Chapter 1: Probability

Soln 1.3.a
or any m, n and $m < n$,

$$\begin{aligned} B_m &= \bigcup_{i=m}^{\infty} A_i = \lim_{s\to\infty} \bigcup_{i=m}^{s} A_i \\ &= \lim_{s\to\infty} \bigcup_{i=n}^{s} A_i \cup \bigcup_{i=m}^{n} A_i \\ &= B_n \cup \bigcup_{i=m}^{n} A_i \end{aligned}$$

Hence $\{B_n\}$ is a monotonic decreasing sequence. The same argument holds for $\{C_n\}$.

Soln 1.3.b

$$\begin{aligned} \omega \in \bigcap_{n=1}^{\infty} B_n &\Leftrightarrow \omega \in B_n,\ \forall n \\ &\Leftrightarrow \omega \in B_1 \\ &\Leftrightarrow \omega \in A_n,\ \forall n \end{aligned}$$

Soln 1.3.c
$\omega \in \bigcup_{n=1}^{\infty} C_n \Leftrightarrow \omega \in C_{m\in M}$. By monotonicity of C_n, there is a finite m such that $\omega \in C_{\{\forall n \geq m\}}$ and $\omega \notin C_{\{1..m\}}$. By definition of C_n, ω is not in $A_1, ..., A_m$ but in $A_m, ...$ The converse holds as well.

Soln 1.12
Let G denotes green side, (GG, RR, GR) be the 3 cards respectively. $s1$ be the side we pick, $s2$ be the other side.

$$\begin{aligned} P(s2 = G|s1 = G) &= \frac{P(s2 = G, s1 = G)}{P(s1 = G)} \\ &= \frac{P(GG)}{P(s1 = G|GG)P(GG) + P(s1 = G|GR)P(GR) + P(s1 = G|RR)P(RR)} \\ &= \frac{1/3}{1/3(1 + 1/2 + 0)} \\ &= \frac{2}{3} \end{aligned}$$

Chapter 2: Random Variable

Soln 2.4.a

$$F_X(x) = \begin{cases} 0 & if\ x \le 0 \\ \frac{x}{4} = 0 + \int_0^x \frac{1}{4} dt & if\ 0 < x < 1 \\ \frac{1}{4} & if\ 1 \le x \le 3 \\ \frac{1}{4} + \frac{8}{3}(x-3) & if\ 3 < x < 5 \\ 1 & x \ge 5 \end{cases}$$

Soln 2.4.b

$$\begin{aligned} F_Y(y) &= P(A_y = \{x | x >= \frac{1}{y}\}) \\ &= \int_{1/y}^{\infty} f_X(x)dx = 1 - F_X(\frac{1}{y}) \end{aligned}$$

For $5 \le x = \frac{1}{y} \Rightarrow y \le \frac{1}{5},\ F_Y(y) = 1 - 1 = 0$

For $3 < x < 5 \Rightarrow \frac{1}{5} < y < \frac{1}{3},\ F_Y(y) = 1 - (\frac{1}{4} - \frac{8}{3}(\frac{1}{y} - 3)) =$

For $1 \le x \le 3 \Rightarrow \frac{1}{3} \le y \le 1,\ F_Y(y) = 1 - \frac{1}{4}$

For $0 < x < 1 \Rightarrow 1 < y,\ F_Y(y) = 1 - \frac{1}{4y}$

Taking the derivative, we get the pdf.

$$f_Y(y) = \begin{cases} 0 & y < \frac{1}{5} \\ -\frac{8}{3y^2} & \frac{1}{5} \le y < \frac{1}{3} \\ 0 & \frac{1}{3} \le y < 1 \\ \frac{1}{4y^2} & 1 \le y \end{cases}$$

Soln 2.6

$F_Y(y) = P(Y \le y) = P(I_A(X) \le y) = P(\{x | I_A(x) \le y\}) = P(A_y)$

If $y \ge 1$, then $\forall x, I_A(x) \le y \Rightarrow A_y = \Omega,\ P(A_{y \ge 1}) = P(\Omega) = 1$

