Week 2

Day 2

Counting Techniques

(The mn Rule [Fundamental Counting Principle])

- First stage = m & Second stage = n
- mn ways to accomplish an experiment

(Extended mn Rule)

- k stages with n_1 ways for the first stage, n_2 ways for the second stage, and n_k ways to for the k^{th} stage
- $\prod_{x=1}^{k} n_x$

Permutation

Distinct Permutation

- Permutation of n objects is $n!$
- Permutation of n objects r at a time

$$
\begin{aligned}
P(n, r) &= \frac{n!}{(n-1)!} \\
&= nPr \\
&= P_r^n
\end{aligned}
$$

- Note that $P(n, n) = n!$

Repeating Permutation

- Permutation of n objects with k types

$$\frac{n!}{n_1! n_2! \cdots n_k!}$$

Example

> How many ways can 9 beads (3 red, 2 yellow, 4 blue) lined up

$$\frac{9!}{3!2!4!} = 1260$$

Circular Permutation

- Permutation in a circle with one fixed object is $(n-1)!$

Example

> 6 different colored beads (**r** **b** **y** **p** **bl** **g**) into a bracelet if **red**, **blue**, **yellow** are together, and **purple**, **black** must not be adjacent

$$
\begin{aligned}
3!(4-1)! &= 36 && \text{rby together} \\
3!2!(3-1)! &= 24 && \text{rby together and pbl together} \\
36 - 24 &= 12 && \text{rby together, pbl not}
\end{aligned}
$$

Combination

- Counting without arrangement

$$C(n, r) = \frac{n!}{r!(n-r)!}$$

Example

> Possible combinations in 7 balls drawn from a set of 42 numbered balls

$$
\begin{aligned}
C(42, 7) &= \binom{42}{7} \\
&= \frac{42!}{7!(42-7)!} \\
&= 26\,978\,328
\end{aligned}
$$

Probability

Uniform Probability Model

$$\begin{aligned} P(A) &= \frac{\text{no. of simple events in } A}{\text{no. of simple events in } S} \\ &= \frac{|A|}{|S|} \end{aligned}$$

- In cases where out comes are <u>not equally likely to occur</u>:

$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

- Note that:
  1. $0 \leq P(A) \leq 1$
  2. $P(\varnothing) = 0$
  3. $P(S) = 1$
- Union and Intersection

$$\begin{aligned} P(A \cup B) &= \frac{|A \cup B|}{|S|} \\ &= P(A) + P(B) \end{aligned}$$

$$P(A \cap B) = \frac{|A \cap B|}{|S|}$$

- Note the **Addition Rule**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Compliment $\longrightarrow$ $P(A^C) = 1 - P(A)$

Conditional Probability

Dependent Events

- An event occurring affects the probability of the following event
- Probability of A given event B has occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Note the **Multiplication Rule**

$$\begin{aligned} P(A \cap B) &= P(A|B) \cdot P(B) \\ &= P(B|A) \cdot P(A) \end{aligned}$$

Independent Events

- Probability of one does not affect the other
- Independent if

$$
P(A|B) = P(A) \\
\text{or} \\
P(B|A) = P(B)
$$

$$
\text{or} \\
P(A \cap B) = P(A) \cdot P(B)
$$

- Mutual Independence
  - Events $A_1, A_2, \cdots, A_n$ are mutually independent if each pair of events A_i and A_j are independent.

$$
P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \cdots \cdot P(A_n)
$$

Day 3

Probability Distribution

- A formula, table, or graph that gives all the possible values k of the discrete random variable X, and the probability $p_X(k) = P(X = k)$ associated with each value
- $p_X(k) \geq 0$
- $\sum_{\text{all } k} p_x(k) = 1$

Example

```
Toss two fair coins and let $X$ be the number of heads observed. Find the probability distribution for $X$.
```

Simple Event	**Coin 1**	**Coin 2**	Probability of Simple Event $P(E_i)$	**Number of Heads Observed** X
E_1	H	H	$\frac{1}{4}$	2
E_2	H	T	$\frac{1}{4}$	1
E_3	T	H	$\frac{1}{4}$	1
E_4	T	T	$\frac{1}{4}$	0

