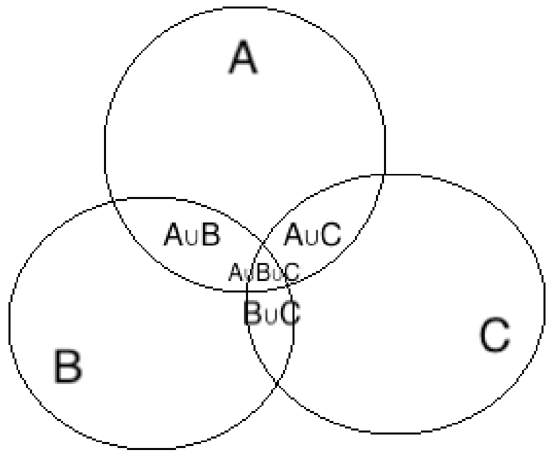
Michael Huang
EN.625.633.81
Homework 1

**1.**

**(a)**




Intuitively, we first start with adding up the full area of $P(A)$, $P(B)$, and $P(C)$. We see that for each of these, we are double/triple counting when trying to take the union:

- We double count $P(A) \cup P(B)$, from both $P(A)$ and $P(B)$.
- We double count $P(B) \cup P(C)$, from both $P(B)$ and $P(C)$.
- We double count $P(A) \cup P(C)$, from both $P(A)$ and $P(C)$.
- We triple count $P(A) \cup P(B) \cup P(C)$, from $P(A)$, $P(B)$, and $P(C)$.

To avoid the double counting, let's first subtract one of each of those relevant sections above. However, through this, we've also subtracted the innermost center, $P(A) \cup P(B) \cup P(C)$, three times, where this was previously triple counted, so we end up with 0 – we need to add one of these sections back. This gives us the final expression of

$$P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

as we sought to show.

**(b)**

$$
\begin{aligned}
P(A \cup B \cup C) && \text{Given} \\
P((A \cup B) \cup C) && \text{Assoc. prop. of union} \\
P(A \cup B) + P(C) - P((A \cup B) \cap C) && \text{Def. of union} \\
P(A) + P(B) - P(A \cap B) + P(C) - P((A \cup B) \cap C) && \text{Def. of union} \\
P(A) + P(B) - P(A \cap B) + P(C) - P((A \cap C) \cup (B \cap C)) && \text{Distr. prop. of intersection} \\
P(A) + P(B) - P(A \cap B) + P(C) && \\
-P(A \cap C) - P(B \cap C) + P(A \cap B \cap A \cap C) && \text{Assoc. prop. of union} \\
P(A) + P(B) - P(A \cap B) + P(C) && \\
-P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) && \text{Prop. of intersection}
\end{aligned}
$$

which ultimately simplifies to

$$P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

as we exactly sought to prove.

**2.**

We are given $p = 0.05$. Given the independent trials and the fact that we want to find the minimum number of throws $n$ such that the chance of having at least one hit in those $n$ throws is greater than 0.5. As calculating the probability of at least one hit being greater than 0.5 is messy as we need to compute for all terms from 1 to $n$, let's instead take the complement of this, which is easier to calculate. We set this up as

$$P(\text{one hit in n}) = 1 - P(\text{no hits in n}) = 1 - 0.95^n$$

We can then solve accordingly:

$$
\begin{aligned}
1 - 0.95^n &\geq 0.5 && \text{Given} \\
-0.95^n &\geq -0.5 && \text{Arithmetic} \\
0.95^n &\leq 0.5 && \text{Arithmetic} \\
n \cdot \ln(0.95) &\leq \ln(0.5) && \text{Property of logarithm} \\
n &\geq \frac{\ln(0.5)}{\ln(0.95)} && \text{As } \ln(0.95) \text{ is negative} \\
n &\geq 13.513407334 && \text{Arithmetic}
\end{aligned}
$$

As we need an integer number of throws, we find the minimum number of throws to be $\boxed{\mathbf{14}}$.

