# Chapter 7 part 1

## Properties of Expectations

Jingchen (Monika) Hu

Vassar College

MATH 241

# Outline

# Expected value of $g(X, Y)$

Recap: expectation of random variable $g(X)$

- Discrete case $E[g(X)] = \sum_{\text{all } x} g(x)f(x)$
- Continuous case $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$

Suppose $g(X, Y)$ is a real-valued function of random variables $X$ and $Y$, then

- Discrete case

$$E[g(X, Y)] = \sum_{\text{all } x} \sum_{\text{all } y} g(x, y)f(x, y)$$

- Continuous case

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y)dxdy$$

# Expectation of sums of two random variables

$$E(X + Y) = E(X) + E(Y)$$

# Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

- It's not difficult to show that if either (or both) of the $X, Y$ is discrete, this formula still holds.
- This results does not require $X$ and $Y$ to be independent.
- This can be generalized to $n$ random variables

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

**Question**

What's the expected value of $X - Y$?

# Expectation of sums of two random variables

$$E(X+Y) = E(X) + E(Y)$$

- It's not difficult to show that if either (or both) of the $X, Y$ is discrete, this formula still holds.
- This results does not require $X$ and $Y$ to be independent.
- This can be generalized to $n$ random variables

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

**Question**

What's the expected value of $X - Y$?

$$E(X-Y) = E[X + (-Y)] = E(X) + E(-Y) = E(X) - E(Y)$$

## Question

Suppose that $n$ people throw their hats into the center of a room. The hats are mixed up, and each person randomly selects one. Find the expected number of people that select their own hat.

## Question

Suppose that $n$ people throw their hats into the center of a room. The hats are mixed up, and each person randomly selects one. Find the expected number of people that select their own hat.

Let $X$ denote the total number of matches.

$$X_i = \begin{cases} 1 & \text{if the } i \text{ th person selects his own hat} \\ 0 & \text{otherwise} \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_n$$

## Question

Suppose that $n$ people throw their hats into the center of a room. The hats are mixed up, and each person randomly selects one. Find the expected number of people that select their own hat.

Let $X$ denote the total number of matches.

$$X_i = \begin{cases} 1 & \text{if the } i \text{ th person selects his own hat} \\ 0 & \text{otherwise} \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_n$$

For any $i = 1, 2, \ldots, n$,

$$E(X_i) = 1/n \Longrightarrow$$

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = n/n = 1$$

## Question

Let $X_1, ..., X_n$ be independent and identically distributed random variables having distribution function $F$ and expected value $\mu$. Such a sequence of random variables is said to constitute a sample from the distribution $F$. Then quantity

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

is called the sample mean. Compute $E[\bar{X}]$.

## Question

Let $X_1, ..., X_n$ be independent and identically distributed random variables having distribution function $F$ and expected value $\mu$. Such a sequence of random variables is said to constitute a sample from the distribution $F$. Then quantity

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

is called the sample mean. Compute $E[\bar{X}]$.

$$E[\bar{X}] = E[\sum_{i=1}^{n} \frac{X_i}{n}] = \frac{1}{n} E[\sum_{i=1}^{n} X_i] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \mu$$

### Question

Let $X_1, ..., X_n$ be independent and identically distributed random variables having distribution function $F$ and expected value $\mu$. Such a sequence of random variables is said to constitute a sample from the distribution $F$. Then quantity

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

is called the sample mean. Compute $E[\bar{X}]$.

$$E[\bar{X}] = E[\sum_{i=1}^{n} \frac{X_i}{n}] = \frac{1}{n} E[\sum_{i=1}^{n} X_i] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \mu$$

- The expected value of the sample mean is $\mu$ (the distribution mean).
- When the distribution mean $\mu$ is unknown, the sample mean is often used in statistics to estimate it.

# Recap

Expectation of sum

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

$X_1, X_2, \ldots, X_n$ are independent $\Longrightarrow$ $\not\Longleftarrow$

$$E[X_1 X_2 \cdots X_n] = E[X_1]E[X_2] \cdots E[X_n]$$

# Outline

# Covariance

## Definition

*Covariance of two random variables $X$ and $Y$ is defined as*

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

- Simplification

$$\begin{aligned} Cov(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E[XY + \mu_X\mu_Y - X\mu_Y - Y\mu_X]) \\ &= E[XY] - \mu_X\mu_Y \end{aligned}$$

- Recall

$$\begin{aligned} E[XY] &= \int\int xy\ f(x,y)\ dx\ dy \quad \text{if continuous} \\ &= \sum_x \sum_y xy\ f(x,y) \quad \text{if discrete} \end{aligned}$$