If $y < 1$, then $\forall x \notin A, I_A(x) = 0 < y$, but $\forall x \in A, I_A(x) = 1 > y$. So $A_y = \{x | x \in A^c\} \Rightarrow P(A_{y<1}) = \int_{A^c} f_X(x)dx$.

$$F_Y(y) = \begin{cases} 1 & if\ y \ge 1 \\ c = \int_{A^c} f_X(x)dx & if\ y < 1 \end{cases}$$

Soln 2.7

Since X, Y are independent, $f_{X,Y}(x, y) = f_X(x) f_Y(y) = 1$

$P(Z < z) = 1$ if $z \geq 1$, 0 if $z \leq 0$

If $0 < z < 1$, $P(Z > z) = P(min\{X, Y\} > z) = \int_z^1 \int_z^1 f_{X,Y}(x, y)dxdy = (z-1)^2$

Then $P(Z \leq z) = 1 - P(Z > z) = 1 - (z-1)^2 = 2z - z^2$

$$F_Z(z) = \begin{cases} 0 & if\ z < 0 \\ 2z - z^2 & if\ 0 \leq z < 1 \\ 1 & if\ z \geq 1 \end{cases}$$

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \begin{cases} 0 & otherwise \\ \frac{d}{dz}(2z - z^2) = 2 - 2z & if\ 0 \leq z < 1 \end{cases}$$

Soln 2.16
Suppose $X \sim Poi(\lambda), Y \sim Poi(\mu)$ and $X \perp Y$.

$$P(X = k|X + Y = n) = \frac{P(X + Y = n|X = k)P(X = k)}{P(X + Y = n)}$$

Note that $P(X + Y = n|X = k) = P(Y = n - k) = \frac{\mu^{n-k}e^{-\mu}}{(n-k)!}$
$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
$P(X + Y = n) = \sum_{i=0}^{\infty} P(X = i)P(Y = n - i) = \sum_{i=0}^{\infty} \frac{\mu^{n-i}e^{-\mu}\lambda^i e^{-\lambda}}{(n-i)!i!}$
Hence,

$$\begin{aligned} P(X = k|X + Y = n) &= \frac{\frac{\mu^{n-k}e^{-\mu}\lambda^k e^{-\lambda}}{(n-k)!k!}}{\sum_{i=0}^{\infty} \frac{\mu^{n-i}e^{-\mu}\lambda^i e^{-\lambda}}{(n-i)!i!}} \\ &= \frac{\lambda^k \mu^{n-k}\binom{n}{k}}{\sum_{i=0}^{n} \lambda^i \mu^{n-i}\binom{n}{i}} \\ &= \binom{n}{k} \frac{\lambda^k \mu^{n-k}}{(\lambda + \mu)^n} \\ &= \binom{n}{k} \frac{\lambda^k \mu^{n-k}}{(\lambda + \mu)^k (\lambda + \mu)^{n-k}} \\ &= \binom{n}{k} \left(\frac{\lambda}{\lambda + \mu}\right)^k \left(\frac{\mu}{\lambda + \mu}\right)^{n-k} \\ &= \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{(given that } p = \frac{\lambda}{\lambda + \mu}) \end{aligned}$$

Chapter 3: Expectation

Soln 3.1
To Find X_n, Let I_i be the random variable for multiplication factor at ith trial and c be starting fortune. Since I_i are i.i.d uniform, $X_n \sim B(n, \frac{1}{2})$

$$P(X_n = c2^i \frac{1}{2^{n-i}}) = B(n, \frac{1}{2}) = \binom{n}{i} \frac{1}{2^n}$$

$$E(X_n) = c\sum_{i=0}^{n} \binom{n}{i} \frac{2^i}{2^{n-i}} \frac{1}{2^n} = c\sum_{i=0}^{n} \binom{n}{i} \frac{1}{4^{n-i}} = c(\frac{5}{4})^n$$

