Lecture 13

1 The Central Limit Theorem

As suggested by its name, the following theorem is a major result in Probability Theory.

Theorem 1.1 (the Central Limit Theorem)**.** *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of mutually independent identically distributed random variables with finite expectation μ and finite variance $\sigma^2 > 0$. For every positive integer n, let $Y_n = \frac{X_1+\ldots+X_n-n\mu}{\sigma\sqrt{n}}$ and let F_n be the cumulative distribution function of Y_n, that is, $F_n(a) = \mathbb{P}(Y_n \leq a)$ for every $a \in \mathbb{R}$. Then $\lim_{n\to\infty} F_n(a) = \Phi(a)$ holds for every $a \in \mathbb{R}$.*

Example 1: Toss a fair coin 1000 times, all coin tosses being mutually independent. Let X be the total number of coin tosses (of these 1000 tosses) whose outcome is heads. In Lecture 1 we used Chebyshev's inequality and Chernoff's inequality to bound from below the probability of the event $450 < X < 550$. Indeed, using Chebyshev's inequality we proved that

$$\mathbb{P}(450 < X < 550) \geq 0.9$$

and using Chernoff's inequality we proved that

$$\mathbb{P}(450 < X < 550) \geq 1 - 2e^{-5} \approx 0.98652.$$

We will now use the Central Limit Theorem (Theorem 1.1) to approximate this probability. For every $1 \leq i \leq 1000$, let $X_i = 1$ if the outcome of the ith coin toss is heads and $X_i = 0$ otherwise. Note that the X_i's are mutually independent identically distributed random variables. Moreover, $\mathbb{E}(X_i) = 1/2$ and $Var(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 1/4$ for every $1 \leq i \leq 1000$. Let $Y = \frac{X-500}{\sqrt{250}}$. It then follows by Theorem 1.1 that

$$\begin{aligned}
\mathbb{P}(450 < X < 550) &= \mathbb{P}\left(\frac{450 - 1000 \cdot 1/2}{\sqrt{1/4} \cdot \sqrt{1000}} < \frac{X - 1000 \cdot 1/2}{\sqrt{1/4} \cdot \sqrt{1000}} < \frac{550 - 1000 \cdot 1/2}{\sqrt{1/4} \cdot \sqrt{1000}}\right) \\
&= \mathbb{P}\left(-\frac{50}{\sqrt{250}} < Y < \frac{50}{\sqrt{250}}\right) = \mathbb{P}\left(Y < \sqrt{10}\right) - \mathbb{P}\left(Y \leq -\sqrt{10}\right) \\
&\approx \Phi(\sqrt{10}) - \Phi(-\sqrt{10}) = 2\Phi(\sqrt{10}) - 1 \approx 0.9984. \qquad (1)
\end{aligned}$$

Note that the approximation of $\mathbb{P}(450 < X < 550)$ obtained from the Central Limit Theorem is better than the ones obtained from Chebyshev's inequality and from Chernoff's inequality. Moreover, it is an approximation, i.e., a lower and an upper bound rather than just a lower bound. On the other hand, it is not clear how good this approximation really is as the Central Limit Theorem is a qualitative rather than quantitative result (similarly to the laws of large numbers and unlike concentration inequalities such as Chebyshev's inequality and Chernoff's inequality). That is, in order to approximate probabilities stemming from small values of n (such as 1000) we actually need a quantitative version of the Central Limit Theorem. There are several such theorems. Below, we state one of them.

Theorem 1.2 (Berry-Esseen Theorem). *Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of mutually independent identically distributed random variables. Suppose that $\mathbb{E}(X_i) = 0$, $Var(X_i) = \sigma^2 > 0$ and $\mathbb{E}(|X_i|^3) = \rho$ hold for every $i \in \mathbb{N}$, where σ and ρ are real numbers. For every positive integer n, let $Y_n = \frac{X_1+\ldots+X_n}{\sigma\sqrt{n}}$ and let F_n be the cumulative distribution function of Y_n, that is, $F_n(a) = \mathbb{P}(Y_n \leq a)$ for every $a \in \mathbb{R}$. Then $|F_n(a) - \Phi(a)| \leq \frac{\rho}{\sigma^3\sqrt{n}}$ holds for every $a \in \mathbb{R}$ and every positive integer n.*

We will now use Theorem 1.2 to better quantify the approximation of $\mathbb{P}(450 < X < 550)$, given in Example 1. For every $1 \leq i \leq 1000$ let $Z_i = X_i - 1/2$, and note that $\mathbb{E}(Z_i) = 0$. Moreover, $Var(Z_i) = Var(X_i) = 1/4$ and thus $\sigma := \sqrt{Var(Z_i)} = 1/2$. Finally, $\mathbb{P}(|Z_i| = 1/2) = 1$ and thus $\mathbb{E}(|Z_i|^3) = (1/2)^3 = 1/8$. Let $Y = \frac{Z_1+\ldots+Z_{1000}}{\sqrt{1/4}\cdot\sqrt{1000}} = \frac{X-500}{\sqrt{250}}$. Theorem 1.2 implies that

$$\mathbb{P}\left(Y \leq \sqrt{10}\right) \geq \Phi(\sqrt{10}) - \frac{1/8}{(1/2)^3 \cdot \sqrt{1000}} \geq \Phi(\sqrt{10}) - 0.03163 \tag{2}$$

and

$$\mathbb{P}\left(Y \leq -\sqrt{10}\right) \leq \Phi(-\sqrt{10}) + \frac{1/8}{(1/2)^3 \cdot \sqrt{1000}} \leq \Phi(-\sqrt{10}) + 0.03163. \tag{3}$$