**3.**

As the outcomes of each player shooting is deemed to be independent, we can model the PMF for the result for the number of attempts by breaking it into two scenarios: Player A makes the first shot, or Player B makes the first shot. Since Player A goes first, we can say that the make is on an odd number of attempts, and Player B making it would be on an even number of attempts. We therefore break this down into

$$P(N) = P(N = 2k - 1) + P(N = 2k)$$

where $N$ is the number of attempts. Given that they make with probability $p_1$ and $p_2$, we can break each scenario down using the complements of these probabilities, with the miss probabilities multiplied together until the applicable make probability:

$$P(N = 2k - 1) = ((1 - p_1) \cdot (1 - p_2))^{k-1} \cdot p_1$$

for odd-numbered makes and

$$P(N = 2k) = ((1 - p_1) \cdot (1 - p_2))^{k-1} \cdot (1 - p_1) \cdot p_2$$

for even numbered-makes, with $k \geq 1$. Putting this together, and substituting $n = 2k$ or $k = \frac{n}{2}$ and $n = 2k - 1$ or $k = \frac{n+1}{2}$, we have the PMF for the make being on the $N$th shot being

$$\boxed{P(N = n) = \begin{cases} \left((1 - p_1) \cdot (1 - p_2)\right)^{\frac{n-1}{2}} p_1, & \text{if } n \text{ is odd} \\ \left((1 - p_1) \cdot (1 - p_2)\right)^{\frac{n}{2}-1} (1 - p_1) p_2, & \text{if } n \text{ is even} \end{cases}}$$

for $n = 1, 2, 3, \ldots$

**4.**

We can use the uniform distribution here, i.e. we can model the point at which we cut with $X \sim \text{Unif}(0, 1)$. We want to find $P(\text{Longer Piece} > 2 \cdot \text{Shorter Piece})$,which we can divide into the following scenarios:

Scenario 1: We cut above the halfway point, so $X > \frac{1}{2}$:

$$X > 2(1 - X)$$

$$X > 2 - 2X$$

$$3X > 2$$

$$X > \frac{2}{3}$$

Scenario 2: We cut below the halfway point, so $X < \frac{1}{2}$:

$$1 - X > 2X$$

$$1 - X > 2X$$

$$1 > 3X$$

$$X < \frac{1}{3}$$

Putting this together, we can find the total probability of the scenario to be

$$P(X \geq \frac{2}{3}) + P(X \leq \frac{1}{3})$$

which we very intuitively interpret using the Uniform distribution to be

$$\frac{1}{3} + \frac{1}{3} = \boxed{\mathbf{\frac{2}{3}}}$$

**5.**

We are given the binomial r.v. $np$, i.e. $X \sim \text{Binom}(n, p)$. We can first define that we have $n$ trials in the binomial form and take the expected value by setting up a Bernoulli r.v. with equivalent success rate $p$. We can then sum the total number of successes into $X$ by breaking it down into $X_i$ for each $i$th trial, i.e.

$$X = X_1 + X_2 + X_3 + \ldots + X_n$$

with success being 1 and failure being 0. By definition, we have a bunch of Bernoulli r.v.s that we have now strung together into $n$ trials, which is exactly what a binomial r.v. is. We take the expected value of $X$:

$$E[X] = E[X_1 + X_2 + X_3 + \ldots + X_n]$$

And use the property of expectation of r.v.s being the expectation of its components, i.e. linearity of expectation:

$$E[X] = E[X_1] + E[X_2] + E[X_3] + \ldots + E[X_n]$$

And using the definition of the expected value of a Bernoulli variable, we can find that

$$E[X] = p + p + p + \ldots + p = np$$

as we sought to find.

For variance, we know that each $X_i$ is independent by definition of the binomial variable, and that $Var(X_i) = E[X_i^2] - (E[X_i])^2$. We know that for each trial, $E[X_i^2] = E[X_i]$ since the value is just 0 or 1. Therefore, we know that $Var(X_i) = E[X_i^2] - (E[X_i])^2 = E[X_i] - (E[X_i])^2 = p - p^2$. Putting this together:

$$Var[X] = Var[X_1 + X_2 + X_3 + \ldots + X_n]$$

$$Var(X) = Var(X_1) + Var(X_2) + Var(X_3) + \ldots + Var(X_n)$$

$$Var(X) = p - p^2 + p - p^2 + p - p^2 + \ldots + p - p^2$$

$$Var(X) = n(p - p^2) = np(1 - p)$$

exactly as we sought to find as well.

