

Problem 3.6:

Proof. $B_i \cap B_j = \emptyset \forall i \neq j$, so $(A \cap B_i) \cap (A \cap B_j) = \emptyset \forall i \neq j$, i.e. $\{A \cap B_i\}_{i \in I}$ is a collection of pairwise-disjoint events. Since it is indexed by a finite or countable set I , then, we have that:

$$\begin{aligned} \sum_{i \in I} P(A \cap B_i) &= P\left(\bigcup_{i \in I} A \cap B_i\right) \\ &= P\left(A \cap \left(\bigcup_{i \in I} B_i\right)\right) && \text{by the distributive property of } \cap \\ &= P(A \cap \Omega) && \text{by assumption} \\ &= P(A) && \text{since } A \subseteq \Omega \end{aligned}$$

□

Problem 3.8:

Proof.

$$\begin{aligned} P\left(\bigcup_{k=1}^n E_k\right) &= 1 - P\left(\left(\bigcup_{k=1}^n E_k\right)^C\right) \\ &= 1 - P\left(\bigcap_{k=1}^n E_k^C\right) && \text{by De Morgan's law} \\ &= 1 - \prod_{k=1}^n P(E_k^C) && \text{since the events are independent} \\ &= 1 - \prod_{k=1}^n (1 - P(E_k)) \end{aligned}$$

□

Problem 3.11:

Proof. Let C denote "criminal", NC denote "not criminal", and $+$ denote "tested positive". Then, we know that $P(+|NC) = \frac{1}{3,000,000}$, $P(+|C) = 1$ (by assumption), and that $P(C) = 1 - P(NC) = \frac{1}{250,000,000}$. So, we have that:

$$\begin{aligned} P(NC|+) &= \frac{P(+|NC)P(NC)}{P(+|NC)P(NC) + P(+|C)P(C)} \\ &= \frac{\frac{1}{3,000,000} \cdot \frac{249,999,999}{250,000,000}}{\frac{1}{3,000,000} \cdot \frac{249,999,999}{250,000,000} + 1 \cdot \frac{1}{250,000,000}} \\ &\approx 0.98814 \end{aligned}$$

So, $P(C|+) = 1 - P(NC|+) \approx 0.01186$. This is the probability that the suspect is the criminal, given that they tested positive. \square

Problem 3.12:

Proof. Consider the mutually exclusive and collectively exhaustive set of events $\{C_1, C_2, C_3\}$, where C_n corresponds to the car being behind the n th door. Also, consider the set of events $\{D_1, D_2, D_3\}$ where D_n corresponds to Monty Hall opening the n th door. Suppose you choose the 1st door. Then, we know that:

$$\begin{aligned} P(D_3|C_1) &= \frac{1}{2} && \text{since Monty Hall is equally likely to choose to open the 1st or the 2nd door} \\ P(D_3|C_2) &= 1 && \text{since Monty Hall must choose to open the 3rd door since you have chosen the 1st door} \\ P(D_3|C_3) &= 0 && \text{since Monty Hall cannot reveal the car} \end{aligned}$$

Clearly, we have that:

$$P(C_1) = P(C_2) = P(C_3) = \frac{1}{3}$$

So, we have that:

$$\begin{aligned} P(C_2|D_3) &= \frac{P(D_3|C_2)P(C_2)}{P(D_3|C_1)P(C_1) + P(D_3|C_2)P(C_2) + P(D_3|C_3)P(C_3)} \\ &= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} \\ &= \frac{2}{3} \\ P(C_1|D_3) &= \frac{P(D_3|C_1)P(C_1)}{P(D_3|C_1)P(C_1) + P(D_3|C_2)P(C_2) + P(D_3|C_3)P(C_3)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} \\ &= \frac{1}{3} \end{aligned}$$

Note that we may replace D_3 with D_2 in the above equations and arrive at the same result. Since we know that one of D_3 and D_2 must occur, it follows that the probability of winning is higher if the contestant changes to the other unopened door (the 2nd door) instead of sticking with the original choice (the 1st door) since:

$$P(C_2|D_3) = P(C_2|D_2) = \frac{2}{3} > \frac{1}{3} = P(C_1|D_2) = P(C_1|D_3)$$

If there were 10 doors and Monty opened 8 with goats after the contestant's first choice, the contestant would have a $\frac{9}{10}$ chance of winning by switching, and a $\frac{1}{10}$ chance of winning by sticking with the original choice. This result can be found by a trivial generalization of the above proof (it is useful in this case to think of events D_n as Monty Hall *not* opening the n th door). \square

Problem 3.16:

Proof.

