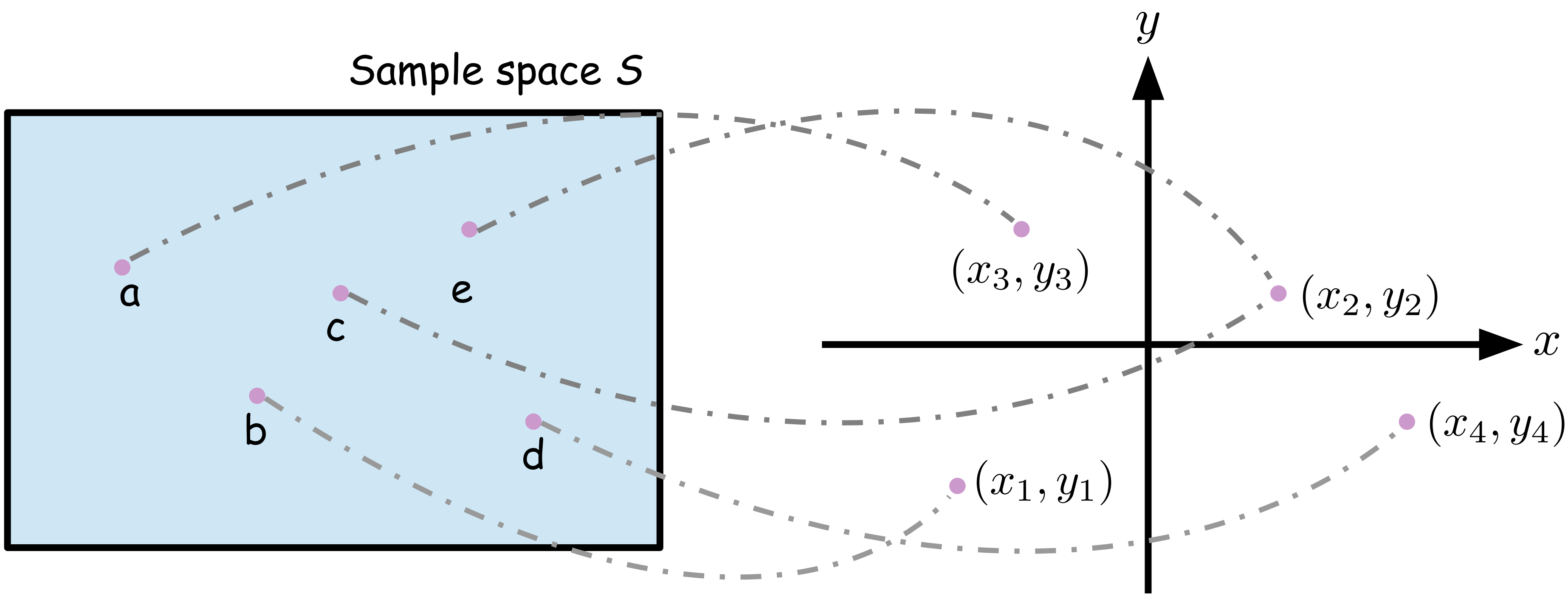
통계분석
Statistical Analysis

Joint Distribution

- We have discussed an **univariate probability distribution** of a **single random variable**.
- We can extend the probability theory to the case of more than two random variables
- The probability distribution involving more than two random variables is called the **joint distribution (multivariate distribution)**.

Joint Distribution



Joint Distribution

Tossing a coin twice

- Sample space

$$S = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$$

- Random variables

$$X = \text{the number of heads} \in \{0, 1, 2\}$$
$$Y = \text{the number of tails} \in \{0, 1, 2\}$$

$$\mathrm{HH} \longrightarrow (x, y) = (2, 0)$$
$$\mathrm{HT} \longrightarrow (x, y) = (1, 1)$$
$$\mathrm{TH} \longrightarrow (x, y) = (1, 1)$$
$$\mathrm{TT} \longrightarrow (x, y) = (0, 2)$$

