Lecture 7

Part 2 Probability and Distributions

Combination and permutation formula: A different approach for probability

Combination

Question: Number of ways to select 3 people from a group of 10 to form a committee.

- A combination is a selection of items from a collection, where the order does not matter.
- It represents the number of ways to choose k items from n distinct elements.
- Denoted as $C(n, k)$ or $\binom{n}{k}$.

Combination Formula

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

Example: Number of ways to select 3 people from a group of 10 to form a committee.

$$C(10,3) = \frac{10!}{3!(10-3)!} = \frac{10!}{3!7!} = 120$$

Permutation

Question: Number of ways to select and arrange 3 people from a group of 10 (e.g., as president, vice-president, secretary).

- A permutation is an arrangement of objects where order matters.
- It represents the number of ways to arrange k items from n distinct elements.
- Denoted as $P(n,k)$ or ${}_nP_k$.

Permutation Formula

$$P(n, k) = \frac{n!}{(n-k)!}$$

Example: Number of ways to select and arrange 3 people from a group of 10 (e.g., as president, vice-president, secretary).

$$P(10, 3) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 720$$

Examples with Colored Balls

Suppose we have a box containing 5 red balls, 4 blue balls, and 3 green balls.

> Question 1: What is the probability of drawing 3 balls, one of each color?
>
> Question 2: What is the probability of drawing 4 balls with at least one green ball?

Examples with Colored Balls

Suppose we have a box containing 5 red balls, 4 blue balls, and 3 green balls.

Question 1: What is the probability of drawing 3 balls, one of each color?

$$\frac{C(5,1)\cdot C(4,1)\cdot C(3,1)}{C(12,3)} = \frac{5\cdot 4\cdot 3}{220} = \frac{60}{220} = \frac{3}{11}$$

Question 2: What is the probability of drawing 4 balls with at least one green ball?

$$1 - \frac{C(9,4)}{C(12,4)} = 1 - \frac{126}{495} = \frac{369}{495} = \frac{41}{55}$$

Let's continue our discussion about distribution.

Variance

- Let X be a discrete random variable with probability distribution $p(x)$ and $\mu = E(X)$.
- The (population) variance of X is defined as:

$$\sigma^2 = V(X) = E((X-\mu)^2) = \sum_{\text{all } x}((x-\mu)^2 \times p(x))$$

- A shortcut formula for the variance is given below:

$$V(X) = E(X^2) - (E(X))^2$$

$$= \left(\sum_{\text{all } x}(x^2 \times p(x))\right) - \mu^2$$

Variance and Standard Deviation Calculation

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$
\begin{aligned}
V(X) &= \left(\sum_{\text{all } x} (x^2 \times p(x))\right) - \mu^2 \\
&= \left(0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8}\right) - 1.5^2 \\
&= 0.75
\end{aligned}
$$

$$SD(X) = \sqrt{V(X)} = \sqrt{0.75} = 0.866 = \sigma$$

Laws of Variance

- If X and Y are random variables (discrete or continuous) and c is any constant, then:

$$
\begin{aligned}
&1. \quad V(c) = 0 \\
&2. \quad V(cX) = c^2 V(X) \\
&3. \quad V(X + c) = V(X)
\end{aligned}
$$

- And if X and Y are independent, then:

$$4. \quad V(X + Y) = V(X) + V(Y)$$

$$5. \quad V(X - Y) = V(X) + V(Y)$$

Example

- Let $Z = 3X - 2Y - 7$ with $V(X) = 2, V(Y) = 1$ and X and Y independent. Then:

$$
\begin{aligned}
V(Z) &= V(3X - 2Y - 7) \\
&= V(3X - 2Y) \\
&= V(3X) + V(2Y) \\
&= 9V(X) + 4V(Y) \\
&= 9 \times 2 + 4 \times 1 \\
&= 22
\end{aligned}
$$

Bivariate Distribution

- If X and Y are discrete random variables, then the **bivariate distribution** of X and Y is a table or formula that lists the joint probabilities
  $P(\{X = x\} \cap \{Y = y\})$, denoted $p(x, y)$, for all pairs of x and y.
- A bivariate distribution must satisfy two requirements:

1. $0 \leq p(x, y) \leq 1$ for all x and y
2. $\sum_{\text{all } x} \sum_{\text{all } y} p(x, y) = 1$

Example

- Flip a coin three times.
- Let X be the number of heads.
- Let Y be the number of sequence changes within the three flips, i.e., the number of times we change from $H \Rightarrow T$ or $T \Rightarrow H$.
- For example:
  - HHH: $x = 3$ (3 heads) and $y = 0$ (0 sequence changes since $H \Rightarrow H \Rightarrow H$).
  - HHT: $x = 2$ (2 heads) and $y = 1$ (1 sequence change since $H \Rightarrow H \Rightarrow T$).
  - HTH: $x = 2$ (2 heads) and $y = 2$ (2 sequence changes since $H \Rightarrow T \Rightarrow H$).