Probability Distribution Function (pdf) $p_X(k)$

$$p_X(k) = \begin{cases} \frac{1}{4} & \text{if } k = 0 \\ \frac{1}{2} & \text{if } k = 1 \\ \frac{1}{4} & \text{if } k = 2 \end{cases}$$

Probability Distribution Table

k	$p_X(k)$
0	$\frac{1}{4}$
1	$\frac{1}{2}$
2	$\frac{1}{4}$

Cumulative Distribution $F_X(k)$

- formula, table or graph that gives all the possible values k and $F_X(k) = P(X \leq k)$, the probability that X is at most k

k	$F_X(k)$
0	$p_X(0) = \frac{1}{4}$
1	$p_X(0) + p_X(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$
2	$p_X(0) + p_X(1) + p_X(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

Mean or Expected Value

- The average value of X in the population

$$\mu = E(X) = \sum_{\text{all } k} k \cdot p_X(k)$$

Standard Deviation and Variance

- Standard Deviation
  - Measures the spread or variability of the random variable

$$\begin{aligned}\sigma &= \sqrt{E((X-\mu)^2)} \\ &= \sqrt{\sum_{\text{all } k} (k-\mu)^2 \cdot p_X(k)}\end{aligned}$$

- Variance

$$\begin{aligned}\sigma^2 &= E((X-\mu)^2) \\ &= \sum_{\text{all } k} (k-\mu)^2 \cdot p_X(k)\end{aligned}$$

Binomial Distribution

- Experiment consists of n identical trials
- Each trial results in **one of two** outcomes
- The probability of success on a single trial is equal to p and remains from trial to trial. Failure, $q = 1 - p$
- Trials are independent
- Each trial is called a Bernoulli Trial

Example

```
33 distinguishable biased coins ⟶ 0.60 heads.
```

Coin 1	Coin 2	Coin 3	Number of Heads	Probability
H	H	H	3	0.216
H	H	T	2	0.144
H	T	H	2	0.144
H	T	T	1	0.096
T	H	H	2	0.144
T	H	T	1	0.096
T	T	H	1	0.096
T	T	T	0	0.064

Probability Distribution Function

- If p is the probability of success in n Bernoulli Trials, then the probability of k successes:

$$p_X(k) = P(X = k) = nCp \cdot p^k \cdot (1-p)^{n-k}$$

$$\text{for } k = 0, 1, \cdots, n$$
$$\text{aka. } X \sim B(n, p)$$

Mean, Variance, Standard Deviation

$$\mu = np$$

$$\sigma^2 = np(1-p)$$

$$\sigma = \sqrt{np(1-p)}$$

Excel

- Probability Dist Func

```
=BINOM.DIST(X,N,p,FALSE)
```

- Cumulative Dist Func

```
=BINOM.DIST(X,N,p,TRUE)
```

Examples

> 6 chips (N Total), 2 defective (D Defective). $X = \text{defective} = 0, 1, 2$.
> Find probability distribution

1. $\binom{D}{k}$ $\rightarrow$ ways to choose k defective chips
2. $\binom{N-D}{n-k}$ $\rightarrow$ ways to choose the rest $(n-k)$
3. $\binom{N}{n}$ $\rightarrow$ total number of possible ways to choose any n chips

$$\frac{\binom{2}{k}\binom{4}{3-k}}{\binom{6}{3}}$$

k	0	1	2
$p_X(k)$	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

> 4 laptops, probability distribution

k	0	1	2	3	4	5
$p_X(k)$	0.1	0.4	0.2	0.15	0.1	0.05

> Find mean, variance, standard deviation

$$\begin{aligned}\mu &= \sum_{k=0}^{5} k \cdot p_X(k) \\ &= 0(0.10) + 1(0.40) + 2(0.20) + 3(0.15) + 4(0.10) + 5(0.05) \\ &= 1.90\end{aligned}$$

$$\begin{aligned}\sigma^2 &= \sum_{k=0}^{5} (k - 1.90)^2 px(k) \\ &= (0 - 1.90)^2(0.10) + (1 - 1.90)^2(0.40) + \cdots + (5 - 1.90)^2(0.05) \\ &= 1.79 \\ \sigma &= \sqrt{1.79} \\ &\approx 1.34\end{aligned}$$

