Homework 1

1. **Estimating logarithm function.** For $x \in [0, 1)$, we shall use the identity that

$$\ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots.$$

(a) **(5 points)** Prove that $\ln(1 - x) \leqslant -x - \frac{x^2}{2}$.

For all $x \in [0, 1)$, $-x$ is negative, $-\frac{x^2}{2}$ is negative, $-\frac{x^3}{3}$ is negative, and so on. This means that the estimation for $ln(1 - x)$ becomes increasingly negative.

This also means that $-x - \frac{x^2}{2}$ will always be greater than or equal to $-x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots$ because of the fact that the two equations share their first two terms ($-x - \frac{x^2}{2}$), but the first equation has more negative terms.

In the case that $x = 0$, the two will be equal. In any other case where $x \in [0, 1)$, equation 1 will be less than equation 2. Therefore, it can be said that $ln(1 - x) \leqslant -x - \frac{x^2}{2}$.

(b) **(10 points)** For $x \in [0, 1/2]$, prove that

$$\ln(1-x) \geqslant -x - \frac{x^2}{2 \cdot 2^0} - \frac{x^2}{2 \cdot 2^1} - \frac{x^2}{2 \cdot 2^2} - \frac{x^2}{2 \cdot 2^3} - \cdots = -x - x^2.$$

From part a, we know that, for $x \in [0, 1/2]$ **(which is a subset of** $x \in [0, 1)$**),**
$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots$.

If we simplify the equation given in this part and compare it with the one from part a, we can see the following:
$ln(1-x) \qquad\qquad \geqslant -x - \frac{x^2}{2*2^0} - \frac{x^2}{2*2^1} - \cdots$
$=$
$-x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots \geqslant -x - \frac{x^2}{2} - \frac{x^2}{4} - \cdots$

For each equation, the first two terms are identical $(-x - \frac{x^2}{2})$**. This means that the only terms to compare are the 3rd terms and beyond. If we set up an inequality using the 3rd term:**
$-\frac{x^3}{3} \geqslant -\frac{x^2}{4}$
$\frac{x^3}{3} \leqslant \frac{x^2}{4}$
$\frac{x}{3} \leqslant \frac{1}{4}$
$x \leqslant \frac{3}{4}$

This means that $-x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots \geqslant -x - \frac{x^2}{2} - \frac{x^2}{4} - \cdots$ **for any** $x \leqslant \frac{3}{4}$**. In this problem,** $x \in [0, 1/2]$**, which is less than** $3/4$**. Therefore, we can conclude that** $\ln(1-x) \geqslant -x - \frac{x^2}{2 \cdot 2^0} - \frac{x^2}{2 \cdot 2^1} - \frac{x^2}{2 \cdot 2^2} - \frac{x^2}{2 \cdot 2^3} - \cdots = -x - x^2$.

2. **Tight Estimations.** Provide meaningful upper-bounds and lower-bounds for the following expressions.

   (a) **(10 points)** $S_n = \sum_{i=1}^{n} \ln i$,

   This summation looks like the following:
   $ln(1) + ln(2) + \cdots + ln(n)$
   Which means that the function $f(x) = ln(x)$, an increasing function.

   For an increasing function $f(x)$,
   $\int_0^x f(t)\,\mathrm{d}t. \leqslant f(x) \leqslant \int_1^{x+1} f(t)\,\mathrm{d}t$

   Therefore,
   $\int_0^x ln(t)\,\mathrm{d}t. \leqslant ln(x) \leqslant \int_1^{x+1} ln(t)\,\mathrm{d}t$

   However, since the natural log of 0 is not possible, we increase the bounds of the lower-bound integral of $ln(x)$. To compensate, we add 1 to the resulting equation, giving our lower bound a possible error of up to 1. The resulting evaluation follows.
   $1 + \int_1^x ln(t)\,\mathrm{d}t. \leqslant ln(x) \leqslant \int_1^{x+1} ln(t)\,\mathrm{d}t$
   Left Side: $1 + \int_1^x ln(t)\,\mathrm{d}t = 1 + (t * ln(t) - t)\Big|_x^1 = 2 + x * ln(x) - x$
   Right Side: $\int_1^{x+1} ln(t)\,\mathrm{d}t = (t * ln(t) - t)\Big|_1^{x+1} = (x+1) * ln(x+1) - x$