Example

Outcome	x	y
HHH	3	0
HHT	2	1
HTH	2	2
THH	2	1
TTH	1	1
THT	1	2
HTT	1	1
TTT	0	0

Example

		y			
		0	1	2	
	0	$\frac{1}{8}$	0	0	$\frac{1}{8}$
x	1	0	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
	2	0	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
	3	$\frac{1}{8}$	0	0	$\frac{1}{8}$
		$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	1

Marginal Probability Distribution

- Just like we did in past class, we can calculate marginal probabilities for X and Y by adding across the rows and down the columns, respectively.
- Specifically, given $p(x, y)$ (the bivariate distribution of X and Y), the **marginal probability distribution** of X is:

$$p_X(x) = P(X = x) = \sum_{\text{all } y} p(x, y)$$

Marginal Probability Distribution

- So for our example,

$$p_X(1) = P(X = 1) = p(1, 0) + p(1, 1) + p(1, 2) = \frac{3}{8}$$

- Considering Y for the moment, notice that the events $\{Y = 0\}$, $\{Y = 1\}$ and $\{Y = 2\}$ are a *partition* of the sample space!
- So, calculating a marginal probability distribution is just a direct consequence of the *Law of Total Probability*.

Marginal Probability Distribution

- The marginal distribution of X is:

x	0	1	2	3
$p_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

- The marginal distribution of Y is:

y	0	1	2
$p_Y(y)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$

Independence of Random Variables

- Two discrete random variables, X and Y, are **independent** if and only if

$$p(x, y) = p_X(x) \times p_Y(y)$$

for all x and y.

- Note that this has to be true for *all* x and y. If there is just *one* pair of x and y for which the above is not true, then X and Y are **not** independent.

Example

- In our previous coin flipping example, X and Y are clearly not independent since if we consider the pair $x = 0$ and $y = 0$:

$$p(0, 0) = \frac{1}{8}$$

but

$$p_X(0) \times p_Y(0) = \frac{1}{8} \times \frac{2}{8} = \frac{1}{32}$$

- That is, we have found one pair for which

$$p(x, y) \neq p_X(x) \times p_Y(y)$$

Sum of Two Random Variables

- Consider two real estate agents, Albert and Bob.
  - Let X be the number of houses sold by Albert in a week.
  - Let Y be the number of houses sold by Bob in a week.

		x			
		0	1	2	$P_Y(y)$
	0	0.12	0.42	0.06	0.6
y	1	0.21	0.06	0.03	0.3
	2	0.07	0.02	0.01	0.1
$P_X(x)$		0.4	0.5	0.1	1

Sum of Two Random Variables

- From the marginal probability distributions of X and Y, it is straightforward to calculate the following:
  - $\mathbb{E}(X) = 0.7$
  - $V(X) = 0.41$
  - $\mathbb{E}(Y) = 0.5$
  - $V(Y) = 0.45$
- Suppose we are interested in the *total* number of houses Albert and Bob sell in a week.

Sum of Two Random Variables

- That is, we are interested in the quantity $X + Y$, which itself is a random variable.
- From the bivariate distribution table, we know the possible values of $X + Y$ are $0, 1, 2, 3$ or 4.
- Suppose we want to find the probability that a total of two houses were sold in a week, i.e., $P(X + Y = 2)$.

