Problem Set #1

Charlie Walker

Problem 1

3.6

By independence,

$$\sum_{i\in I} P(A\cap B_i) = P(A)[\sum_{i\in I} P(B_i)]$$

By additivity,

$$P(A)[\sum_{i\in I} P(B_i)] = P(A)P(\bigcup_{i\in I} B_i)$$

Since $\bigcup_{i\in I} B_i = \Omega$ and $B_i \cap B_j = \emptyset$ for all $i \neq j$, $P(\bigcup_{i\in I} B_i) = 1 \implies P(A)P(\bigcup_{i\in I} B_i) = P(A)$, as desired.

3.8

$$\begin{aligned}
1 - \prod_{k=1}^{n}(1 - P(E_k)) &= 1 - \prod_{k=1}^{n} P(E_k^c) \\
&= 1 - P(\bigcap_{k=1}^{n} E_k^c) \\
&= P((\bigcap_{k=1}^{n} E_k^c)^c) \\
&= P(\bigcup_{k=1}^{n} E_k)
\end{aligned}$$

Where line (2) follows from independence, and line (4) by DeMorgan's Laws.

3.11

Assume that the DNA test is perfectly accurate in identifying someone who was at the crime scene, i.e. $P(pos|guilty) = 1$. From the description, we have:

$$P(guilty) = \frac{1}{250m}$$
$$P(pos) = \frac{1}{3m}$$

We want the probability that an individual was at the crime scene, given their DNA test was positive, that is $P(guilty|positive)$. Bayes' Rule gives:

$$\begin{aligned}
P(guilty|pos) &= \frac{P(pos|guilty)P(guilty)}{P(pos)} \\
&\approx 1.2\%
\end{aligned}$$

3.12

First, some setup. Define the events:

D1 = Morty opens door 1
D2 = Morty opens door 2
D3 = Morty opens door 3
C1 = Car behind door 1
C2 = Car behind door 2
C3 = Car behind door 3

Assume you choose door 1. Then,

$$P(D3|C1) = P(D2|C1) = \frac{1}{2}$$
$$P(D3|C3) = 0$$
$$P(D3|C2) = 1$$

And,

$$\begin{aligned} P(C1|D3) &= \frac{P(D3|C1)P(C1)}{P(D3)} \\ &= \frac{1}{3} \\ P(C2|D3) &= \frac{P(D3|C2)P(C2)}{P(D3)} \\ &= \frac{2}{3} \end{aligned}$$

For the 10-door case, if Monty opens 8 doors, the contestant has a $\frac{9}{10}$ chance of winning if they switch, and a $\frac{1}{10}$ chance if they do not.

3.16

$$\begin{aligned} E[(X-\mu)^2)] &= E[X^2 - 2X\mu + \mu^2] \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

3.33

Chebyshev's Inequality for a binomial random variable with mean np and variance $(1-p)p$ gives:

$$P(|B - np| \geq \epsilon) \leq \frac{(1-p)p}{\epsilon^2} \implies P\left(\left|\frac{B}{n} - p\right| \geq \epsilon\right) \leq \frac{(1-p)p}{n\epsilon^2}$$

3.36

Let B_n denote individual trials, $B_n \in \{0, 1\}$. Each enrolment event is a Bernoulli trial with success probability $P(B_n = 1) = 0.801$. Define $X_n = B_1 + B_2 + ... + B_n$. For $n = 6242, E[X] = (6242)(0.801) = 5000, Var[X] = (6242)(0.801)(0.199) = 994.97$. By the central limit theorem, $X \sim N(5000, 994.97)$.

$$\begin{aligned}
P(X > 5500) &= P\left(\frac{X - 5000}{\sqrt{994.97}} > \frac{500}{\sqrt{994.97}}\right) \\
&= 1 - \Phi\left(\frac{500}{\sqrt{994.97}}\right) \\
&= 1 - \Phi(15.85) \\
&\approx 0.003\%
\end{aligned}$$

Problem 2

(a) Throw two 6-sided die. Define events:

A: sum of points is 7

B: die #1 is 3

C: die #2 is 4

Then,

$$\begin{aligned}
P(A) = P(B) = P(C) &= \frac{1}{6} \\
P(A \cap B) = P(A \cap C) = P(B \cap C) &= \frac{1}{36} \\
P(A \cap B \cap C) = \frac{1}{6} \neq P(A)P(B)P(C) &= \frac{1}{216}
\end{aligned}$$

(b)

Problem 3

Benford's Law states that $P(d) = log_{10}(1 + \frac{1}{d}), d \in \{1, 2, ..., 9\}$. $\sum_{i=1}^{9} log_{10}(1 + \frac{1}{d}) = log(2) + log(\frac{3}{2}) + log(\frac{4}{3}) + ... + log(\frac{10}{9}) = 1 \implies$ Benford's Law is a well-defined discrete probability distribution.