   Therefore, the final bounds are:
   $2 + x * ln(x) - x \leqslant \sum_{i=1}^{n} \ln i \leqslant (x+1) * ln(x+1) - x$

(b) **(10 points)** $A_n = n!$

$L_n \leqslant ln(n!) \leqslant U_n$
$e^{L_n} \leqslant n! \leqslant e^{U_n}$

From part (a), we know that:
$2 + x * ln(x) - x \leqslant \sum_{i=1}^{n} \ln i \leqslant (x+1) * ln(x+1) - x$

Therefore,
$e^{2+x*ln(x)-x} \leqslant n! \leqslant e^{(x+1)*ln(x+1)-x}$

(c) **(10 points)** $B_n = \binom{2n}{n} = \frac{(2n)!}{(n!)^2}$

When estimating fractions, we know that:
$\frac{L_{X_n}}{U_{Y_n}} \leqslant \frac{X_n}{Y_n} \leqslant \frac{U_{X_n}}{L_{Y_n}}$

From part (b), we know that:
$e^{2+x*ln(x)-x} \leqslant n! \leqslant e^{(x+1)*ln(x+1)-x}$

So if $X_n = (2n)!$**,**
Using the information from part (b), it follows that:
$e^{2+(2x)*ln(2x)-2x} \leqslant (2n)! \leqslant e^{(2x+1)*ln(2x+1)-2x}$

And if $Y_n = (n!)^2$**,**
Using the information from part (b), it follows that:
$(e^{2+x*ln(x)-x})^2 \leqslant (n!)^2 \leqslant (e^{(x+1)*ln(x+1)-x})^2$

Therefore,
$\frac{e^{2+(2x)*ln(2x)-2x}}{(e^{(x+1)*ln(x+1)-x})^2} \leqslant \frac{(2n)!}{(n!)^2} \leqslant \frac{e^{(2x+1)*ln(2x+1)-2x}}{(e^{2+x*ln(x)-x})^2}$

3. **Understanding Joint Distribution.** Recall that in the lectures we considered the joint distribution $(\mathbb{T}, \mathbb{B})$ over the sample space $\{4, 5, \ldots, 10\} \times \{\mathsf{T}, \mathsf{F}\}$, where $\mathbb{T}$ represents the time I wake up in the morning, and $\mathbb{B}$ represents whether I have breakfast or not. The following table summarizes the joint probability distribution.

t	b	$\mathbb{P}[\mathbb{T} = t, \mathbb{B} = b]$
4	T	0.01
4	F	0.05
5	T	0
5	F	0.04
6	T	0.1
6	F	0.20
7	T	0.25
7	F	0.10
8	T	0.10
8	F	0.05
9	T	0.03
9	F	0.05
10	T	0.01
10	F	0.01

Calculate the following probabilities.

(a) **(5 points)** Calculate the probability that I wake up at 8 a.m. or earlier, but do not have breakfast. That is, calculate $\mathbb{P}[\mathbb{T} \leqslant 8, \mathbb{B} = \mathsf{F}]$,

$\mathbb{P}[\mathbb{T} \leqslant 8, \mathbb{B} = \mathsf{F}] = \mathbb{P}[\mathbb{T} = 4, \mathbb{B} = \mathsf{F}] + \mathbb{P}[\mathbb{T} = 5, \mathbb{B} = \mathsf{F}] + \mathbb{P}[\mathbb{T} = 6, \mathbb{B} = \mathsf{F}] + \mathbb{P}[\mathbb{T} = 7, \mathbb{B} = \mathsf{F}] + \mathbb{P}[\mathbb{T} = 8, \mathbb{B} = \mathsf{F}]$