Sum of Two Random Variables

- From the table, we can find $P(X+Y=2)$ by summing up all the joint probabilities for the values of x and y which give $x+y=2$.
- That is,

$$\begin{aligned} P(X+Y=2) &= p(0,2)+p(1,1)+p(2,0) \\ &= 0.07+0.06+0.06 \\ &= 0.19 \end{aligned}$$

Sum of Two Random Variables

- We can repeat this for $X + Y = 0, 1, 3$ and 4 to obtain the probability distribution for $X + Y$:

$x + y$	0	1	2	3	4
$p(x + y)$	0.12	0.63	0.19	0.05	0.01

- From this we can calculate the mean and variance of $X + Y$:

$$E(X + Y) = 1.2$$
$$V(X + Y) = 0.56$$

Functions of Two Random Variables

- Note that we could use the same approach to calculate the probability distribution of any function of two discrete random variables.
- For example:
  - $g(X, Y) = XY$
  - $g(X, Y) = \sqrt{XY^3}$
  - $g(X, Y) = \frac{X}{Y+1}$
  - etc.

Expected Value

- If X and Y are two discrete random variables with bivariate distribution $p(x,y)$ and $g(X,Y)$ is some function of X and Y, the **expected value** of $g(X,Y)$ is given by:

$$E(g(X,Y)) = \sum_{\text{all } x} \sum_{\text{all } y} (g(x,y) \times p(x,y))$$

Covariance

- Let X and Y be discrete random variables with joint probability distribution $p(x, y)$.
- If we denote $E(X) = \mu_X$ and $E(Y) = \mu_Y$, then the **(population) covariance** between X and Y is:

$$\begin{aligned}\sigma_{XY} &= Cov(X, Y) \\ &= E((X - \mu_X)(Y - \mu_Y)) \\ &= \sum_{\text{all } x} \sum_{\text{all } y} ((x - \mu_X)(y - \mu_Y) \times p(x, y))\end{aligned}$$

Covariance

- Just like with the variance, there is a shortcut formula for calculating the covariance:

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$= \left(\sum_{\text{all } x} \sum_{\text{all } y} (xy \times p(x,y)) \right) - \mu_X \mu_Y$$

Correlation Coefficient

- The **(population) correlation coefficient** is defined in exactly the same way as before:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

- Remember that the correlation always lies between -1 and 1, i.e., $-1 \leq \rho_{XY} \leq 1$.

Example

- Flip a coin three times.
- X is the number of heads, Y is the number of sequence changes.
- We know that:
  - $\mu_X = \frac{3}{2}$
  - $\sigma_X^2 = \frac{3}{4}$
  - $\mu_Y = 1$
  - $\sigma_Y^2 = \frac{1}{2}$

Example

$$Cov(X,Y) = \left(\sum_{\text{all } x} \sum_{\text{all } y} (xy \times p(x,y)) \right) - \mu_X \mu_Y$$

$$= \left(0 \times 0 \times \frac{1}{8} + 1 \times 1 \times \frac{2}{8} + 1 \times 2 \times \frac{1}{8} \right.$$

$$\left. +2 \times 1 \times \frac{2}{8} + 2 \times 2 \times \frac{1}{8} + 3 \times 0 \times \frac{1}{8} \right)$$

$$- \frac{3}{2} \times 1$$

$$= 0$$

Independence and Being Uncorrelated

- This implies $\rho_{XY} = 0$ so X and Y are *uncorrelated*.
- But remember we showed previously that X and Y were *not* independent!
- Independence and being uncorrelated are *not* the same thing.
- In fact, independence is a stronger condition than being uncorrelated.
- Specifically, independence always implies a correlation of zero, whereas being uncorrelated does not always imply independence.

Linear Combination of Random Variables

- The quantity $Z = aX + bY$, where a and b are constants, is called a **linear combination** of the random variables X and Y.
- It can be shown that:

$$
\begin{aligned}
E(aX + bY) &= aE(X) + bE(Y) \\
V(aX + bY) &= a^2V(X) + b^2V(Y) + 2ab \times Cov(X, Y) \\
&= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho_{XY}\sigma_X\sigma_Y
\end{aligned}
$$

Application in Finance: Portfolio Diversification

- In finance, variance or standard deviation is often used to assess the risk of an investment.
- Analysts reduce risk by diversifying their investments - that is, combining investments where the correlation is small.

Portfolio Diversification

- An investor forms a portfolio by putting 25% of his money in stock A and 75% in stock B, with population parameters given below.

	Expected Value of Return	Standard Deviation of Return
Stock A	8%	12%
Stock B	15%	22%

Expected Portfolio Return

- Let R_A and R_B denote the returns of stocks A and B, respectively.
- If we let R_P denote the return of the portfolio, then we can write:

$$R_P = 0.25R_A + 0.75R_B$$

- We are given that $E(R_A) = 8$ and $E(R_B) = 15$.

