Assignment 5 - Key: Central Limit Theorem

Question 1: The number of accidents in a certain city is modeled by a Poisson random variable with average rate of 10 accidents per day. Suppose that the number of accidents in different days are independent. Use the central limit theorem to find the probability that there will be more than 3800 accidents in a certain year. Assume that there are 365 days in a year.

Solution:

$$\begin{aligned}
Y &= X_1 + X_2 + \cdots + X_n, \quad n = 365 \\
X_i &\sim Poisson(\lambda = 10). \quad \text{Thus:} \quad EX_i = 10 \\
\text{Var}(X_i) &= \lambda = 10 \\
EY &= 365 \times 10 = 3650 \\
\text{Var}(Y) &= 365 \times 10 = 3650 \\
\frac{Y - 3650}{\sqrt{3650}} & \text{ is approximately} \quad N(0,1) \quad \text{(by the CLT)} \\
P(Y \geq 3800) &= P\left(\frac{Y - 3650}{\sqrt{3650}} \geq \frac{3800 - 3650}{\sqrt{3650}}\right) \\
&= 1 - \Phi\left(\frac{3800 - 3650}{\sqrt{3650}}\right) \\
&\approx 1 - \Phi(2.48) \\
&\approx 0.0065
\end{aligned}$$

Question 2: In a communication system, each code-word consists of 1000 bits. Due to the noise, each bit may be received in error with probability 0.1. It is assumed bit errors occur independently. Since error correcting codes are used in this system, each code-word can be decoded reliably if there are less than or equal to 125 errors in the received code-word, otherwise the decoding fails. Using the CLT, find the probability of decoding failure.

Solution: Let $Y = X_1 + X_2 + \cdots + X_n, \quad n = 1000.$

$$\begin{aligned}
X_i &\sim \quad \text{Bernoulli}(p = 0.1) \\
EX_i &= p = 0.1 \\
\text{Var}(X_i) &= p(1-p) = 0.09 \\
EY &= np = 100 \\
\text{Var}(Y) &= np(1-p) = 90
\end{aligned}$$

By the CLT:

$$\frac{Y - EY}{\sqrt{\text{Var}(Y)}} = \frac{Y - 100}{\sqrt{90}} \quad \text{(can be approximated by} \quad N(0,1)). \quad \text{Thus,}$$

$$\begin{aligned}
P(Y > 125) &= P\left(\frac{Y-100}{\sqrt{90}} > \frac{125-100}{\sqrt{90}}\right) \\
&= 1 - \Phi\left(\frac{25}{\sqrt{90}}\right) \\
&\approx 0.0042
\end{aligned}$$

Question 3: The amount of time needed for a certain machine to process a job is a random variable with mean E(Xi) = 10 minutes and Var(Xi) = 2 minutes2. The time needed for different jobs are independent from each other. Find the probability that the machine processes less than or equal to 40 jobs in 7 hours.

Solution:

$$\begin{aligned}
Y &= X_1 + X_2 + \cdots + X_{40} \\
EX_i &= 10, \text{Var}(X_i) = 2 \\
EY &= 40 \times 10 = 400 \\
\text{Var}(Y) &= 40 \times 2 = 80 \\
P(\text{Less than or equal to 40 jobs in 7 hours}) &= P(Y > 7 \times 60) \\
&= P(Y > 420) \\
&= P\left(\frac{Y-400}{\sqrt{80}} > \frac{420-400}{\sqrt{80}}\right) \\
&\approx 1 - \Phi\left(\frac{20}{\sqrt{80}}\right) \approx 0.0127
\end{aligned}$$

Question 4: You have a fair coin. You toss the coin n times. Let X be the portion of times that you observe heads. How large n has to be so that you are 95% sure that $0.45 \le X \le 0.55$? In other words, how large n has to be so that

$P(0.45 \le X \le 0.55) \ge 0.95$?

Solution:

$$\begin{aligned}
X &= \frac{X_1 + X_2 + \cdots + X_n}{n} \\
&= \frac{Y}{n} \quad \text{where} \quad Y = X_1 + X_2 + \cdots + X_n \\
X_i &\sim Bernoulli\left(\frac{1}{2}\right) \\
EX_i &= \frac{1}{2} \\
\mathrm{Var}(X_i) &= \frac{1}{4} \\
EY &= \frac{n}{2} \\
\mathrm{Var}(Y) &= \frac{n}{4} \\
P(0.45n \le Y \le 0.55n) &= P\left(\frac{0.45n - 0.5n}{\frac{\sqrt{n}}{2}} \le \frac{Y - 0.5n}{\frac{\sqrt{n}}{2}} \le \frac{0.55n - 0.5n}{\frac{\sqrt{n}}{2}}\right) \\
&\approx \Phi(0.1\sqrt{n}) - \Phi(-0.1\sqrt{n}) = 0.95 \\
2\Phi(0.1\sqrt{n}) - 1 &= 0.95 \\
\Phi(0.1\sqrt{n}) &= 0.975 \\
0.1\sqrt{n} &\approx 1.96 \\
n &\ge 385
\end{aligned}$$