> 5000 tickets @ Php 100 each. Jackpot = Php 25,00. Expected gain of 4 tickets?

$$\begin{aligned}\mu = E(X) &= \sum_{k\in\{-400,\, 24\,600\}} k \cdot p_X(k) \\ &= (-400) \cdot \frac{4\,996}{5\,000} + 24\,600 \cdot \frac{4}{5\,000} \\ &= -380\end{aligned}$$

Probability Distribution Continuous Variable

Probability Density Function $f_X(x)$

- For all values x of X
  - $f_X(x) \geq 0$
  - $\int_{-\infty}^{\infty} f_X(x)\, dx = 1 \quad \text{(the total area under the curve)}$
- Integration is actually done over all values x that X can assume

$$P(a \leq X \leq b) = \int_a^b f_X(x)\, dx$$

- Rule Satisfaction
  - $P(X = a) = 0$
  - $P(X \geq a) = P(X > a)$
  - $P(X \leq a) = P(X < a)$

- $P(X > a) = 1 - F_X(a)$
- $\lim_{x \to \infty} F_X(x) = 1$
- $\lim_{x \to -\infty} F_X(x) = 0$

Excel

```
=NORM.DIST(X^,N,p,TRUE) - NORM.DIST(X⌄,N,P,TRUE)
```

Cumulative Distribution Function $F_X(x)$

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt$$

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$$

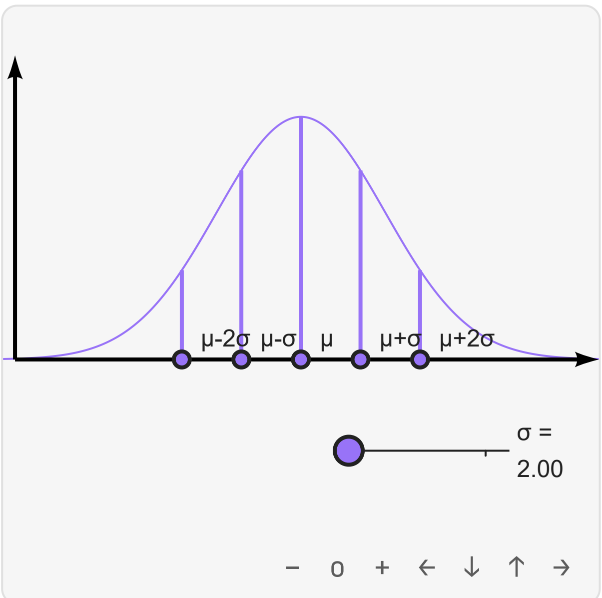
$$\sigma = \sqrt{E((X-\mu)^2)} = \sqrt{\int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_X(x)\, dx}$$

$$\sigma^2 = \mathrm{Var}(X) = E((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_X(x)\, dx$$

Excel

```
=NORM.INV(AREA,σ,μ)
```

Normal Distribution



$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\dfrac{-(x-\mu)^2}{2\sigma^2}}, \; -\infty < x < \infty$$

- Large values of σ reduce the height of the curve and increase the spread
- $X \sim N\left(\mu, \sigma\right)$

```
#### ~={blue}Gaussian Curve=~
```

$$X \sim N\left(0, 1\right)$$

$$or$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \quad -\infty < x < \infty$$

- $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.6827$
- $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.9545$
- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.9973$

Standardization

- Expressing a normal random variable $X \sim N\left(\mu, \sigma\right)$ as the number of standard deviations it lies to the left or the right of its mean

$$Z = \frac{X - \mu}{\sigma} \sim N\left(0, 1\right)$$