Joint Probability Table: $p(x, y)$

$Y \backslash X$	0	1	2
0	0	0	$\frac{1}{4}$
1	0	$\frac{1}{2}$	0
2	$\frac{1}{4}$	0	0

Bivariate Distribution (Discrete)

- Two discrete random variables X, Y

$$X = x \in D_1 = \{x_1, x_2, \cdots, x_m\}$$

$$Y = y \in D_2 = \{y_1, y_2, \cdots, y_n\}$$

- Joint distribution (joint probability mass function)

$$P(X = x_i, Y = y_j) = p(x_i, y_j)$$

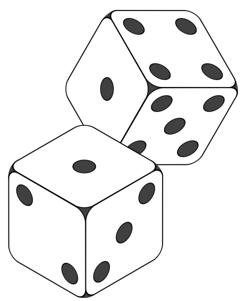
The probability of the joint outcome $X = x_i, Y = y_j$

$$0 \leq p(x_i, y_j) \leq 1$$

$$\sum_{x \in D_1} \sum_{y \in D_2} p(x, y) = 1$$

Bivariate Distribution: Example

Two fair dices



$$X = x \in D = \{1, 2, 3, 4, 5, 6\}$$

$$Y = y \in D = \{1, 2, 3, 4, 5, 6\}$$

For any x and y in D $\quad p(x, y) = \dfrac{1}{36}$

Example 1. Roll two dice. Let X be the value on the first die and let Y be the value on the second die. Then both X and Y take values 1 to 6 and the joint pmf is $p(i, j) = 1/36$ for all i and j between 1 and 6. Here is the joint probability table:

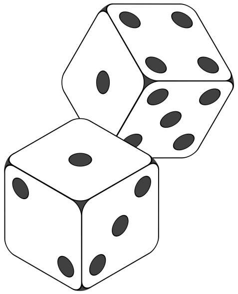
$X \backslash Y$	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Bivariate Distribution: Example

Example 2. Roll two dice. Let X be the value on the first die and let T be the total on both dice. Here is the joint probability table:

$X\backslash T$	2	3	4	5	6	7	8	9	10	11	12
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36

Two fair dices



$$X = x \in D_1 = \{1, 2, 3, 4, 5, 6\}$$
$$T = t \in D_2 = \{2, 3, 4, \cdots, 11, 12\}$$

Marginal Probability Mass Function

- Marginal probability mass function (pmf) of ***X***

$$p_X(x) = \sum_{y \in D_2} p(x, y)$$

- Marginal probability mass function (pmf) of ***Y***

$$p_Y(y) = \sum_{x \in D_1} p(x, y)$$

Marginal PMF: Example

Example 1. Roll two dice. Let X be the value on the first die and let Y be the value on the second die. Then both X and Y take values 1 to 6 and the joint pmf is $p(i,j) = 1/36$ for all i and j between 1 and 6. Here is the joint probability table:

$X\backslash Y$	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

$$p_X(x) = \sum_{y \in D_2} p(x,y) = 6 \cdot \frac{1}{36} = \frac{1}{6}$$

$$p_Y(y) = \sum_{x \in D_1} p(x,y) = 6 \cdot \frac{1}{36} = \frac{1}{6}$$

Marginal PMF: Example

Example 2. Roll two dice. Let X be the value on the first die and let T be the total on both dice. Here is the joint probability table:

$X \backslash T$	2	3	4	5	6	7	8	9	10	11	12	
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0	$1/6$
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	$1/6$
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	$1/6$
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	$1/6$
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	$1/6$
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

