Problem 1.

3.6. First we shall prove that

$$A = \bigcup_{i\in I}(A\cap B_i).$$

Indeed, it is obvious that $\bigcup\limits_{i\in I}(A\cap B_i)$ is a subset of A because $(A\cap B_i)\subseteq A$ or every $i\in I$. On the other hand, for every $a\in A$, of course $a\in\Omega$ because $A\subseteq\Omega$. Then, since $\{B_i\}_{i\in I}$ is a partition of Ω, it follows that $a\in B_{i_a}$ for some $i_a\in I$ (this i_a is unique because these B_i's are pairwise disjoint). Hence, $a\in A\cap B_{i_a}$ for every $a\in A$, so we deduce that $A\subseteq\bigcup\limits_{i\in I}(A\cap B_i)$. Therefore, $A=\bigcup\limits_{i\in I}(A\cap B_i)$. Finally, for any $i\neq j\in I$, because $B_i\cap B_j=\emptyset$, it follows that $(A\cap B_i)\cap(A\cap B_j)=\emptyset$, hence $P(\bigcup\limits_{i\in I}(A\cap B_i))=\sum\limits_{i\in I}P(A\cap B_i)=P(A)$.

3.8. For any event A, denote $\overline{A}$ as the complement of A. We will show that if $X_1,X_2,\dots,X_m$ are independent random variables then $\overline{X_1},X_2,\dots,X_m$ are also independent. For any subset F of $\{2\dots,m\}$ We have

$$\begin{aligned}
P(\overline{X_1}\cap\prod_{i\in F}X_i) &= P(\prod_{i\in F}X_i) - P(X_1\cap\prod_{i\in F}X_i)\\
&= \prod_{i\in F}P(X_i) - P(X_1)\prod_{i\in F}P(X_i)\\
&= (1-P(X_1))\prod_{i\in F}P(X_i)\\
&= P(\overline{X_1})\prod_{i\in F}P(X_i).
\end{aligned}$$

Repeating the proof above multiple times, then we deduce that $\overline{X_1},\dots,\overline{X_m}$ are independent. Now since $E_1,\dots,E_n$ are independent, so are $\overline{E_1},\dots,\overline{E_n}$. We have

$$\begin{aligned}
P\left(\bigcup_{i=1}^n E_k\right) &= 1-P\left(\overline{\bigcup_{i=1}^n E_k}\right)\\
&= 1-P\left(\bigcap_{i=1}^n \overline{E_k}\right)\\
&= 1-\prod_{i=1}^n P(\overline{E_i})\\
&= 1-\prod_{i=1}^n (1-P(E_i)).
\end{aligned}$$

3.16. We have:

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[(X-\mu)^2]\\
&= \mathbb{E}[X^2-2\mu X+\mu^2]\\
&= \mathbb{E}[X^2]-2\mu\mathbb{E}[X]+\mu^2\\
&= \mathbb{E}[X^2]-\mu^2. \quad (\text{because } \mathbb{E}[X]=\mu)
\end{aligned}$$

3.11. We have

$$\begin{aligned}
P(s = \text{crime} \mid s \text{ tested}+) &= \frac{P(s = \text{crime}) \cap s \text{ tested}+)}{P(s \text{ tested}+)} \\
&= \frac{P(s = \text{crime})}{P(s \text{ tested}+)} (\text{because crime} \Rightarrow \text{tested } +) \\
&= \frac{1/(250 \cdot 10^6)}{1/(3 \cdot 10^6)} \\
&= \frac{3}{250}.
\end{aligned}$$

3.12. Note that the probability of winning if the contestant chooses to keep the initial door is always $\frac{1}{3}$. About switching, since there are 3 doors and a door with a goat will be revealed, the contestant wins when switching if and only if his initial choice is a goat, and vice versa. In other words, the probability of winning when switching is equal to the probability of choosing the wrong door at the beginning, which is $\frac{2}{3}$. Hence, the contestant should always switch. If there are 10 doors and 8 goats, the same argument still holds: the contestant has $\frac{1}{10}$ chance of winning if he keeps the initial door, while switching yields a $\frac{9}{10}$ chance of winning.