Similarly to (1), it follows by (2) and (3) that

$$\begin{aligned}
\mathbb{P}(450 < X < 550) &= \mathbb{P}\left(\frac{450 - 1000\cdot 1/2}{\sqrt{1/4}\cdot\sqrt{1000}} < \frac{X - 1000\cdot 1/2}{\sqrt{1/4}\cdot\sqrt{1000}} < \frac{550 - 1000\cdot 1/2}{\sqrt{1/4}\cdot\sqrt{1000}}\right) \\
&= \mathbb{P}\left(-\frac{50}{\sqrt{250}} < Y < \frac{50}{\sqrt{250}}\right) = \mathbb{P}\left(Y < \sqrt{10}\right) - \mathbb{P}\left(Y \leq -\sqrt{10}\right) \\
&\geq \left[\Phi(\sqrt{10}) - 0.03163\right] - \left[\Phi(-\sqrt{10}) + 0.03163\right] \\
&= 2\Phi(\sqrt{10}) - 1 - 0.06326 \geq 0.935.
\end{aligned}$$

This is still better than the lower bound we obtained using Chebyshev's inequality, but not as good as the one we obtained using Chernoff's inequality.

Example 2: Roll a fair die 360000 times, all die rolls being mutually independent. Let X be the random variable counting the number of die rolls whose outcome is 6. We would like to find real numbers a and b for which

$$\mathbb{P}(54000 \leq X \leq 63000) \approx \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx.$$

For every $1 \leq i \leq 360000$, let X_i be the indicator random variable for the event "the outcome of the ith die roll is 6", that is, $X_i = 1$ if the outcome of the ith die roll is 6 and $X_i = 0$ otherwise. Then $X = \sum_{i=1}^{360000} X_i$. For every $1 \leq i \leq 360000$, it holds that $\mathbb{E}(X_i) = \mathbb{P}(X_i = 1) = 1/6$ and $Var(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 1/6 - 1/36 = 5/36$. It thus follows by the Central Limit Theorem that

$$\begin{aligned}
\mathbb{P}(54000 \leq X \leq 63000) &= \mathbb{P}\left(\frac{54000 - 360000 \cdot 1/6}{\sqrt{5/36} \cdot \sqrt{360000}} \leq \frac{X - 360000 \cdot 1/6}{\sqrt{5/36} \cdot \sqrt{360000}} \leq \frac{63000 - 360000 \cdot 1/6}{\sqrt{5/36} \cdot \sqrt{360000}}\right) \\
&= \mathbb{P}\left(-12\sqrt{5} \leq \frac{X - 360000 \cdot 1/6}{\sqrt{5/36} \cdot \sqrt{360000}} \leq 6\sqrt{5}\right) \\
&\approx \frac{1}{\sqrt{2\pi}} \int_{-12\sqrt{5}}^{6\sqrt{5}} e^{-x^2/2} dx.
\end{aligned}$$

We conclude that $a = -12\sqrt{5}$ and $b = 6\sqrt{5}$ satisfy the desired property.

Example 3: We will use the Central Limit Theorem to prove that

$$\lim_{n \to \infty} e^{-n} \left(1 + n + \frac{n^2}{2!} + \ldots + \frac{n^n}{n!}\right) = \frac{1}{2}.$$

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of mutually independent random variables, where $X_i \sim Poi(1)$ for every $i \in \mathbb{N}$. For every positive integer n, let $Y_n = \sum_{i=1}^{n} X_i$. Then, as was shown in Probability Theory 1, $Y_n \sim Poi(n)$ for every $n \in \mathbb{N}$. In particular,

$$\mathbb{P}(Y_n \leq n) = \sum_{j=0}^{n} \mathbb{P}(Y_n = j) = e^{-n} \left(1 + n + \frac{n^2}{2!} + \ldots + \frac{n^n}{n!}\right). \tag{4}$$

Moreover, it follows by the Central Limit Theorem that

$$\lim_{n \to \infty} \mathbb{P}(Y_n \leq n) = \lim_{n \to \infty} \mathbb{P}(Y_n - n \leq 0) = \lim_{n \to \infty} \mathbb{P}\left(\frac{Y_n - n}{\sqrt{n}} \leq 0\right) = \Phi(0) = 1/2. \tag{5}$$

Combining (4) and (5) shows that

$$\lim_{n \to \infty} e^{-n} \left(1 + n + \frac{n^2}{2!} + \ldots + \frac{n^n}{n!}\right) = \lim_{n \to \infty} \mathbb{P}(Y_n \leq n) = \frac{1}{2}$$

as claimed