$$p_X(x) = \sum_{y \in D_2} p(x, y) = 6 \cdot \frac{1}{36} = \frac{1}{6}$$

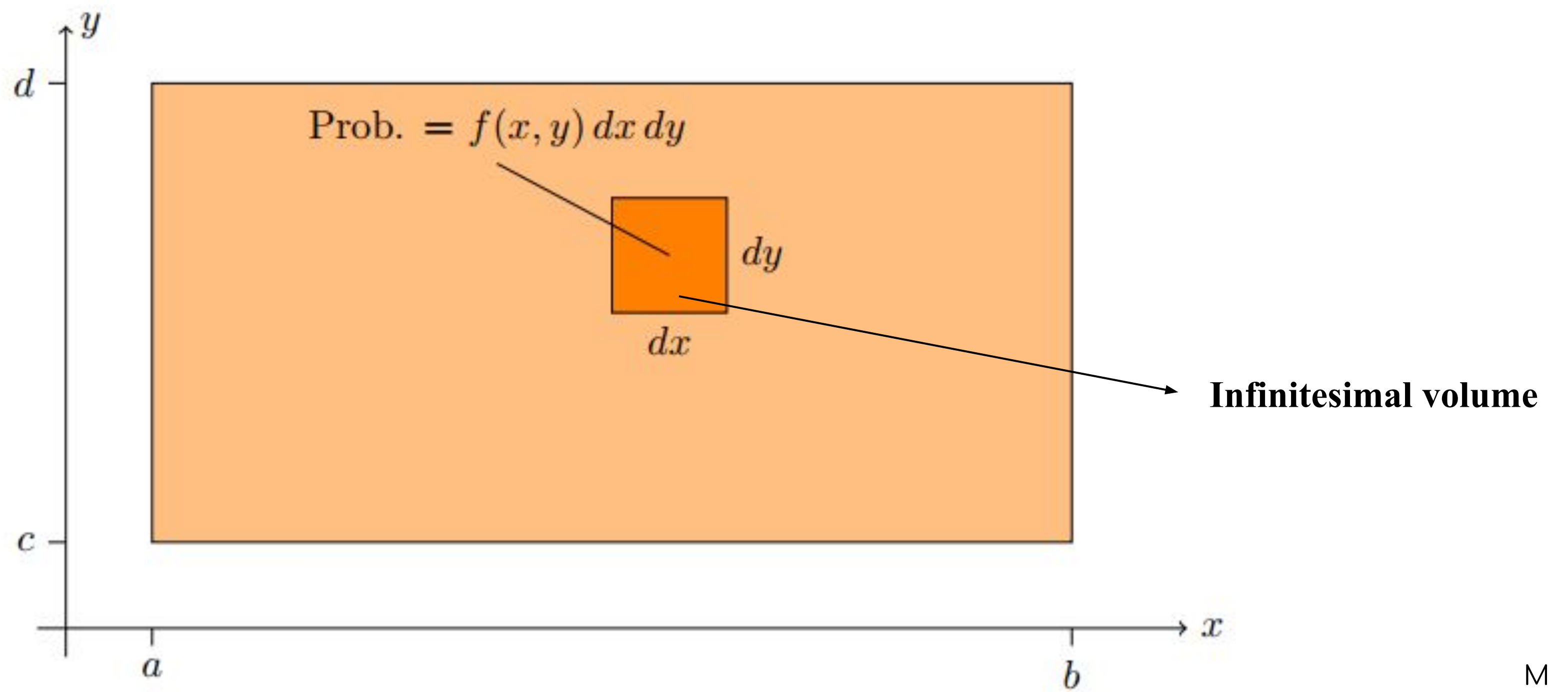
$$p_Y(y) = \sum_{x \in D_1} p(x, y)$$

Joint Bivariate Distribution (Continuous)

Two continuous random variables X, Y $\begin{cases} X = x \in (-\infty, \infty) \\ Y = y \in (-\infty, \infty) \end{cases}$

$$f(x,y) = \text{joint probability } \underline{\text{density}} \text{ at } (x,y)$$

$$\begin{aligned} dp(x,y) &= f(x,y)dxdy \\ &= \text{probability that } X \in (x, x+dx), Y \in (y, y+dy) \end{aligned}$$