Soln 3.2
$V(X) = E(X-\mu)^2 = 0 \Rightarrow (X-\mu)^2 = 0 \Rightarrow X = \mu$ almost everywhere.

$$P(X = \mu) = \int f_X(x)dx = 1$$

Conversely, if $P(X = c) = 1 \Rightarrow X = c$ almost everywhere $\Rightarrow V(X) = E(X-c)^2 = 0$

Soln 3.4
Let $I_n \in \{1, -1\}$ be random variable of a step. Then $E(X_n) = E(\sum_i I_i) = \sum E(I_i) = n(-p + (1-p)) = n(1-2p)$.

To find $V(X_n)$, note that $E(I_i I_i) = 1$, for $i \neq j, I_i I_j \sim B(2, p)$, which means

$$P(I_i I_j = (-1)^i (1)^{2-i}) = \binom{2}{i} p^i (1-p)^{2-i}$$

$$\begin{aligned}
\Longrightarrow E(I_i I_j) &= \sum_{0}^{2} (-1)^i (1)^{2-i} \binom{2}{i} p^i (1-p)^{2-i}, \quad (i \neq j) \\
&= (1-p)^2 - 2p(1-p) + p^2 \\
&= (1-2p)^2
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow V(X_n) &= EX_n^2 - (EX_n)^2 \\
&= E(\sum I_i I_j) - n^2(1-2p)^2 \\
&= \sum E(I_i I_j) - n^2(1-2p)^2 \\
&= n + \binom{n}{2}(1-2p)^2 - n^2(1-2p)^2 \\
&= n - n(\frac{n+1}{2})(1-2p)^2
\end{aligned}$$

Soln 3.5
Let X be the number of toss before the first head appears. $EX = \sum_{k=1}^{\infty}(1-p)^{k-1}p = \frac{1}{p} = 2$

Chapter 4: Inequalities

Soln 4.1

Given $X \sim \lambda e^{-\lambda x}$

$$\begin{aligned} P(|X-\mu| \geq k\sigma) &= P(X \geq k\sigma + \mu) + P(X \leq \mu - k\sigma) \\ &= \int_{k\sigma+\mu}^{\infty} \lambda e^{-\lambda x} + \int_{0}^{\mu-k\sigma} \lambda e^{-\lambda x} \\ &= e^{1-k} - e^{k-1} + 1, \quad (\sigma = \mu = \frac{1}{\lambda}) \end{aligned}$$

Note that Chebyshev's inequality gives a tight upper bound of $\frac{1}{k}$.
To see this, obviously for $k > 1$, we have $e^{1-k} - e^{k-1} + 1 < \frac{1}{k}$ for $k > 0$, equality holds when $k = 1$
For $0 < k < 1$, Let $y = 1-k$, then $0 < y < 1$, we only need to show that $e^{y} - e^{-y} + 1 < \frac{1}{1-y}$.

$$\begin{aligned} 1 + e^{y} - e^{-y} &= 1 + \sum \frac{y^i}{i!} - \sum \frac{(-y)^i}{i!} \\ &= 1 + \sum_{i\ odd} \left(\frac{y^i}{i!} + \frac{y^i}{i!} \right) \quad \text{(we can rerrange since series converge)} \\ &< 1 + \sum_{i\ odd} y^{i+1} + y^i \\ &= 1 + y + y^2 + ... \\ &= \frac{1}{y-1} \end{aligned}$$

Chapter 9: Parametric Inference

Soln 9.2.a
A j-th moment $\alpha_j(a,b) = EX^j = \frac{1}{b-a}\int_a^b x^j dx = \frac{b^{j+1}-a^{j+1}}{(j+1)(b-a)}$

The methods of moments estimator $(\hat{a}, \hat{b})$

$$\begin{aligned}\alpha_k(\hat{a}, \hat{b}) &= \frac{\hat{b}^{k+1} - \hat{a}^{k+1}}{(k+1)(\hat{b} - \hat{a})} \\ &= \frac{1}{n} \sum_i \hat{X}_i^k \quad (\forall k)\end{aligned}$$