# Properties of $Cov(X,Y) = E[XY] - E[X]E[Y]$

- $Cov(X,Y) = Cov(Y,X)$
- $Cov(X,c) = 0$
- $Cov(X,X) = Var(X)$
- $Cov(aX,bY) = ab\ Cov(X,Y)$

$$
\begin{aligned}
Cov(aX,bY) &= E[abXY] - E[aX]E[bY] \\
&= ab\ E[XY] - ab\ E[X]E[Y]
\end{aligned}
$$

- $Cov(X+a,Y+b) = Cov(X,Y)$

$$
\begin{aligned}
Cov(X+a,Y+b) =& E[(X+a)(Y+b)] - E[X+a]E[Y+b] \\
&= E[XY + aY + bX + ab] \\
&\quad - (E[X]+a)(E(Y)+b) \\
&= E[XY] + E[aY] + E[bX] + ab \\
&\quad - E[X]E[Y] - a\ E[Y] - b\ E[X] - ab \\
&= E[XY] - E[X]E[Y]
\end{aligned}
$$

# Covariance of sums of random variables

$$Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov\left(X_i, Y_j\right)$$

A special case

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{1\leq i<j\leq n} Cov(X_i, X_j)$$

Some more special cases

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$
$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$

## Question

Suppose $Z_1$ and $Z_2$ are two standard normal random variables. Let

$$X = Z_1 + Z_2,\ Y = Z_1 - Z_2$$

Find $Cov(X, Y)$.

## Question

Suppose $Z_1$ and $Z_2$ are two standard normal random variables. Let

$$X = Z_1 + Z_2,\ Y = Z_1 - Z_2$$

Find $Cov(X, Y)$.

Method 1.

$$\begin{aligned}
Cov(X, Y) &= Cov(Z_1 + Z_2, Z_1 - Z_2) \\
&= Cov(Z_1, Z_1) + Cov(Z_2, Z_1) + Cov(Z_1, -Z_2) + Cov(Z_2, -Z_2) \\
&= Cov(Z_1, Z_1) + Cov(Z_2, Z_1) - Cov(Z_1, Z_2) - Cov(Z_2, Z_2) \\
&= Var(Z_1) - Var(Z_2) = 0
\end{aligned}$$

## Question

Suppose $Z_1$ and $Z_2$ are two standard normal random variables. Let

$$X = Z_1 + Z_2,\ Y = Z_1 - Z_2$$

Find $Cov(X, Y)$.

Method 1.

$$\begin{aligned} Cov(X, Y) &= Cov(Z_1 + Z_2, Z_1 - Z_2) \\ &= Cov(Z_1, Z_1) + Cov(Z_2, Z_1) + Cov(Z_1, -Z_2) + Cov(Z_2, -Z_2) \\ &= Cov(Z_1, Z_1) + Cov(Z_2, Z_1) - Cov(Z_1, Z_2) - Cov(Z_2, Z_2) \\ &= Var(Z_1) - Var(Z_2) = 0 \end{aligned}$$

Method 2.

$$E[XY] = E[Z_1^2 - Z_2^2] = E[Z_1^2] - E[Z_2^2] = 0$$
$$E[X] = E[Z_1] + E[Z_2] = 0,\ E[Y] = E[Z_1] - E[Z_2] = 0$$

# Zero covariance and independence

- $X$ and $Y$ are independent $\Longrightarrow Cov(X,Y) = 0$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

- $X_1, X_2, \ldots, X_n$ are independent $\Longrightarrow$

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

- $Cov(X,Y) = 0 \not\Longrightarrow X$ and $Y$ are independent
  Counter example?

## Question

Let $X_1, \ldots, X_n$ be independent random variables having the same variance $\sigma^2$, and

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

Find $Var(\bar{X})$.

## Question

Let $X_1, \ldots, X_n$ be independent random variables having the same variance $\sigma^2$, and

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

Find $Var(\bar{X})$.

$$\begin{aligned}
Var(\bar{X}) &= \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\left[\sum_{i=1}^{n} Var(X_i)\right] \\
&= \frac{1}{n^2}\left(n\sigma^2\right) = \frac{\sigma^2}{n}
\end{aligned}$$

# Correlation

Since $Cov(X, Y)$ depends on the magnitude of $X$ and $Y$ we would prefer to have a measure of association that is not effected by arbitrary changes in the scales of the random variables.