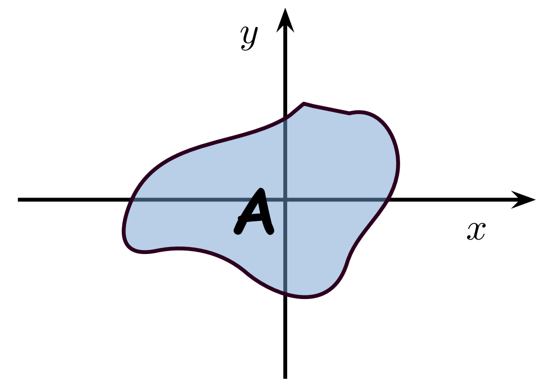


Joint Bivariate Distribution (Continuous)

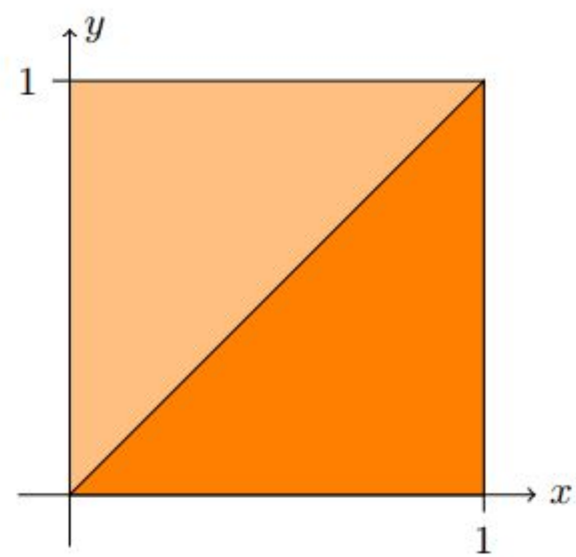
Probability from Joint Distribution

$$P\left[(x,y) \in A\right] = \int\int_A f(x,y)dxdy$$

probability that $(x,y) \in A$ for some subset A of range



Example 4. Suppose X and Y both take values in [0,1] with uniform density $f(x,y) = 1$. Visualize the event '$X > Y$' and find its probability.



The event '$X > Y$' in the unit square.

Marginal Probability Density Function

- Marginal probability density function (pdf) of ***X***

$$p_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

X=x is fixed when y is integrated out.

- Marginal probability density function (pdf) of ***Y***

$$p_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Y=y is fixed when x is integrated out.

Independent Random Variables

- Recall the Independence of probabilities

$$P(A \cap B) = P(A) \cdot P(B)$$

For $\forall x, y \quad p(x,y) = p_X(x) \cdot p_Y(y)$ [discrete]

For $\forall x, y \quad f(x,y) = f_X(x) \cdot f_Y(y)$ [continuous]

Random variables *X* and *Y* are independent.

If the above relation is not satisfied for all x and y, X and Y are NOT dependent.

Independent Random Variables: Example

Example 1. Roll two dice. Let X be the value on the first die and let Y be the value on the second die. Then both X and Y take values 1 to 6 and the joint pmf is $p(i,j) = 1/36$ for all i and j between 1 and 6. Here is the joint probability table:

$X\backslash Y$	1	2	3	4	5	6	$p_X(x)$
1	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
2	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
3	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
4	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
5	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
6	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
$p_Y(y)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	

For all $x, y,\ p(x,y) = p_X(x) \cdot p_Y(y) \rightarrow$ X and Y are independent.

Independent Random Variables: Example

Example 2. Roll two dice. Let X be the value on the first die and let T be the total on both dice. Here is the joint probability table:

$X\backslash T$	2	3	4	5	6	7	8	9	10	11	12	
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0	$1/6$
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	$1/6$
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	$1/6$
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	$1/6$
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	$1/6$
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	$1/6$
	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

For some (x,y), $p(x, y) \neq p_X(x) \cdot p_Y(y) \rightarrow$ X and Y are dependent.