$\mathbb{P}[\mathbb{T} \leqslant 8, \mathbb{B} = \mathsf{F}] = 0.05 + 0.04 + 0.2 + 0.1 + 0.05 = 0.44$

(b) **(5 points)** Calculate the probability that I wake up at 8 a.m. or earlier. That is, calculate $\mathbb{P}[\mathbb{T} \leqslant 8]$,

$\mathbb{P}[\mathbb{T} \leqslant 8] = \mathbb{P}[\mathbb{T} = 4] + \mathbb{P}[\mathbb{T} = 5] + \mathbb{P}[\mathbb{T} = 6] + \mathbb{P}[\mathbb{T} = 7] + \mathbb{P}[\mathbb{T} = 8]$
$= (0.01 + 0.05) + (0 + 0.04) + (0.1 + 0.2) + (0.25 + 0.1) + (0.1 + 0.05)$

$\mathbb{P}[\mathbb{T} \leqslant 8] = 0.9$

(c) **(5 points)** Calculate the probability that I skip breakfast conditioned on the fact that I woke up at 8 a.m. or earlier. That is, compute $\mathbb{P}\left[\mathbb{B} = \mathsf{F} \mid \mathbb{T} \leqslant 8\right]$.

By Bayes Rule,
$\mathbb{P}\left[\mathbb{B} = \mathsf{F} \mid \mathbb{T} \leqslant 8\right] = \frac{\mathbb{P}[\mathbb{B} = \mathsf{F}, \mathbb{T} \leqslant 8]}{\mathbb{P}[\mathbb{T} \leqslant 8]}$

We know the numerator from part (a), and the denominator from part (b). Therefore,
$\mathbb{P}\left[\mathbb{B} = \mathsf{F} \mid \mathbb{T} \leqslant 8\right] = \frac{\mathbb{P}[\mathbb{B} = \mathsf{F}, \mathbb{T} \leqslant 8]}{\mathbb{P}[\mathbb{T} \leqslant 8]} = \frac{0.44}{0.9} = 0.4888$

4. **Random Walk.** There is a frog sitting at the origin $(0, 0)$ in the first quadrant of a two-dimensional Cartesian plane. The frog first jumps uniformly at random along the X-axis to some point $(\mathbb{X}, 0)$, where $\mathbb{X} \in \{1, 2, 3, 4, 5, 6\}$. Then, it jumps uniformly at random along the Y-axis to some point $(\mathbb{X}, \mathbb{Y})$, where $\mathbb{Y} \in \{1, 2, 3, 4, 5, 6\}$. So $(\mathbb{X}, \mathbb{Y})$ represents the final position of the frog after these two jumps. Note that $\mathbb{X}$ and $\mathbb{Y}$ are two independent random variables that are uniformly distributed over their respective sample spaces.

   (a) **(5 points)** What is the probability that the frog jumps more than 3 units along the Y-axis. That is, compute $\mathbb{P}[\mathbb{Y} > 3]$.

   There are 6 equally likely distances for the frog to jump on the Y-axis. Out of those 6, 3 distances are greater than 3 – 4, 5, and 6. Therefore,
   $\mathbb{P}[\mathbb{Y} > 3] = 0.5$

   (b) **(10 points)** What is the probability that the final position of the frog is above the line $X + Y = 7$. That is compute $\mathbb{P}[\mathbb{X} + \mathbb{Y} > 7]$?

   There are 36 separate combinations of jumps. Out of those 36 combinations, there are 6 that satisfy $X + Y = 7$:
   1 and 6
   2 and 5
   3 and 4
   4 and 3
   5 and 2
   6 and 1

   Therefore,
   $\mathbb{P}[\mathbb{X} + \mathbb{Y} > 7] = \frac{6}{36} = \frac{1}{6} = 0.1666$

(c) **(10 points)** What is the probability that the frog has jumped 2 units along X-axis conditioned on the fact that its final position is above the line $X + Y = 7$? That is, compute $\mathbb{P}\left[\mathbb{X} = 4 \mid \mathbb{X} + \mathbb{Y} > 7\right]$?