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

\square

Problem 3.33:

Proof. Recall that we define $B := \sum_{i=1}^n B_i$, where B_1, \dots, B_n are all independent, Bernoulli distributed variables with parameter p . So, we know that $E[B_i] := \mu = p$, and $\text{Var}[B_i] := \sigma^2 = p(1 - p)$. By the weak law of large numbers, we know that, for every $\epsilon > 0$:

$$\begin{aligned}
 P\left(\left|\frac{\sum_{i=1}^n B_i}{n} - \mu\right| \geq \epsilon\right) &\leq \frac{\sigma^2}{n\epsilon^2} \\
 P\left(\left|\frac{B}{n} - p\right| \geq \epsilon\right) &\leq \frac{p(1-p)}{n\epsilon^2} \quad \text{by substitution}
 \end{aligned}$$

\square

Problem 3.36:

Proof. For each Bernoulli trial, $E[B_i] := \mu = p = 0.801$ (by assumption) and $\text{Var}[B_i] := \sigma^2 = p(1 - p) = 0.159399$. Define the following:

$$S_{6242} := \sum_{i=1}^{6242} B_i \quad y := \frac{5500 - 6242\mu}{\sqrt{6242}\sigma} = 15.8563$$

Since 6242 is large, we know that:

$$P(S_{6242} \geq 5500) \approx 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx \approx 0$$

This is the probability that the number of students enrolling will exceed 5500, provided the university admits 6242 students. \square

Problem 2:

Part a.

Proof. Consider a sample space $\Omega := \{1, 2, 3, 4, 5, 6, 7, 8\}$. Say $P(\omega) = 1/|\Omega| \forall \omega \in \Omega$. Then, define the following events:

$$A := \{1, 2, 3, 4\} \quad B := \{3, 4, 5, 6\} \quad C := \{3, 4, 7, 8\}$$

Then, the following equations hold:

$$\begin{aligned} P(A \cap B) &= P(A)P(B) = \frac{1}{4} \\ P(A \cap C) &= P(A)P(C) = \frac{1}{4} \\ P(B \cap C) &= P(B)P(C) = \frac{1}{4} \\ P(A \cap B \cap C) &= \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C) \quad \text{as required} \end{aligned}$$

\square

Part b.

Proof. Consider a sample space $\Omega := \{1, 2, 3, 4, 5, 6, 7, 8\}$. Say $P(\omega) = 1/|\Omega| \forall \omega \in \Omega$. Then, define the following events:

$$A := \{1, 2, 3, 4\} \quad B := \{3, 4, 5, 6\} \quad C := \{2, 4, 7, 8\}$$

Then, the following equations hold:

$$\begin{aligned}
P(A \cap B) &= P(A)P(B) = \frac{1}{4} \\
P(A \cap C) &= P(A)P(C) = \frac{1}{4} \\
P(A \cap B \cap C) &= P(A)P(B)P(C) = \frac{1}{8} \\
P(B \cap C) &= \frac{1}{8} \neq \frac{1}{4} = P(B)P(C) \quad \text{as required}
\end{aligned}$$

□

Problem 3:

Proof. Consider the state space $\Omega := \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. We define the probability measure P s.t. $P(\omega) = \log_{10}(1 + \frac{1}{\omega}) \forall \omega \in \Omega$. Then, we have that:

$$\begin{aligned}
P(\Omega) &= \sum_{i=1}^9 P(i) \\
&= \sum_{i=1}^9 \log_{10}\left(1 + \frac{1}{i}\right) \\
&= \log_{10}\left(\frac{2}{1}\right) + \log_{10}\left(\frac{3}{2}\right) + \log_{10}\left(\frac{4}{3}\right) + \dots + \log_{10}\left(\frac{9}{8}\right) + \log_{10}\left(\frac{10}{9}\right) \\
&= \log_{10}\left(\frac{2}{1} \cdot \frac{3}{2} \cdot \frac{4}{3} \cdot \dots \cdot \frac{9}{8} \cdot \frac{10}{9}\right) \\
&= \log_{10}(10) \\
&= 1
\end{aligned}$$

Thus, Benford's Law is, in fact, a well-defined discrete probability distribution, since we may impose the additivity axiom. □

Problem 4:

Part a.

