Due Mon April 18 in class.
Problems with $(*)$ are required but no need to hand in.

1. $(*)$ Read lecture notes 3.1–3.2.

2. Let $\{X_n\}_{n\in\mathbb{N}}$ be i.i.d. standard exponential random variables. That is, $\mathbb{P}(X_1 > x) = e^{-x}, x > 0$. Show that

$$\max_{i=1,\dots,n} X_i - \log n$$

converges weakly as $n \to \infty$ to the so-called Gumbel distribution, which has cumulative distribution function $F(x) = \exp(-e^{-x}), x \in \mathbb{R}$.

Hint: prove by definition.

3. (Prelim April 2014 (4a)) Let $\{X_n\}_{n\in\mathbb{N}}$ be i.i.d. random variables distributed uniformly in $(0,\theta)$, for some $\theta > 0$. That is, $\mathbb{P}(X_1 \le x) = x/\theta, x \in (0,\theta)$. Set $M_n := \max_{i=1,\dots,n} X_i$. Prove that

$$Z_n := n\left(\theta - \frac{n+1}{n} M_n\right)$$

converges weakly to a non-degenerate distribution, as $n \to \infty$. Identify the limit distribution.

Hint: prove by definition.

4. Suppose $\{X_n\}_{n\in\mathbb{N}}$ are independent random variables with

$$\mathbb{P}(X_n = \pm n) = \frac{1}{2n^2}, \mathbb{P}(X_n = \pm 1) = \frac{1}{2}\left(1 - \frac{1}{n^2}\right), n \in \mathbb{N}.$$

   (a) Show that $\lim_{n\to\infty} \mathrm{Var}(S_n)/n = 2$.

   (b) Show that $S_n/\sqrt{n} \Rightarrow \mathcal{N}(0,1)$. Note that none of the two central limit theorems can be applied directly.

   Hint: Occasionally X_n may take big values $\pm n$: the magnitudes of these values cause the conditions of the central limit theorem to fail, but fortunately such big values disappear eventually (easily seen from first Borel–Cantelli lemma) and hence should not have any impact to weak convergence (since the normalization goes to infinite).

   The idea is again to treat the big (and rare) values separately, by considering $Y_n := X_n \mathbf{1}_{\{|X_n|\le 1\}}$ and $T_n := Y_1 + \cdots + Y_n$.

Comment: this is an example to keep in mind that $Z_n \Rightarrow Z$ does not necessarily imply that $\lim_{n\to\infty} \mathbb{E}|Z_n|^\alpha = \mathbb{E}|Z|^\alpha$ ($\alpha = 2$ in this case).

5. (Prelim May 2015, 3(a)) Let $\{X_n\}_{n\in\mathbb{N}}$ be i.i.d. random variables with probability density function, for some fixed $\theta > 0$,
$$p(x) = \theta(\theta+1)x^{\theta-1}(1-x), 0 < x < 1.$$
Consider
$$T_n := \frac{2S_n/n}{1-S_n/n}, n \in \mathbb{N}.$$
Show that
$$\sqrt{n}(T_n - \theta) \Rightarrow \mathcal{N}(0, \sigma^2(\theta)),$$
and identify the expression of $\sigma^2(\theta)$.

Hint: first verify $\mathbb{E}X_1 = \theta/(\theta+2)$. Then $\sqrt{n}(T_n-\theta)$ can be rewritten in the form of
$$\sqrt{n}(T_n - \theta) = \frac{S_n - \mathbb{E}S_n}{\sqrt{n}} \cdot Z_n$$
for some random variable Z_n.

6. (Prelim August 2015, 5) Let $\{U_n\}_{n\in\mathbb{N}}$ be a collection of i.i.d. random variables with $\mathbb{E}U_n = 0$ and $\mathbb{E}U_n^2 = \sigma^2 \in (0,\infty)$. Consider random variables $\{X_n\}_{n\in\mathbb{N}}$ defined by $X_n := U_n + U_{2n}, n \in \mathbb{N}$, and the partial sum $S_n = X_1 + \cdots + X_n$. Find appropriate constants $\{a_n, b_n\}_{n\in\mathbb{N}}$ such that
$$\frac{S_n - b_n}{a_n} \Rightarrow \mathcal{N}(0,1).$$

7. Let $\{X_n\}_{n\in\mathbb{N}}$ be i.i.d. random variables with $\mathbb{P}(X_1 = \pm 1) = 1/2$. Show that, for all $a, b \in \mathbb{R}\setminus\{0\}$, there exists a non-negative number $\sigma^2(a,b)$, such that
$$\frac{aS_n + bS_{2n}}{\sqrt{n}} \Rightarrow \mathcal{N}(0, \sigma^2(a,b)).$$
Provide the explicit expression of $\sigma^2(a,b)$.

Hint: you may choose from one of the following two approaches. The standard approach is to apply the Lindeberg–Feller central limit theorem. Or, you may use the following version of continuous mapping theorem: if for each $n \in \mathbb{N}$, X_n, Y_n are *independent* random variables, and $X_n \Rightarrow X, Y_n \Rightarrow Y$ where X and Y are independent *continuous* random variables, then $X_n + Y_n \Rightarrow X + Y$.

Comment: the result to be established implies that
$$\left(\frac{S_n}{\sqrt{n}}, \frac{S_{2n}}{\sqrt{n}}\right)$$
as a random vector converges weakly to a bivariate Gaussian random vector, say (Z_1, Z_2). The standard central limit theorem yields $S_n/\sqrt{n} \Rightarrow Z_1$ and $S_{2n}/\sqrt{n} \Rightarrow Z_2$ immediately, with explicit variances. The result to be established here, however, is a bivariate central limit theorem. It is a stronger result as it characterizes the joint dependence. In particular, Z_1 and Z_2 are not independent, and the expression $\sigma^2(a,b)$ to be derived shall tell the covariance between them.