By Bayes Rule:
$\mathbb{P}\left[\mathbb{X} = 4 \mid \mathbb{X} + \mathbb{Y} > 7\right] = \frac{\mathbb{P}[\mathbb{X}=4, \mathbb{X}+\mathbb{Y}>7]}{\mathbb{P}[\mathbb{X}+\mathbb{Y}>7]}$
And by Chain Rule:
$\frac{\mathbb{P}[\mathbb{X}=4, \mathbb{X}+\mathbb{Y}>7]}{\mathbb{P}[\mathbb{X}+\mathbb{Y}>7]} = \frac{\mathbb{P}[\mathbb{X}=4] * \mathbb{P}\left[\mathbb{X}+\mathbb{Y}>7 \mid \mathbb{X}=4\right]}{\mathbb{P}[\mathbb{X}+\mathbb{Y}>7]}$

We know the denominator from part (b), and we can reduce $\mathbb{P}\left[\mathbb{X} + \mathbb{Y} > 7 \mid \mathbb{X} = 4\right]$ to $\mathbb{P}[\mathbb{Y} > 3]$, which we know from part (a). The final equation is:
$\frac{\mathbb{P}[\mathbb{X}=4] * \mathbb{P}[\mathbb{Y}>3]}{\mathbb{P}[\mathbb{X}+\mathbb{Y}>7]} = \frac{\frac{1}{6} * \frac{1}{2}}{\frac{1}{6}} = \frac{1}{2} = 0.5$

5. **Coin Tossing Word Problem.** We have three (independent) coins represented by random variables $\mathbb{C}_1, \mathbb{C}_2$, and $\mathbb{C}_3$.

   (i) The first coin has $\mathbb{P}[\mathbb{C}_1 = H] = \frac{1}{4}, \mathbb{P}[\mathbb{C}_2 = T] = \frac{3}{4}$,
   (ii) The second coin has $\mathbb{P}[\mathbb{C}_2 = H] = \frac{3}{4}$ and $\mathbb{P}[\mathbb{C}_2 = T] = \frac{1}{4}$, and
   (iii) The third coin has $\mathbb{P}[\mathbb{C}_3 = H] = \frac{1}{4}$ and $\mathbb{P}[\mathbb{C}_3 = T] = \frac{3}{4}$.

   Consider the following experiment.

   (A) Toss the first coin. Let the outcome of the first coin-toss be ω_1.
   (B) If $\omega_1 = H$, then we toss the second coin twice. Otherwise, (i.e., if $\omega_1 = T$) toss the third coin twice. Let the two outcomes of this step be represented by ω_2 and ω_3.
   (C) Output $(\omega_1, \omega_2, \omega_3)$.

   Based on this experiment, compute the probabilities below.

   (a) **(5 points)** In the experiment mentioned above, what is the probability that a majority of the three outcomes $(\omega_1, \omega_2, \omega_3)$ are H (head)?

   There are four different possibilities where the majority of the three outcomes can be H:
   (H H H), (H H T), (H T H), (T H H)

   By the Chain Rule,
   $\mathbb{P}[\omega_1 = H, \omega_2 = H, \omega_3 = H] = \mathbb{P}[\omega_1 = H] * \mathbb{P}\left[\omega_2 = H \mid \omega_1 = H\right] * \mathbb{P}\left[\omega_3 = H \mid \omega_1 = H, \omega_2 = H\right]$
   $\mathbb{P}[\omega_1 = H, \omega_2 = H, \omega_3 = H] = \frac{1}{4} * \frac{3}{4} * \frac{3}{4} = \frac{9}{64} = 0.140625$