Proof. Consider the state space $\Omega := \{T, HT, HHT, HHHT, HHHHT, \dots\}$. Consider an outcome $H...HT$ with $n \geq 0$ H s. Then we define the random variable X s.t. $X(H...HT) = 2^{(n+1)}$. Also, since the coin is fair, we know that

$P(H...HT) = 1/2^{(n+1)}$. It follows that:

$$\begin{aligned} E[X] &= \frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 4 + \frac{1}{8} \cdot 8 + \dots \\ &= 1 + 1 + 1 + \dots \\ &= +\infty \end{aligned}$$

If we substitute $\ln(X)$ for X , we get that:

$$\begin{aligned} E[\ln(X)] &= \frac{1}{2} \cdot \ln(2) + \frac{1}{4} \cdot \ln(4) + \frac{1}{8} \cdot \ln(8) + \dots \\ &= \ln(2^{(1/2)}) + \ln(4^{(1/4)}) + \ln(8^{(1/8)}) + \dots \\ &= \ln(2^{(1/2)} \cdot 4^{(1/4)} \cdot 8^{(1/8)} \cdot \dots) \\ &= \ln(4) \\ &\approx 1.3863 \end{aligned}$$

□

Problem 5:

Consider one of the investors. If she invests in one dollar of her own currency, then her expected wealth is clearly one dollar. However, if she invests in one dollar of foreign currency, then her expected wealth is $\frac{1}{2} \cdot \frac{1}{1.25} + \frac{1}{2} \cdot 1.25 = 1.025$. So, she should invest in foreign currency. This logic applies for both investors.

Problem 6:

Part a.

Proof. Let U be uniform on $[0, 1]$. That is, the density of U is $f(u) = 1$ on the unit interval and 0 everywhere else. Let $X = 1/\sqrt{U}$. Then we have:

$$\begin{aligned} E(X) &= \int_0^1 \frac{1}{\sqrt{u}} du = 2 \\ E(X^2) &= \int_0^1 \frac{1}{u} du = +\infty \end{aligned}$$

□

Part b.

Proof. Consider the CDF of X as defined in part a for $x \geq 1$:

$$F_X(x) = P(X < x) = P\left(\frac{1}{\sqrt{U}} < x\right) = P\left(U > \frac{1}{x^2}\right) = 1 - \frac{1}{x^2}$$

Using the CDF, we know that the median of X is $\sqrt{2}$. So, we may define a random variable Y to be some constant $c \in [\sqrt{2}, 2]$, say 1.5, with probability 1. Then $P(Y > X) = \frac{5}{9} > \frac{1}{2}$, but $E[Y] = 1.5 < 2 = E[X]$. \square

Part c.

Consider random variables $X = Y = Z = N(0, 1)$. Then $P(X > Y)P(Y > Z)P(X > Z) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} > 0$ by the symmetry of the normal distribution and $E(X) = E(Y) = E(Z) = 0$

Problem 7:

Part a.

Proof. Consider the CDF of Y :

$$\begin{aligned} P(Y \leq y) &= P(X \leq y)P(Z = 1) + P(X \geq -y)P(Z = -1) \\ &= P(X \leq y)P(Z = 1) + P(X \leq y)P(Z = -1) \quad \text{by the symmetry of the normal distribution} \\ &= P(X \leq y) \cdot \frac{1}{2} + P(X \leq y) \cdot \frac{1}{2} \\ &= P(X \leq y) \end{aligned}$$

So, since $X \sim N(0, 1)$, we must have $Y \sim N(0, 1)$. \square

Part b.