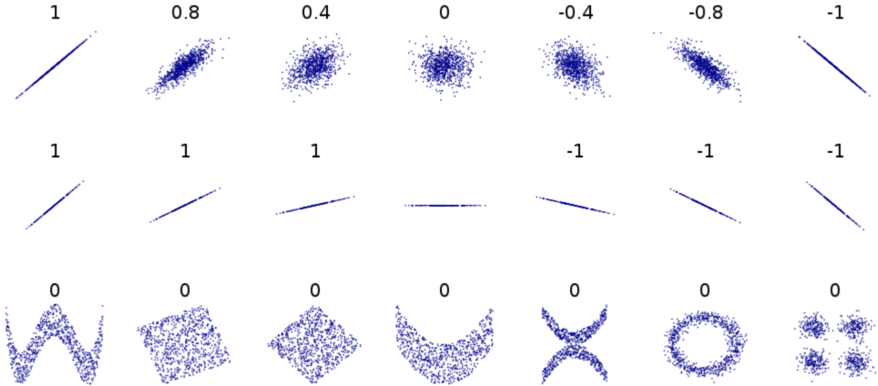
## Definition

*The most common measure of linear association is correlation which is defined as*

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\,\sqrt{Var(Y)}}$$

- range: $-1 \leq \rho(X, Y) \leq 1$
- the magnitude (i.e. absolute value) of the $\rho(X, Y)$ measures the strength of the linear association
- the sign determines if it is a positive or negative relationship.
- if $\rho(X, Y) = 0$, then $X$ and $Y$ are said to be uncorrelated.

# Correlation



### Question

Find the value of $b$ in $Y = a + bX$ such that $\rho(X, Y) = 1$. What about $\rho(X, Y) = -1$?

## Question

Find the value of $b$ in $Y = a + bX$ such that $\rho(X, Y) = 1$. What about $\rho(X, Y) = -1$?

$$
\begin{aligned}
\rho(X, a + bX) &= \frac{Cov(X, a + bX)}{\sqrt{Var(X)}\sqrt{Var(a + bX)}} \\
&= \frac{bVar(X)}{\sqrt{Var(X)}\sqrt{b^2 Var(X)}} \\
&= \frac{bVar(X)}{|b| Var(X)}
\end{aligned}
$$

## Question

Find the value of $b$ in $Y = a + bX$ such that $\rho(X, Y) = 1$. What about $\rho(X, Y) = -1$?

$$
\begin{aligned}
\rho(X, a + bX) &= \frac{Cov(X, a + bX)}{\sqrt{Var(X)}\sqrt{Var(a + bX)}} \\
&= \frac{bVar(X)}{\sqrt{Var(X)}\sqrt{b^2Var(X)}} \\
&= \frac{bVar(X)}{|b|Var(X)}
\end{aligned}
$$

- When $b > 0$, $\rho(X, a + bX) = \rho(X, Y) = 1$
- When $b < 0$, $\rho(X, a + bX) = \rho(X, Y) = -1$

# Outline

# Conditional expectation

- The discrete case: for all $p_Y(y) > 0$
  - Conditional pmf: $p_{X|Y}(x \mid y) = P\{X = x \mid Y = y\} = \frac{p(x,y)}{p_Y(y)}$

## Definition

The conditional expectation of $X$ given that $Y = y$ is

$$E[X \mid Y = y] = \sum_x xP\{X = x \mid Y = y\} = \sum_x xp_{X|Y}(x \mid y)$$

- The continuous case: for all $f_Y(y) > 0$
  - Conditional pdf: $f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)}$

## Definition

The conditional expectation of $X$ given that $Y = y$ is

$$E[X \mid Y = y] = \int_{-\infty}^{\infty} xf_{X|Y}(x \mid y)dx$$

# Properties of expectations remain

- Expectation of a function of a random variable
  - The discrete case:

$$E[g(X) \mid Y = y] = \sum_x g(x) p_{X|Y}(x \mid y)$$

  - The continuous case:

$$E[g(X) \mid Y = y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x \mid y) dx$$

- Expectation of sum of random variables

$$E[\sum_{i=1}^{n} X_i \mid Y = y] = \sum_{i=1}^{n} E[X_i \mid Y = y]$$

- Think about the condition $Y = y$ as taking expectation of $X$ on a reduced sample space consisting only of outcomes for which $Y = y$

# Computing expectations by conditioning

$$E[X] = E[E[X \mid Y]]$$

- A very important property of conditional expectation
- Think of $E[X|Y]$ as a random variable (when $Y = y$)
- The discrete case:

$$E[X] = \sum_{y} E[X \mid Y = y] P\{Y = y\}$$

- Intuition:
  - $E[E[X \mid Y]]$ is a weighted average of $E[X \mid Y]$, where weights are $P\{Y = y\}$ (the probability of the condition)
  - Similar to the "law of total probability" $P(E) = \sum_{i=1}^{n} P(E \mid F_i) P(F_i)$
- The continuous case:

$$E[X] = \int_{-\infty}^{\infty} E[X \mid Y = y] f_Y(y) dy$$

# Computing probabilities by conditioning

- Let $E$ denote an arbitrary event, and define the indicator random variable $X$ as

$$X = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{if } E \text{ does not occur} \end{cases}$$