**6.**

We are given an experiment with $X_1, X_2, \ldots, X_n \sim$ Unif$[0,1]$ with all $X_i$ being independent. To find the pdf of $X_{(1)} = \min(X_1, \ldots, X_i)$

We can approach this by finding the cdf and taking the derivative of the cdf. By definition, we can find the cdf of the uniform distribution to be

$$F_X(x) = P(X \leq x) = P(X_{(1)} \leq x)$$

To make this easier, we instead take the complement and continue deriving:

$$\begin{aligned}
P(X_{(1)} \leq x) &= 1 - P(X_{(1)} > x) && \text{Complement} \\
&= 1 - P(X_1 > x \cap X_2 > x \cap \ldots \cap X_n > x) && \text{Definition of } X_{(1)} \\
&= 1 - P(X_1 > x) \cap P(X_2 > x) \cap \ldots \cap P(X_n > x) && \text{Property of set intersection} \\
&= 1 - (1 - x) \cdot (1 - x) \cdot \ldots \cdot (1 - x) && \text{Definition of uniform distr.} \\
&= 1 - (1 - x)^n && \text{Property of set intersection}
\end{aligned}$$

Now, to find the pdf, we differentiate:

$$\begin{aligned}
f_{X_{(1)}}(x) &= \frac{d}{dx}(1 - (1 - x)^n) && \text{Definition of pdf} \\
&= \frac{d}{dx}(1) - \frac{d}{dx}(1 - x)^n && \text{Differentiation} \\
&= 0 - \frac{d}{dx}(1 - x)^n && \text{Differentiation} \\
= -nu^{n-1} \cdot \frac{du}{dx} &= -n(1 - x)^{n-1} \cdot (0 - 1) && \text{Chain rule, with } u = 1 - x \\
&= \boxed{n(1 - x)^{n-1}} && \text{Arithmetic}
\end{aligned}$$

for $0 \leq x \leq 1$, and 0 otherwise.

### 7.

#### (a)

We are given $Y \sim e^X$, with $X \sim N(0,1)$. We'll use the methodology from the transformation section. From this, we know that we can determine that the pdf

$$f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}(y)}{dy}|$$

We know that $Y \sim g(X) \sim e^X$, so

$$g(y) = e^y$$

$$g^{-1}(y) = \ln(y)$$
$$\frac{d}{dy}g^{-1}(y) = \frac{d}{dy}\ln(y) = \frac{1}{y}$$

Now we can substitute accordingly. We know as well by definition that for $X \sim N(0,1)$, $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Putting this all together:

$$f_Y(y) = f_X(g^{-1}(y))|\frac{dg^{-1}(y)}{dy}|$$

$$f_Y(y) = f_X(\ln(y))|\frac{1}{y}|$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\ln(y)^2}{2}}|\frac{1}{y}|$$

$$\boxed{f_Y(y) = \frac{1}{y} \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{\ln(y)^2}{2}}}$$

for $y > 0$, as by definition $Y \sim e^X > 0$ always, and 0 otherwise.

#### (b)

We want to find $E[Y] = E[e^X]$. We know that in general, $E[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy$. Given the pdf we found in part (a), let's start by plugging in:

$$E[Y] = \int_{-\infty}^{\infty} y \cdot \frac{1}{y} \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{\ln(y)^2}{2}}dy$$

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{\ln(y)^2}{2}} dy$$

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{\ln(y)^2}{2}} dy$$

We can perform integration by substitution. Using $x = \ln(y)$ (and subsequently, $y = e^x$ and $dy = e^x dx$):

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \cdot e^x dx$$

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{x-\frac{x^2}{2}} dx$$

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{x-\frac{x^2}{2}} dx$$

We want to make the exponent of the $e$ term more interpretible, in order to get something with the form of the normal pdf, as we know that generally for $Z \sim N(\mu, \sigma^2)$, $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$. We complete the square accordingly to get to this form:

$$x - \frac{x^2}{2} = -\frac{1}{2}(x^2 - 2x)$$

Set up with $(x-1)^2 = x^2 + 1 - 2x$:

$$x^2 - 2x + 1 = (x-1)^2$$

$$x^2 - 2x = (x-1)^2 - 1$$

$$-\frac{1}{2}(x^2 - 2x) = -\frac{1}{2}((x-1)^2 - 1) = \frac{1}{2} - \frac{(x-1)^2}{2}$$

So substituting back:

$$E[Y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{1}{2}-\frac{(x-1)^2}{2}} dx$$

$$E[Y] = e^{\frac{1}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-1)^2}{2}} dx$$

And we use the fact that we know the form of the normal pdf to find it is a pdf with $\mu = 1$, $\sigma = 1$. The integral of this kind of distribution is simply 1, so:

$$E[Y] = e^{\frac{1}{2}} \cdot 1 = \boxed{e^{\frac{1}{2}}}$$