3.33. Denote X_i as the indicator variable of trial i. Note that $\mathbb{E}[B] = np$ (the expected value of each trial is p) and

$$\begin{aligned}
\text{Var}[B] &= \sum_{i=1}^{n} \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_i X_j] - \mathbb{E}^2[B] \\
&= np + n(n-1)p^2 - n^2 p^2 \\
&= np(1-p).
\end{aligned}$$

. By Chebyshev Inequality, we have:

$$\begin{aligned}
P\left(\left|\frac{B}{n} - p\right| \geq \varepsilon\right) &= P(|B - np| \geq n\varepsilon) \\
&\leq \frac{\text{Var}[B]}{n^2\varepsilon^2} \\
&= \frac{p(1-p)}{n\varepsilon^2}.
\end{aligned}$$

3.36. For any student $i = 1, \ldots, 6242$, let X_i be the indicator function of the event "student i chooses to enroll". Thus, the total number of students who enroll is S_{6242}. We can deduce that $S_{6242} \sim B(6242, 0.801)$, and by the same calculation as 3.33, we know that $\text{Var}[B] = 6242 \cdot 0.801 \cdot 0.199 = 995$. We need to find

$$\begin{aligned}
&P(S_{6242} \geq 5500) \\
&= P\left(\frac{S_{6242} - 5000}{\sqrt{995}\sqrt{6242}} \geq 0.2\right) \\
&= \varphi(-0.2)
\end{aligned}$$

Problem 2.

(a) Suppose we toss a coin twice. Let A be the event that the first toss is H, B be the event the second toss is H and C be the event there is exactly one H in the two tosses. It is easy to see that $P(A) = P(B) = P(C) = \frac{1}{2}$. We deduce that $A \cap B$ is the event that both tosses are H's, $B \cap C$ is the event TH, $A \cap C$ is the event HT, hence all these events have probability $\frac{1}{4}$. Thus, we have $P(A \cap B) = P(A)P(B), P(B \cap C) = P(B)P(C), P(A \cap C) = P(A)P(C)$. However, $P(A \cap B \cap C) = 0 \neq P(A)P(B)P(C)$.

(b) Let Ω be the space of 12 equally likely outcomes, denoted by $a_i, i = 1, \ldots, 12$. Let $A = \{a_4, \ldots, a_{12}\}, B = \{a_3, a_4, a_5, a_6\}, C = \{a_2, a_6, a_7, a_8\}$. Now we have:

$$P(A) = \frac{3}{4}, \quad P(B) = \frac{1}{3}, \quad P(C) = \frac{1}{3}$$

and

$$P(A \cap B) = \frac{|A \cap B|}{12} = \frac{3}{12} = \frac{1}{4} = P(A \cap B).$$

Similarly, we have

$$P(A \cap C) = \frac{1}{4} = P(A)P(C).$$

and

$$P(A \cap B \cap C) = \frac{1}{12} = P(A)P(B)P(C).$$

But

$$P(B \cap C) = \frac{1}{12} \neq P(B)P(C).$$

Problem 3. It suffices to show that

$$\sum_{d=1}^{10} \log_{10}\left(1 + \frac{1}{d}\right) = 1$$

or equivalently,

$$\sum_{d=1}^{10} \log_{10}(d+1) - log_{10}(d) = 1.$$

But $\sum_{d=1}^{10} \log_{10}(d+1) - log_{10}(d)$ is just $\log_{10}(10) - \log_{10}(1) = 1$. Hence the Benford's Law is a well-defined discrete probability distribution.

Problem 4.

(a) Observe that the probability that the first tail appears on the nth flip is $\frac{1}{2^n}$ (because it must be a sequence $HH \ldots HT$, with $(n-1)$ H's). This is also the probability that the person wins 2^n. We have

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{n=1}^{\infty} 2^n \cdot P(X = 2^n) \\
&= \sum_{n=1}^{\infty} 1 \\
&= \infty
\end{aligned}$$

(b) Now, we have

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

We have

$$\begin{aligned}
S &= \sum_{n=1}^{\infty} \frac{n}{2^n} \\
&= \sum_{n=1}^{\infty} \left(\frac{n}{2^{n-1}} - \frac{n}{2^n}\right) \\
&= \sum_{n=1}^{\infty} \left(\frac{n-1}{2^{n-1}} - \frac{n}{2^n}\right) + \sum_{n=1}^{\infty} \frac{1}{2^{n-1}} \\
&= 2.
\end{aligned}$$

Thus, we conclude that $\mathbb{E}[\ln X] = 2\ln 2$.

