MH2500 AY19/20

Solution 1.

(a)

$$\frac{\binom{3}{3}}{\binom{7}{3}} = \frac{1}{35}.$$

(b)

$$\begin{aligned}
&P(\text{remaining all blue}|\text{drawn all red})\\
&= P(i = 3|\text{drawn all red})\\
&= \frac{P(\text{drawn all red}|i = 3)P(i = 3)}{P(\text{drawn all red}|i = 3)P(i = 3) + P(\text{drawn all red}|i = 2)P(i = 2) + P(\text{drawn all red}|i = 1)P(i = 1)}\\
&= \frac{1/35}{\binom{3}{1}/\binom{7}{1} + \binom{3}{2}/\binom{7}{2} + \binom{3}{3}/\binom{7}{3}} = \frac{1}{21}.
\end{aligned}$$

Solution 2.

(a) X is uniformly distributed in $(-1,1)$. $f_X(x) = \frac{1}{2}$, if $x \in (-1,1)$.

(b) For a given $X = x$,

$$|x\cos\theta - y\sin\theta| < 1 \implies \frac{x\cos\theta - 1}{\sin\theta} < y < \frac{x\cos\theta + 1}{\sin\theta}.$$

Since X, Y are uniformly distributed in D_θ, therefore

$$Y|X = x \sim \text{Unif}\left(\frac{x\cos\theta - 1}{\sin\theta}, \frac{x\cos\theta + 1}{\sin\theta}\right)$$

The conditional density function is

$$f_{Y|X}(y|x) = \frac{1}{\frac{x\cos\theta+1}{\sin\theta} - \frac{x\cos\theta-1}{\sin\theta}} = \frac{\sin\theta}{2}, \qquad \text{for } y \in \left(\frac{x\cos\theta - 1}{\sin\theta}, \frac{x\cos\theta + 1}{\sin\theta}\right)$$

(c)

$$\mathbb{E}(Y|X) = \int_{\frac{x\cos\theta-1}{\sin\theta}}^{\frac{x\cos\theta+1}{\sin\theta}} y \cdot \frac{\sin\theta}{2} dy = \frac{x}{\tan\theta}.$$

(d)

$$\mathbb{E}(Y) = \mathbb{E}_X\mathbb{E}(Y|X) = \mathbb{E}\left(\frac{X}{\tan\theta}\right) = \frac{1}{\tan\theta}\mathbb{E}(X) = 0.$$

(e)

$$\mathbb{E}(Y^2) = \mathbb{E}_X\mathbb{E}_Y(Y^2|X) = \int_{-1}^{1} f_X(x) \int_{\frac{x\cos\theta-1}{\sin\theta}}^{\frac{x\cos\theta+1}{\sin\theta}} y^2 \frac{\sin\theta}{2} dydx = \frac{\sin\theta}{4}\int_{-1}^{1}\int_{\frac{x\cos\theta-1}{\sin\theta}}^{\frac{x\cos\theta+1}{\sin\theta}} y^2 dydx = \frac{1+\cos^2\theta}{3\sin^2\theta}.$$

Therefore, $Var(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \frac{1+\cos^2\theta}{3\sin^2\theta}$.

(f) Since $\mathbb{E}(Y) = 0$

$$Cov(X,Y) = \mathbb{E}(XY) = \mathbb{E}_X\mathbb{E}(XY|X) = \mathbb{E}_X(X\mathbb{E}(Y|X)) = \mathbb{E}_X\left(X \cdot \frac{X}{\tan\theta}\right) = \frac{1}{3\tan\theta} > 0.$$

X, Y are positively correlated.

Solution 3.

(a)
$$P(X_2 = k) = \binom{n}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{n-k}.$$

(b)
$$P(X_2 = k, X_4 = \ell) = \frac{n!}{k!\ell!(n-k-\ell)!}\left(\frac{1}{6}\right)^k\left(\frac{1}{6}\right)^\ell\left(\frac{4}{6}\right)^{n-k-\ell}.$$

(c)
$$P(X_2 = k | X_4 = \ell) = \frac{(n-\ell)!}{k!(n-k-\ell)!}\left(\frac{1}{5}\right)^k\left(\frac{4}{5}\right)^{n-k-\ell}.$$