Problem 4

(a)

$$
\begin{aligned}
E[X] &= \sum_{n=1}^{\infty} 2^n P(i) \\
&= \sum_{n=1}^{\infty} 2^n \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{n-1} \\
&= \frac{1}{2}\cdot 2 + \frac{1}{4}\cdot 4 + \frac{1}{8}\cdot 8 + ... \\
&= 1 + 1 + 1 + ... \\
&= +\infty
\end{aligned}
$$

(b)

$$
\begin{aligned}
E[\ln X] &= \sum_{n=1}^{\infty} \ln(2^n) P(n) \\
&= \sum_{n=1}^{\infty} \ln(2^n) \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{n-1} \\
&= \frac{\ln(2)}{2} \sum_{n=1}^{\infty} n \left(\frac{1}{2}\right)^{n-1} \\
&= 2\ln(2)
\end{aligned}
$$

Problem 5

Both investors should invest in the foreign currency:

$$
\begin{aligned}
E[USD/CHF] &= (0.5)(1.25) + (0.5)(0.8) = 1.025 \\
E[CHF/USD] &= (0.5)(0.8) + (0.5)(1.25) = 1.025
\end{aligned}
$$

Problem 7

(a)

$$
\begin{aligned}
E[Y] &= E[XZ] = E[X]E[Z] = 0 \\
Var[Y] &= Var[XZ] \\
&= E[XZ^2] - E[XZ]^2 \\
&= E[X^2]E[Z^2] - E[X]^2 E[Z]^2 \\
&= E[X^2]E[Z^2] \\
&= 1 \\
&\implies Y \sim N(0,1)
\end{aligned}
$$

(b)

$$\begin{aligned}P(|X| = |Y|) &= P(|X| = |XZ|) \\ &= P(|X| = |X||Z|) \\ &= P(|X| = |X|) = 1\end{aligned}$$

(c)
(d)

$$\begin{aligned}Cov[X, Y] &= E[X, Y] - E[X]E[Y] \\ &= E[XY] \\ &= E[X^2Z] = 0\end{aligned}$$

(e)
Problem 8
For m, (temporarily substituting y for m to avoid confusion between m and M random variables), $F(y) = P(Y \leq y) = 1 - P(Y > y) = 1 - P(min(X_1, ..., X_n) > y)$. Since all X_i are i.i.d, we have $F(y) = 1 - P(X_1 > y)P(X_2 > y)...P(X_n > y) = 1 - P(X_1 > y)^n$. The uniform distribution of X_i's over [0,1] yields the cumulative distribution:

$$F(m) = \begin{cases} 0 & y < 0 \\ 1 - \left(\frac{1-m}{1}\right)^n & m \in (0,1) \\ 1 & m > b \end{cases}$$

The density function follows:

$$f(m) = \begin{cases} n\left(\frac{1-m}{1}\right)^{n-1} & m \in [0,1] \\ 0 & otherwise \end{cases}$$

Integrating over the distribution function, $E[m] = \int_{-\infty}^{\infty} mf(m)dm = \frac{1}{n+1}$
For M, (temporarily substituting x for M, for above reason), $F(x) = P(X \leq x) = 1 - P(X > x)$

Problem 9
(a) Let S_n denote individual states (trials), $S_n \in 0, 1$. Each state is a Bernoulli trial with success probability $P(S_n = 1) = P(S_n = 0) = 0.5$. Define $X_n = S_1 + S_2 + ... + S_n$. For $n = 1000, E[X] = (1000)(0.5) = 500, Var[X] = (1000)(0.5)(0.5) = 250$. By the central limit theorem, $X \sim N(500, 250)$. The desired probability is $P(490 < X_{1000} <$

510).

$$
\begin{aligned}
P(X > 490) &= 1 - P(X < 490) \\
&= 1 - P\left(\frac{X - 490}{\sqrt{250}} < \frac{490 - 600}{\sqrt{250}}\right) \\
&= 1 - \Phi(-0.632) \\
&= \Phi(0.632) = 0.736
\end{aligned}
$$

By the symmetry of the normal distribution, $P(X < 510) = 0.736 \implies P(490 < X_{1000} < 510) = 2(0.736) - 1 = 0.472$.

(b) Chebyshev's Inequality for X_n, the sum of n i.i.d Bernoulli trials with success probability 0.5 (defined in more detail above) gives:

$$
P\left(\left|\frac{X_n}{n} - \frac{1}{2}\right| \geq \epsilon\right) \leq \frac{1}{4n\epsilon^2} \tag{1}
$$

Where ϵ in this case is $(0.1)(0.5) = (0.005)$. For the left hand side of (1) to be less than 0.01, $n \geq 5000$.

Problem 10

Proof by contradiction. Assume $\theta < 0$. Define the function $f : \mathbb{R} \to \mathbb{R}, f(x) = e^{\theta x}$. By Jensen's Inequality and convexity of f, $f(E[X]) \leq E[f(x)] = 1 \implies e^{\theta E[X]} \leq E[e^{\theta X}]$. For $E[X] < 0$ and $\theta < 0, e^{\theta E[X]} > 1 = E[e^{\theta X}]$, contradiction Jensen's Inequality. Thus, θ must be greater than 0.