- Then $E[X] = P(E), E[X \mid Y = y] = P(E \mid Y = y)$ for any random variable $Y$

- The discrete case:

$$P(E) = \sum_y P(E|Y = y)p(Y = y)$$

  related to $P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$

- The continuous case:

$$P(E) = \int_{-\infty}^{\infty} P(E \mid Y = y) f_Y(y) dy$$

# Conditional variance

- Similarly to the conditional expectation, we can define the conditional variance of $X$ given that $Y = y$

## Definition

$$Var(X \mid Y = y) = E[(X - E[X \mid Y = y])^2 \mid Y = y]$$

- A very useful conditional variance formula

$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$

## Question

Type $i$ light bulbs function for a random amount of time having mean $\mu_i$ and standard deviation $\sigma_i$, $i = 1, 2$. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb with probability $p$ and a type 2 bulb with probability $1 - p$. Let $X$ denote the lifetime of this bulb. Find (a) $E[X]$ (b) $Var(X)$.

### Question

Type $i$ light bulbs function for a random amount of time having mean $\mu_i$ and standard deviation $\sigma_i$, $i = 1, 2$. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb with probability $p$ and a type 2 bulb with probability $1-p$. Let $X$ denote the lifetime of this bulb. Find (a) $E[X]$ (b) $Var(X)$.

Let $Y$ be the type of light bulb chosen. $Y = 1$ for choosing type 1 and $Y = 2$ for choosing type 2.

## Question

Type $i$ light bulbs function for a random amount of time having mean $\mu_i$ and standard deviation $\sigma_i$, $i = 1, 2$. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb with probability $p$ and a type 2 bulb with probability $1 - p$. Let $X$ denote the lifetime of this bulb. Find (a) $E[X]$ (b) $Var(X)$.

Let $Y$ be the type of light bulb chosen. $Y = 1$ for choosing type 1 and $Y = 2$ for choosing type 2.

- The mean is $E[X] = E[E[X \mid Y]] = \mu_1 p + \mu_2(1 - p)$

### Question

Type $i$ light bulbs function for a random amount of time having mean $\mu_i$ and standard deviation $\sigma_i$, $i = 1, 2$. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb with probability $p$ and a type 2 bulb with probability $1-p$. Let $X$ denote the lifetime of this bulb. Find (a) $E[X]$ (b) $Var(X)$.

Let $Y$ be the type of light bulb chosen. $Y = 1$ for choosing type 1 and $Y = 2$ for choosing type 2.

- The mean is $E[X] = E[E[X \mid Y]] = \mu_1 p + \mu_2(1-p)$
- The variance is

$$Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y])$$

## Question

Type $i$ light bulbs function for a random amount of time having mean $\mu_i$ and standard deviation $\sigma_i$, $i = 1, 2$. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb with probability $p$ and a type 2 bulb with probability $1-p$. Let $X$ denote the lifetime of this bulb. Find (a) $E[X]$ (b) $Var(X)$.

Let $Y$ be the type of light bulb chosen. $Y = 1$ for choosing type 1 and $Y = 2$ for choosing type 2.

- The mean is $E[X] = E[E[X \mid Y]] = \mu_1 p + \mu_2(1-p)$
- The variance is

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

## Question

Type $i$ light bulbs function for a random amount of time having mean $\mu_i$ and standard deviation $\sigma_i$, $i = 1, 2$. A light bulb randomly chosen from a bin of bulbs is a type 1 bulb with probability $p$ and a type 2 bulb with probability $1-p$. Let $X$ denote the lifetime of this bulb. Find (a) $E[X]$ (b) $Var(X)$.

Let $Y$ be the type of light bulb chosen. $Y = 1$ for choosing type 1 and $Y = 2$ for choosing type 2.

- The mean is $E[X] = E[E[X \mid Y]] = \mu_1 p + \mu_2(1-p)$
- The variance is

$$\begin{aligned} Var(X) &= E[Var(X \mid Y)] + Var(E[X \mid Y]) \\ &= E[Var(X \mid Y)] + E[(E[X \mid Y])^2] - (E[E[X \mid Y]])^2 \\ &= \sigma_1^2 p + \sigma_2^2(1-p) + \mu_1^2 p + \mu_2^2(1-p) - (\mu_1 p + \mu_2(1-p))^2 \end{aligned}$$