Expected Portfolio Return

- Therefore, the expected value of R_P is:

$$\begin{aligned}\mathbb{E}(R_P) &= \mathbb{E}(0.25R_A + 0.75R_B)\\ &= 0.25 \times E(R_A) + 0.75 \times E(R_B)\\ &= 0.25 \times 8 + 0.75 \times 15\\ &= 13.25\end{aligned}$$

- That is, the expected portfolio return is 13.25%.

Variance of Portfolio Return

- Calculate the variance when the two stock returns are perfectly positively correlated, i.e., $\rho_{AB} = 1$:

$$
\begin{aligned}
V(R_P) &= 0.25^2\sigma_A^2 + 0.75^2\sigma_B^2 + 2 \times 0.25 \times 0.75 \times \rho_{AB}\sigma_A\sigma_B \\
&= 0.25^2 \times 12^2 + 0.75^2 \times 22^2 + 2 \times 0.25 \times 0.75 \times \rho_{AB} \times 12 \\
&= 281.25 + 99 \times \rho_{AB} \\
&= 281.25 + 99 \times 1 \\
&= 380.25\%^2
\end{aligned}
$$

Variance of Portfolio Return

- Calculate the variance when the two stock returns are perfectly uncorrelated, i.e., $\rho_{AB} = 0$:

$$
\begin{aligned}
V(R_P) &= 0.25^2\sigma_A^2 + 0.75^2\sigma_B^2 + 2 \times 0.25 \times 0.75 \times \rho_{AB}\sigma_A\sigma_B \\
&= 0.25^2 \times 12^2 + 0.75^2 \times 22^2 + 2 \times 0.25 \times 0.75 \times \rho_{AB} \times 12 \\
&= 281.25 + 99 \times \rho_{AB} \\
&= 281.25 + 99 \times 0 \\
&= 281.25\%^2
\end{aligned}
$$

Bernoulli Trial

- A **Bernoulli trial** is a random experiment that has the following special properties:
  - On each trial there are only two possible outcomes, which we call success and failure.
  - On any given trial, the probability of a success is p and the probability of a failure is $1 - p$.
  - The trials are independent - that is, the result of one trial does not affect the result of any other trial.

Binomial Distribution

- If a fixed number, n, of Bernoulli trials are performed, the random variable representing the number of successes in the n trials is called a **binomial random variable** and its probability distribution is called the **binomial distribution**.
- If X denotes a binomial random variable, then we use the notation $X \sim Bin(n, p)$, where p is the probability of success on any given trial.

Some Examples

- Flip a coin ten times and let X be the number of heads.
  - $X \sim Bin(n = 10, p = 0.5)$.
- Pull a card from a deck, with replacement, eight times and let X be the number of clubs.
  - $X \sim Bin(n = 8, p = 0.25)$.

Binomial Probability Distribution

- If $X \sim Bin(n, p)$ then the possible values that X can take are $0, 1, 2, 3, \ldots, n$.
- The **binomial probability distribution** is given by the following formula:

$$P(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Note that $n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$.

Expected Value and Variance

- We could use the usual formula to calculate the expected value:

$$
\begin{aligned}
E(X) &= \sum_{\text{all } x} (x \times p(x)) \\
&= \sum_{x=0}^{n} \left(x \times \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right) \\
&= \dots
\end{aligned}
$$

- And similarly for the variance.
- But we don't really want to.

Expected Value and Variance

- Instead, let's define a new random variable for each Bernoulli trial as follows:

$$X_i = \begin{cases} 1 & \text{if trial } i \text{ is a success} \\ 0 & \text{if trial } i \text{ is a failure} \end{cases}$$

- Each X_i is called a **Bernoulli** or **indicator variable**.
- We know that the X_i are independent and we also know that

$$X = \sum_{i=1}^{n} X_i$$

Expected Value and Variance

- Using the laws of expected value and variance:

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np$$

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

Example

- A student sitting a statistics quiz decides to answer each of the ten multiple choice questions entirely by chance.
- Each question has five options, only one of which is correct.
- Let X be the number of questions the student answers correctly.
- Then $X \sim Bin(n = 10, p = 0.2)$.