   $\mathbb{P}[\omega_1 = H, \omega_2 = H, \omega_3 = T] = \mathbb{P}[\omega_1 = H] * \mathbb{P}\left[\omega_2 = H \mid \omega_1 = H\right] * \mathbb{P}\left[\omega_3 = T \mid \omega_1 = H, \omega_2 = H\right]$
   $\mathbb{P}[\omega_1 = H, \omega_2 = H, \omega_3 = T] = \frac{1}{4} * \frac{3}{4} * \frac{1}{4} = \frac{3}{64} = 0.046875$

   $\mathbb{P}[\omega_1 = H, \omega_2 = T, \omega_3 = H] = \mathbb{P}[\omega_1 = H] * \mathbb{P}\left[\omega_2 = T \mid \omega_1 = H\right] * \mathbb{P}\left[\omega_3 = H \mid \omega_1 = H, \omega_2 = T\right]$
   $\mathbb{P}[\omega_1 = H, \omega_2 = T, \omega_3 = H] = \frac{1}{4} * \frac{1}{4} * \frac{3}{4} = \frac{3}{64} = 0.046875$

   $\mathbb{P}[\omega_1 = T, \omega_2 = H, \omega_3 = H] = \mathbb{P}[\omega_1 = T] * \mathbb{P}\left[\omega_2 = H \mid \omega_1 = T\right] * \mathbb{P}\left[\omega_3 = H \mid \omega_1 = T, \omega_2 = H\right]$
   $\mathbb{P}[\omega_1 = T, \omega_2 = H, \omega_3 = H] = \frac{1}{4} * \frac{1}{4} * \frac{1}{4} = \frac{1}{64} = 0.015625$

   Summing everything up, we get that the probability of a majority of the tree outcomes $(\omega_1, \omega_2, \omega_3)$ **are** H **(head) is:**
   $\frac{9}{64} + \frac{3}{64} + \frac{3}{64} + \frac{1}{64} = \frac{16}{64} = \frac{1}{4} = 0.25$

(b) **(5 points)** In the experiment mentioned above, what is the probability that a majority of the three outcomes are H, conditioned on the fact that the first outcome was T?

There is only one case in which the majority of the three outcomes are H, conditioned on the fact that the first outcome was T:
$\mathbb{P}[\omega_1 = T, \omega_2 = H, \omega_3 = H]$

By Chain Rule,
$\mathbb{P}[\omega_1 = T, \omega_2 = H, \omega_3 = H] = \mathbb{P}[\omega_1 = T] * \mathbb{P}\left[\omega_2 = H \mid \omega_1 = T\right] * \mathbb{P}\left[\omega_3 = H \mid \omega_1 = T, \omega_2 = H\right] = \frac{3}{4} * \frac{1}{4} * \frac{1}{4} = \frac{3}{64} = 0.046875$

(c) **(5 points)** In the experiment mentioned above, what is the probability that a majority of the three outcomes are different from the first outcome?

There are only two scenarios where a majority of the three outcomes are different from the first outcome, so the equation will be as follows:
$\mathbb{P}[\omega_1 = H, \omega_2 = T, \omega_3 = T] + \mathbb{P}[\omega_1 = T, \omega_2 = H, \omega_3 = H]$

By Chain Rule,
$\mathbb{P}[\omega_1 = H, \omega_2 = T, \omega_3 = T] + \mathbb{P}[\omega_1 = T, \omega_2 = H, \omega_3 = H] =$
$\mathbb{P}[\omega_1 = H] * \mathbb{P}\left[\omega_2 = T \mid \omega_1 = H\right] * \mathbb{P}\left[\omega_3 = T \mid \omega_1 = H, \omega_2 = T\right] +$
$\mathbb{P}[\omega_1 = T] * \mathbb{P}\left[\omega_2 = H \mid \omega_1 = T\right] * \mathbb{P}\left[\omega_3 = H \mid \omega_1 = T, \omega_2 = H\right]$
$= \frac{1}{4} * \frac{1}{4} * \frac{1}{4} + \frac{3}{4} * \frac{1}{4} * \frac{1}{4} = \frac{1}{64} + \frac{3}{64} = \frac{4}{64} = \frac{1}{16} = 0.0625$