2023.2 Question 11

1. For some $1 \leq i \leq n$, we have

$$\begin{aligned}
\mathrm{P}(Y = x_i) &= \mathrm{P}(Y = X_i, Y = X_1) + \mathrm{P}(Y = X_i, Y = X_2) \\
&= \mathrm{P}(Y = x_i \mid Y = X_1) \cdot \mathrm{P}(Y = X_1) + \mathrm{P}(Y = x_i \mid Y = X_2) \cdot \mathrm{P}(Y = X_2) \\
&= \mathrm{P}(X_1 = x_i) \cdot \mathrm{P}(Y = X_1) + \mathrm{P}(X_2 = X_i) \cdot \mathrm{P}(Y = x_2) \\
&= pa_i + qb_i.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{E}(Y) &= \sum_{i=1}^{n} x_i \, \mathrm{P}(Y = x_i) \\
&= \sum_{i=1}^{n} x_i (pa_i + qb_i) \\
&= p \sum_{i=1}^{n} x_i a_i + q \sum_{i=1}^{n} x_i b_i \\
&= p \, \mathrm{E}(X_1) + q \, \mathrm{E}(X_2) \\
&= p\mu_1 + q\mu_2.
\end{aligned}$$

For the variance, we have

$$\begin{aligned}
\mathrm{E}(Y^2) &= \sum_{i=1}^{n} x_i^2 \, \mathrm{P}(Y = x_i) \\
&= \sum_{i=1}^{n} x_i^2 (pa_i + qb_i) \\
&= p \sum_{i=1}^{n} x_i^2 a_i + q \sum_{i=1}^{n} x_i^2 b_i \\
&= p \, \mathrm{E}(X_1^2) + q \, \mathrm{E}(X_2^2) \\
&= p \left(\mathrm{E}(X_1)^2 + \mathrm{Var}(X_1) \right) + q \left(\mathrm{E}(X_2)^2 + \mathrm{Var}(X_2) \right) \\
&= p \left(\mu_1^2 + \sigma_1^2 \right) + q \left(\mu_2^2 + \sigma_2^2 \right),
\end{aligned}$$

and hence

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 \\
&= p \left(\mu_1^2 + \sigma_1^2 \right) + q \left(\mu_2^2 + \sigma_2^2 \right) - (p\mu_1 + q\mu_2)^2 \\
&= p\sigma_1^2 + q\sigma_2^2 + p\mu_1^2 + q\mu_2^2 - p^2\mu_1^2 - q^2\mu_2^2 - 2pq\mu_1\mu_2 \\
&= p\sigma_1^2 + q\sigma_2^2 + p(1-p)\mu_1^2 + q(1-q)\mu_2^2 - 2pq\mu_1\mu_2 \\
&= p\sigma_1^2 + q\sigma_2^2 + pq\mu_1^2 + pq\mu_2^2 - 2pq\mu_1\mu_2 \\
&= p\sigma_1^2 + q\sigma_2^2 + pq \left(\mu_1 - \mu_2 \right)^2,
\end{aligned}$$

as desired.

2. We have

$$\mathrm{P}(B = 1) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{5}{6} = \frac{1}{2}.$$

Z_1 is the sum of n independent values of B, and counts the number of times when $B = 1$.
Hence, $Z_1 \sim \mathrm{B}\left(n, \frac{1}{2}\right)$.
Since $n \gg 1$, we have

$$Z_1 \sim \mathrm{B}\left(n, \frac{1}{2}\right) \dot{\sim} \mathrm{N}\left(\frac{n}{2}, \frac{n}{4}\right).$$

The probability of Z_1 being within 10 percent of its mean is given by

$$\begin{aligned}
\mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq Z_1\leq\frac{n}{2}+\frac{n}{20}\right) &= \mathrm{P}\left(-\frac{\frac{n}{20}}{\frac{\sqrt{n}}{2}}\leq Z\leq\frac{\frac{n}{20}}{\frac{\sqrt{n}}{2}}\right) \\
&= \mathrm{P}\left(-\frac{\sqrt{n}}{20}\leq Z\leq\frac{\sqrt{n}}{20}\right)
\end{aligned}$$

where $Z\sim\mathrm{N}(0,1)$ is the standard normal.

As $n\to\infty$, $-\frac{\sqrt{n}}{20}\to-\infty$, and $\frac{\sqrt{n}}{20}\to\infty$, and so the probability approaches $\mathrm{P}(-\infty<Z<\infty)$ which is 1.

3. Let $X_1\sim\mathrm{B}\left(n,\frac{1}{6}\right)$, and $X_2\sim\mathrm{B}\left(n,\frac{5}{6}\right)$. Z_2 has $\frac{1}{2}$ chance of taking X_1 and $\frac{1}{2}$ chance of taking X_2.

   We have $\mu_1=\frac{n}{6},\mu_2=\frac{5n}{6},\sigma_1^2=\sigma_2^2=\frac{5n}{36}$.

   Hence,

$$\mathrm{E}(Z_2)=\frac{1}{2}\cdot\frac{n}{6}+\frac{1}{2}\cdot\frac{5n}{6}=\frac{n}{2},$$

   and

$$\mathrm{Var}(Z_2)=\frac{1}{2}\cdot\frac{5n}{36}+\frac{1}{2}\cdot\frac{5n}{36}+\frac{1}{4}\left(\frac{n}{6}-\frac{5n}{6}\right)^2=\frac{n^2}{9}+\frac{5n}{36}.$$

   A normal approximation will not be a good approximation since in this case, Z_2 is bimodal – it is likely to take values close to $\frac{n}{6}$ or $\frac{5n}{6}$, but not near the mean $\frac{n}{2}$.

   The bounds within 10 percent of the mean is $\frac{n}{2}\pm\frac{n}{20}$. We have

$$\begin{aligned}
\mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq Z_2\leq\frac{n}{2}+\frac{n}{20}\right) &= \frac{1}{2}\mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq X_1\leq\frac{n}{2}+\frac{n}{20}\right)+\frac{1}{2}\mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq X_2\leq\frac{n}{2}+\frac{n}{20}\right) \\
&= \frac{1}{2}\mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq X_1\right)+\frac{1}{2}\mathrm{P}\left(X_2\leq\frac{n}{2}+\frac{n}{20}\right) \\
&= \mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq X_1\right).
\end{aligned}$$

   Since n is large, we have $X_1\sim\mathrm{B}\left(n,\frac{1}{6}\right)\dot\sim\mathrm{N}\left(\frac{n}{6},\frac{5n}{36}\right)$, and hence

$$\begin{aligned}
\mathrm{P}\left(\frac{n}{2}-\frac{n}{20}\leq X_1\right) &= \mathrm{P}\left(Z\geq\frac{\frac{n}{2}-\frac{n}{20}-\frac{n}{6}}{\frac{\sqrt{5n}}{6}}\right) \\
&= \mathrm{P}\left(Z\geq\frac{30n-3n-10n}{10\sqrt{5n}}\right) \\
&= \mathrm{P}\left(Z\geq\frac{17\sqrt{n}}{10\sqrt{5}}\right),
\end{aligned}$$

   and as $n\to\infty$, $\frac{17\sqrt{n}}{10\sqrt{5}}\to\infty$, and hence the probability tends to 0, as desired.