Note that $\alpha_1(\hat{a}, \hat{b}) = E\hat{X} = \frac{\hat{a}+\hat{b}}{2}$ and $\alpha_2(\hat{a}, \hat{b}) = E\hat{X}^2 = \frac{\hat{b}^3 - \hat{a}^3}{3(\hat{b}-\hat{a})} = \frac{\hat{a}^2+\hat{a}\hat{b}+\hat{b}^2}{3}$
Solving these two equations, we have

$$\hat{a} = E\hat{X} - \sqrt{3V(\hat{X})}, \quad \hat{b} = E\hat{X} + \sqrt{3V(\hat{X})}$$

If we draw from $[0, 1]$, then $V(\hat{X})$ approximates the real variance which is $\frac{1}{12}(b-a)^2 = \frac{1}{12}$, and the $E\hat{X}$ approximates the expected value 0.5. So we have $\hat{a}$ approximating to 0, and $\hat{b}$ approximating to 1.

Soln 9.2.b
X_i are i.i.d, Likelihood function $L(a, b) = \prod_i f(X_i; a, b)$. We maximize the logarithm.

$$l(a, b) = log(L(a, b)) = \sum_i log(f(X_i; a, b))$$

Notice that if $X_i \notin [a, b]$, then pdf evaluates to 0. So to maximize the likelihood function, $a = min(X_i)$, $b = max(X_i)$

Chapter 23: Stochastic Processes

Soln 23.1
We have directed graph $X_0 \to X_1 \to X_2$. Factoring the joint distribution, we have $P(X_0, X_1, X_2) = P(X_0)P(X_1|X_0)P(X_2|X_1)$.
Hence, $P(X_0 = 0, X_1 = 1, X_2 = 2) = 0.3 \times 0.2 \times 0 = 0$
$P(X_0 = 0, X_1 = 1, X_2 = 1) = 0.3 \times 0.2 \times 0.1 = 0.06$

Soln 23.2

$$\begin{aligned}p_{ij} = P(X_n = j|X_{n-1} = i) &= P(max(Y_0, ..., Y_n) = j|X_{n-1} = i) \\ &= P(max(X_{n-1}, Y_n) = j|X_n n - 1 = i) \\ &= P(max(i, Y_n) = j)\end{aligned}$$

For each pair of $(i, j$, we can find the probability of p_{ij}. Hence $P = \begin{pmatrix} 0.1 & 0.3 & 0.2 & 0.4 \\ 0 & 0.4 & 0.2 & 0.4 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Soln 23.3
o find P^n, we need to find the eigenvalue and eigenvectors of P.

$$\begin{aligned} &det(P - \lambda I) = (1 - a - \lambda)(1 - b - \lambda) - ab = 0 \\ &\Rightarrow \lambda = 1, 1 - a - b, v_1 = (\sqrt{2}, \sqrt{2}), v_{1-a-b} = \frac{1}{\sqrt{a^2 + b^2}}(a, -b) \end{aligned}$$

Then

$$P = V^{-1} D V$$
$$\Rightarrow P^n = V^{-1} D^n V = \begin{pmatrix} \sqrt{2} & -\frac{a}{\sqrt{a^2+b^2}} \\ \sqrt{2} & \frac{b}{\sqrt{a^2+b^2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1 - a - b)^n \end{pmatrix} \begin{pmatrix} \sqrt{2} & -\frac{a}{\sqrt{a^2+b^2}} \\ \sqrt{2} & \frac{b}{\sqrt{a^2+b^2}} \end{pmatrix}^{-1}$$
$$= \frac{1}{a+b} \begin{pmatrix} b + a\epsilon^n & a - a\epsilon^n \\ b - b\epsilon^n & a + b\epsilon^n \end{pmatrix} \qquad \text{(where } \epsilon = 1 - a - b\text{)}$$

Taking n to infinity, we get the answer.