Problem 5. Consider the Swiss investor. Investing 1 CHF will result in the same amount of CHF in the future, so the expected value of his investment is 1. However, if he invests in USD, then the expected value of 1 USD invested (in terms of future CHF) is $\frac{1.25+\frac{1}{1.25}}{2} = \frac{41}{40} > 1$. Thus, it is better for the Swiss investor to invest in America. By the same argument, the U.S. investor should invest in Switzerland.

Problem 6. Throughout this problem, let A be a uniformly distributed random variable on $[0,1]$. That is, for $[a,b] \subseteq [0,1], P(A \in [a,b]) = b-a$.

(a) Define $X = \frac{1}{\sqrt{A}}$ when $A \neq 0$ and $X = 0$ when $A = 0$. Apparently X is also a random variable from $[0,1]$ to $\mathbb{R}$. We have

$$\begin{aligned}\mathbb{E}[X] &= \int_0^1 \frac{1}{\sqrt{x}}dx \\ &= 2\end{aligned}$$

However,

$$\begin{aligned}\mathbb{E}[X^2] &= \int_0^1 \frac{1}{x}dx \\ &= \infty\end{aligned}$$

(b) Define X,Y as followed: $X = 0, Y = 3$ if $A \in [0,\frac{1}{3}]$ and $Y = 0, X = 1$ if $A \in (\frac{1}{3},1]$. Apparently we have

$$P(X > Y) = P(A \in (\frac{1}{3},1]) = \frac{2}{3} > \frac{1}{2}$$

and

$$\begin{aligned}\mathbb{E}[X] - \mathbb{E}[Y] &= \frac{2}{3}\cdot 1 - \frac{1}{3}\cdot 3 \\ &= -\frac{1}{3} \\ &< 0.\end{aligned}$$

(c) Define X,Y,Z as followed: $X = 1, Y = 0, Z = -1$ if $A \in [0,\frac{1}{3})$, $X = -1, Y = 1, Z = 0$ if $A \in [\frac{1}{3},\frac{2}{3})$ and $X = 0, Y = -1, Z = 1$ if $A \in [\frac{2}{3},1]$. It is easy to see that $\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[Z] = 0$. Moreover, we have

$$P(X > Y) = P(A \notin [\frac{1}{3},\frac{2}{3})) = \frac{2}{3}.$$

$$P(Y > Z) = P(A \notin [\frac{2}{3},1]) = \frac{2}{3}.$$

$$P(X > Z) = P(A \in [0,\frac{1}{3})) = \frac{1}{3}.$$

Hence $P(X > Y)P(Y > Z)P(X > Z) > 0$.

Problem 7.

(a) Denote F_Y as the CDF of Y. We have

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(XZ \le y) \\ &= \frac{1}{2}P(X \le y) + \frac{1}{2}P(X \ge -y) \\ &= P(X \le y) \quad \text{(because } X \sim \mathcal{N}(0,1) \text{ so } P(X \le y) = P(x \ge -y)) \\ &= F_X(y). \end{aligned}$$

Since $X \sim \mathcal{N}(0,1)$, it follows that $Y \sim \mathcal{N}(0,1)$ as well.

(b) Note that $P(X=0) = P(Y=0) = 0$, so we have

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

(c) Suppose X and Y are independent. We have

$$\begin{aligned} P(X \le x; Y \le y) &= P(X \le x; XZ \le y) \\ &= \frac{1}{2}P(X \le x; X \le y) + \frac{1}{2}P(X \le x; X \ge -y) \\ &= \frac{1}{2}P(X \le \min\{x,y\}) + \frac{1}{2}P(-y \le X \le x) \quad (1) \end{aligned}$$

If X and Y are independent, then for any x, y, we should have $P(X \le x; Y \le y) = P(X \le x) \cdot P(Y \le y) = F(x)F(y)$, where F is the CDF of $\mathcal{N}(0,1)$. From (1), if we choose $x > 0$ and $y = -x$, then we get

$$\begin{aligned} P(X \le x, Y \le -x) &= \frac{1}{2}P(X \le -x) \\ &= F(-x) \\ &= F(x)F(-x). \end{aligned}$$

This implies that $F(-x) = 0$ or $F(x) = 1$, for any $x > 0$, a contradiction. Thus, X and Y are not independent.

