f(x\,,z)| \leqslant \eta$ whenever $|y - z| \leqslant \varepsilon$. Thus,

$$\left|T_2 - E\left[f(X_n\,, Y);\ |Y_n - Y| \leqslant \varepsilon\right]\right| \leqslant \eta.$$

Finally,

$$\left| \mathrm{E}\left[f(X_n\,, Y);\, |Y_n - Y| \leqslant \varepsilon\right] - \mathrm{E}f(X_n\,, Y) \right| \leqslant \sup|f| \times \mathrm{P}\{|Y_n - Y| > \varepsilon\} \to 0.$$

Combine our efforts to deduce that

$$\lim_{n\to\infty} |\mathrm{E}f(X_n\,, Y_n) - \mathrm{E}f(X_n\,, Y)| = 0.$$

Since Y is non-random and $X_n \Rightarrow X$, $\mathrm{E}f(X_n\,, Y) \to \mathrm{E}f(X\,, Y)$, and the claim follows.

2. This follows readily from the Mann–Wold device.
3. Let X be a (real-valued) symmetric random variable, and define $X_n = X$ and $Y_n = -X$ for all $n \geqslant 1$. Then X_n has the same distribution as X, and so does Y_n. In particular, $X_n \Rightarrow X$ and $Y_n \Rightarrow X$. However, it is clear that $(X_n\,, Y_n) \not\Rightarrow (X, X)$, unless $\mathrm{P}\{X = 0\} = 1$.

Chapter 8

Martingales

8.1. If $E[X_n \mid \mathcal{G}] \to E[X \mid \mathcal{G}]$ a.s., then for all bounded random variables Z, $E[E(X_n \mid \mathcal{G})Z] \to E[E(X \mid \mathcal{G})Z]$, by the bounded convergence theorem. If Z is, in addition, $\mathcal{G}$-measurable, then for all $Y \in L^1(P)$, $E[Y \mid \mathcal{G}]Z = E[YZ \mid \mathcal{G}]$ a.s., whence the weak convergence in $L^1(P)$ of X_n to X.

For the converse note that $\|E[X_n \mid \mathcal{G}] - E[X \mid \mathcal{G}]\|_1 \leqslant \|X_n - X\|_1$ by conditional Jensen.

8.2. Let X, Y, and Z be three independent random variables, each taking the values ± 1 with probability $\frac{1}{2}$ each. Define $W = X$, $U = X + Y$, and $V = X + Z$. Then, $E[U \mid W] = X + E[Y] = X$ a.s. In particular,

$$\begin{aligned} E\left[E(U \mid W) \mid V\right] = E[X \mid V] &= 1 \quad \text{a.s. on } \{V = 2\} \\ &= -1 \quad \text{a.s. on } \{V = -2\}. \end{aligned} \tag{8.1}$$

On the other hand, $E[X; V = 0] = E[X; X = 1, Z = -1] + E[X; X = -1, Z = 1] = 0$. Therefore, to summarize: $E\{E(U \mid W) \mid V\} = V/2$ a.s. A similar analysis shows that $E[U \mid V] = V/2$ a.s also. Thus,

$$E\left[E(U \mid V) \mid W\right] = \frac{1}{2}E[V \mid W] = \frac{1}{2}X \qquad \text{a.s.}$$

Because $P\{V \neq X\} = 1$, we have produced an example wherein

$$E[E(U \mid V) \mid W] \neq E[E(U \mid W) \mid V] \qquad \text{a.s.}$$

8.3. If $f(x, y) = f_1(x)f_2(y)$, then this is easy. In the general case, estimate $f(x, y)$ by functions of the form $f_1(x)f_2(y)$, and appeal to Exercise **8.1**.

8.7. Because $X \geqslant 0$ a.s., $E[X \mid \mathcal{G}] \geqslant E[X\mathbf{1}_{\{X \geqslant \lambda\}} \mid \mathcal{G}] \geqslant \lambda P(X \geqslant \lambda \mid \mathcal{G})$ a.s.

8.8. Define $f_{X|Y}(x|y) = f(x, y)/f_Y(y)$ where $f_Y(y) = \int_{-\infty}^{\infty} f(a, y)\,da$; this shows that our goal is to prove that $E[h(X) \mid Y] = \int_{-\infty}^{\infty} h(x)f_{X|Y}(x|Y)\,dx$ a.s., which is the classical formula on $\{Y = y\}$.

For any positive, bounded function g,

$$\begin{aligned}
\mathrm{E}\left[h(X)g(Y)\right] &= \int_{-\infty}^{\infty} g(y)\left(\int_{-\infty}^{\infty} h(x)f(x,y)\,dx\right)\,dy \\
&= \int_{-\infty}^{\infty} g(y)f_{Y}(y)\left(\int_{-\infty}^{\infty} h(x)f_{X|Y}(x|y)\,dx\right)\,dy \qquad (8.2)\\
&= \mathrm{E}\left[\Pi(Y)g(Y)\right],
\end{aligned}$$

where $\Pi(y) = \int_{-\infty}^{\infty} h(x)f_{X|Y}(x|y)\,dx$. By a monotone class argument, for all bounded $\sigma(Y)$-measurable Z, $\mathrm{E}[h(X)Z] = \mathrm{E}[\Pi(Y)Z]$. Therefore, $\mathrm{E}[h(X)\,|\,Y] = \Pi(Y)$ a.s., which is the desired result.

Apply (8.2) with $h(x) = \mathbf{1}_{(-\infty,a]}(x)$ to find that

$$\mathrm{P}(X \leqslant a\,|\,Y) = \int_{-\infty}^{a} f_{X|Y}(u|y)\,du,$$

as desired.