Proof. Since Z is either -1 or 1 , we have that:

$$|Y| = |XZ| = |X|$$

So, clearly $P(|X| = |Y|) = 1$ \square

Part c.

Proof. Since $Y \sim N(0, 1)$, we have that:

$$P(-1 \leq Y \leq 1) \approx 0.6827$$

However, given $X > 1$, we have that:

$$P(-1 \leq Y \leq 1 \mid X > 1) = 0$$

So, Y is dependent on X . \square

Part d.

Proof.

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

□

Part e.

This assertion is false, as demonstrated by the examples X and Y as defined above.

Problem 8:

Proof. Consider the CDF of M :

$$\begin{aligned}
F_M(x) &= P(M \leq x) \\
&= P((X_1 \leq x) \cap (X_2 \leq x) \cap \dots \cap (X_n \leq x)) && \text{by maximality of } M \\
&= P(X_1 \leq x) \cdot P(X_2 \leq x) \cdot \dots \cdot P(X_n \leq x) && \text{since the } X_i\text{s are i.i.d.} \\
&= x \cdot x \cdot \dots \cdot x && \text{since the } X_i\text{s are uniform} \\
&= x^n
\end{aligned}$$

Taking the derivative of the CDF, we get the PDF:

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

Finally, we have that:

$$\begin{aligned}
E[X] &= \int_0^1 xf_M(x)dx \\
&= \int_0^1 nx^n dx \\
&= \frac{n}{n+1}
\end{aligned}$$

Consider the CDF of m :

$$\begin{aligned}
F_m(x) &= P(m \leq x) \\
&= 1 - P(m \geq x) \\
&= 1 - P((X_1 \geq x) \cap (X_2 \geq x) \cap \dots \cap (X_n \geq x)) \quad \text{by minimality of } m \\
&= 1 - P(X_1 \geq x) \cdot P(X_2 \geq x) \cdot \dots \cdot P(X_n \geq x) \quad \text{since the } X_i\text{s are i.i.d.} \\
&= 1 - (1 - x) \cdot (1 - x) \cdot \dots \cdot (1 - x) \quad \text{since the } X_i\text{s are uniform} \\
&= 1 - (1 - x)^n
\end{aligned}$$

Taking the derivative of the CDF, we get the PDF:

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

Finally, we have that:

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

□

Problem 9:

Part a.

Proof. We can consider the model as a sequence of n i.i.d. Bernoulli trials, labeled $\{X_1, X_2, \dots, X_n\}$. We know that $n = 1000$, $\mu = 0.5$, and $\sigma = 0.5$. Define the following:

$$S_n := \sum_{i=1}^n X_i \quad y_1 := \frac{490 - n\mu}{\sqrt{n}\sigma} \quad y_2 := \frac{510 - n\mu}{\sqrt{n}\sigma}$$

Then, plugging in the appropriate values of n , μ , and σ , we may use the CLT:

$$\begin{aligned}
P(490 \leq S_{1000} \leq 510) &\approx \frac{1}{\sqrt{2\pi}} \int_{y_1}^{y_2} e^{-x^2/2} dx \\
&\approx 0.47291
\end{aligned}$$

□

Part b.

Proof. Using the CLT, we may solve for y :

$$0.99 = \frac{1}{\sqrt{2\pi}} \int_{-y}^y e^{-x^2/2} dx$$

$$y \approx 2.57583$$

We know that $y = \frac{S_n - n\mu}{\sqrt{n}\sigma}$. Also, we know that $\frac{S_n}{n} - \mu = 0.01$. So, we have that:

$$y = \frac{0.01\sqrt{n}}{\sigma}$$

$$n \approx 16587$$

□

Problem 10:

Suppose $E[e^{\theta X}] = 1$. Since $\ln(x)$ concave, by Jensen's Inequality we have that:

$$E[\ln(e^{\theta X})] \leq \ln(E[e^{\theta X}])$$

$$E[\theta X] \leq \ln(1)$$

$$\theta E[X] \leq 0 \quad \text{since } \theta \text{ constant}$$

$$\theta \geq 0 \quad \text{since } E[X] < 0$$

$$\theta > 0 \quad \text{since } \theta \neq 0$$