(d) Note that since X and Z are independent, so are X^2 and Z. Furthermore, $\mathbb{E}[Z] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0$. Observe that

$$\begin{aligned} \mathbb{E}[XY] &= \mathbb{E}[X^2 Z] \\ &= \mathbb{E}[X^2]\mathbb{E}[Z] \\ &= 0 \\ &= \mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

Hence, $\mathrm{Cov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$.

(e) Not true. As we can see X and Y in this problem are both normally distributed, satisfy $\mathrm{Cov}[X,Y] = 0$ but are not independent.

Problem 8. Let F_m, F_M be the CDF's and f_m, f_M be the PDF's of m and M, respectively. It is obious that both m and M only take values on $[0,1]$. For any $x \in [0,1]$, we have

$$\begin{aligned} F_M(x) &= P(M \le x) \\ &= P(\max\{X_1, \ldots, X_n\} \le x) \\ &= P(X_1 \le x;\ X_2 \le x;\ \ldots;\ X_n \le x) \\ &= \prod_{i=1}^{n} P(X_i \le x) \\ &= x^n. \end{aligned}$$

Hence, we deduce that

$$f_M(x) = nx^{n-1}.$$

Next, we have

$$\begin{aligned}
F_m(x) &= P(m \leq x) \\
&= 1 - P(\min\{X_1, \ldots, X_n\} > x) \\
&= 1 - P(X_1 > x;\ X_2 > x;\ \ldots;\ X_n > x) \\
&= 1 - \prod_{i=1}^{n} P(X_i > x) \\
&= 1 - (1-x)^n
\end{aligned}$$

Thus, we deduce that

$$f_m(x) = n(1-x)^{n-1}.$$

Expected value:

$$\begin{aligned}
\mathbb{E}[M] &= \int_0^1 nx^n dx \\
&= \frac{n}{n+1}.
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[m] &= \int_0^1 nx(1-x)^{n-1} dx \\
&= \frac{1}{n+1}.
\end{aligned}$$

Problem 9.

(a) The number of good states over 1000 periods differs by 500 by at most 20% means that the number of good states belongs to the set $\{480, 481, \ldots, 520\}$. Thus, the desired probability is

$$P = \frac{1}{2^{1000}} \sum_{i=480}^{520} \binom{1000}{i}.$$

(b) For $i = 1, 2, \ldots$, let X_i be the random variable defined as followed: $X_i = 1$ if good state happens in period i and $X_i = 0$ otherwise. We have $\mathbb{E}[X_i] = \frac{1}{2}$ and $\text{Var}[X_i] = \mathbb{E}[X_i^2] - \mathbb{E}^2[X_i] = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$ for all i. Let $S_n = X_1 + \ldots + X_n$. The proportion of good states after n periods is $\frac{S_n}{n}$ and it differs from $\frac{1}{2}$ by $\frac{S_n - n\mu}{n}$ (because $\mu = \frac{1}{2}$). Note that $\sigma = \sqrt{\text{Var}[X_i]} = \frac{1}{2}$. We need to find n such that

$$P\left(\frac{S_n - n\mu}{n} \leq 0.01\right) \geq 0.99.$$

or equivalently,

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq 0.02\sqrt{n}\right) \geq 0.99.$$

By the Central Limit Theorem, we know that $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ tends to $\mathcal{N}(0,1)$ as n goes to infinity. Using Excel, we find that the smallest y such that $P(\mathcal{N}(0,1) \leq y) \geq 0.99$ is approximately 2.32. Hence, we solve for n:

$$n = \left(\frac{2.32}{0.02}\right)^2 \sim 13456.$$

Problem 10. Note that for $\theta \neq 0$, $f(x) = e^{\theta x}$ is a differentiable, strictly convex function because $f''(x) = \theta^2 e^{\theta x} > 0$. Hence, by Jensen's Inequality, we deduce that

$$1 = \mathbb{E}[e^{\theta X}] \leq e^{\theta \mathbb{E}[X]}.$$

This implies that $\theta \mathbb{E}[X] \geq 0$, and hence $\theta < 0$ because $\mathbb{E}[X] < 0$.