(d)
$$P(X_2 + X_4 = k) = \binom{n}{k}\left(\frac{2}{6}\right)^k\left(\frac{4}{6}\right)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

(e) Let W_i, Y_i be the indicator variables that the i-th roll results in number 2 and 4 respectively. Therefore,

$$X_2 = \sum_{i=1}^{n} W_i, \qquad X_4 = \sum_{i=1}^{n} Y_i$$

$$\begin{aligned}
Cov(X_2, X_4) &= Cov\left(\sum_{i=1}^{n} W_i, \sum_{j=1}^{n} Y_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(W_i, Y_j) \\
&= \sum_{i=1}^{n} Cov(W_i, Y_i) + \sum_{i \neq j} Cov(W_i, Y_j)
\end{aligned}$$

For $i \neq j$, W_i and Y_j are independent, hence $Cov(W_i, Y_j) = 0$.

$$\begin{aligned}
Cov(W_i, Y_i) &= \mathbb{E}(W_i Y_i) - \mathbb{E}(W_i)\mathbb{E}(Y_i) \\
&= P(i\text{-th throw} = 2 \cap i\text{-th throw} = 4) - P(i\text{-th throw} = 2)P(i\text{-th throw} = 4) = -\frac{1}{36}
\end{aligned}$$

Therefore, $Cov(X_2, X_4) = -\frac{n}{36}$.

Solution 4.

(a)

$$\mathbb{E}\left(\sum_{i=1}^{2m} b_{m,i}Y_i\right) = \mathbb{E}\left(\sum_{i=1}^{m} b_{m,i}Y_i\right) + \mathbb{E}\left(\sum_{i=m+1}^{2m} b_{m,i}Y_i\right) = \sum_{i=1}^{m}\mathbb{E}\left(Y_i\right) - \sum_{i=m+1}^{2m}\mathbb{E}\left(Y_i\right) = 0.$$

$$Var\left(\sum_{i=1}^{2m} b_{m,i}Y_i\right) = \sum_{i=1}^{m} b_{m,i}^2 Var\left(Y_i\right) + \sum_{i=m+1}^{2m} b_{m,i}^2 Var\left(Y_i\right) = \sum_{i=1}^{2m} Var\left(Y_i\right) = 2m\cdot\frac{1}{4} = \frac{m}{2}.$$

By Chebyshev's inequality,

$$P\left(\left|\sum_{i=1}^{2m} b_{m,i}Y_i\right| > c_1\sqrt{m}\right) \leq \frac{m/2}{c_1^2 m} = \frac{1}{2c_1^2} = \frac{1}{10}.$$

We can set $c_1 \geq \sqrt{5}$.

(b) Observe that

$$\sum_{i=1}^{2m} b_{m,i}Y_i = \sum_{i=1}^{m} Y_i - \sum_{i=m+1}^{2m} Y_i = \sum_{i=1}^{m}(Y_i - Y_{m+i})$$

Let $W_i = Y_i - Y_{m+i}$. $\mathbb{E}(W_i) = 0$ and $Var(W_i) = Var(Y_i) + Var(Y_{m+i}) = \frac{1}{2}$. By central limit theorem, $\frac{1}{m}\sum_{i=1}^{m} W_i$ converges in distribution to $\mathrm{N}(0, \frac{1}{2m})$. Therefore,

$$\lim_{m\to\infty} P\left(\left|\sum_{i=1}^{2m} b_{m,i}Y_i\right| > c_2\sqrt{m}\right) = \lim_{m\to\infty} P\left(\left|\frac{1}{m}\sum_{i=1}^{m} W_i\right| > \frac{c_2}{\sqrt{m}}\right) = P\left(|Z| > \frac{\frac{c_2}{\sqrt{m}}}{\sqrt{\frac{1}{2m}}}\right) = P(|Z| > \sqrt{2}c_2),$$

where Z is the standard normal distribution. We now find c_2, such that $P(|Z| > \sqrt{2}c_2) > 0.9$. Let $\Phi(\cdot)$ be the CDF of Z, then

$$P(|Z| > \sqrt{2}c_2) = 2 - 2\Phi(\sqrt{2}c_2) > 0.9 \implies \text{choose } 0 < c_2 < \frac{1}{\sqrt{2}}\Phi^{-1}(0.55).$$