Conditional Distributions

- Recall the conditional probability.

$$P(A|B) = P(A \cap B)/P(B)$$

- For discrete random variables X and Y, when $X = x$, the conditional probability mass function of Y is

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)}$$

- For continuous random variables X and Y, When $X = x$, the conditional probability density function of Y is

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

More than Two Random Variables

- Discrete random variables $X_1, X_2, \cdots, X_n$

$$P(X_1 = x_1, \cdots, X_n = x_n) = p(x_1, \cdots, x_n)$$

- Continuous random variables $X_1, X_2, \cdots, X_n$

$$P\left[(X_1, \cdots, X_n) \in A\right] = \int_A f(x_1, \cdots, x_n) dx_1 \cdots dx_n$$

Expectation Values

Expected values of a function $h(X, Y)$

- $\mu_{h(X,Y)} = E\left[h(X, Y)\right] = \sum_x \sum_y h(x, y) \cdot p(x, y)$ **[discrete]**

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y) dx dy$$ **[continuous]**

- If $h(X, Y) = h(X)$,

$$E[h(X)] = \sum_x \sum_y h(x) p(x, y) = \sum_x h(x) p_X(x)$$ **[discrete]**

$$= \int_{-\infty}^{\infty} h(x) f_X(x) dx$$ **[continuous]**

Covariance

When $h(X,Y) = (X - \mu_X) \cdot (Y - \mu_Y)$

$$\text{Cov}(X,Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$= \sum_x \sum_y (x - \mu_X)(y - \mu_Y) p(x,y)$$ **[discrete]**

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x,y) dx dy$$ **[continuous]**

$$\mu_X = E[X] = \sum_x \sum_y x \cdot p(x,y) = \sum_x x \cdot \left(\sum_y p(x,y) \right) = \sum_x x \cdot p_X(x)$$

$$\mu_Y = E[Y] = \sum_x \sum_y y \cdot p(x,y) = \sum_y y \cdot \left(\sum_x p(x,y) \right) = \sum_x x \cdot p_Y(y)$$

Covariance

When $h(X, Y) = (X - \mu_X) \cdot (Y - \mu_Y)$

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$= \sum_x \sum_y (x - \mu_X)(y - \mu_Y) p(x, y) \quad \text{[discrete]}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy \quad \text{[continuous]}$$

Properties of Covariance

- $\mathrm{Cov}(aX + b, cY + d) = ac\mathrm{Cov}(X, Y)$ for constants a, b, c, d
- $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X) = \sigma_X^2$
- $\mathrm{Cov}(X, Y) = E(XY) - \mu_X \mu_Y$
- $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$
- If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$

  But, Cov(X,Y)=0 does not mean that X and Y are independent.

Properties of Covariance

- If X and Y are independent, then Cov(X,Y) = 0
- In other words, if Cov(X,Y)=0, then X and Y are NOT independent.
- But, Cov(X,Y)=0 does not mean that X and Y are independent.

Properties of Covariance: Example

Tossing a fair coin three times

Sample space

$$S = \{\text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}\}$$

Example 1. Flip a fair coin 3 times. Let X be the number of heads in the first 2 flips and let Y be the number of heads on the last 2 flips (so there is overlap on the middle flip). Compute $\text{Cov}(X,Y)$.

$X\backslash Y$	0	1	2	$p(x_i)$
0	1/8	1/8	0	1/4
1	1/8	2/8	1/8	1/2
2	0	1/8	1/8	1/4
$p(y_j)$	1/4	1/2	1/4	1

$$\mu_X = E(X) = \sum_{x,y} x p(x,y)$$

$$= \sum_x x \left(\sum_y p(x,y) \right) = \sum_x x p_X(x) = 1$$

$$\mu_Y = E(Y) = \sum_{x,y} y p(x,y)$$

$$= \sum_y y \left(\sum_x p(x,y) \right) = \sum_y y p_Y(y) = 1$$

Properties of Covariance: Example

Example 1. Flip a fair coin 3 times. Let X be the number of heads in the first 2 flips and let Y be the number of heads on the last 2 flips (so there is overlap on the middle flip). Compute $\mathrm{Cov}(X,Y)$.

$X\backslash Y$	0	1	2	$p(x_i)$
0	1/8	1/8	0	1/4
1	1/8	2/8	1/8	1/2
2	0	1/8	1/8	1/4
$p(y_j)$	1/4	1/2	1/4	1

$$\mathrm{Cov}(X,Y) = E\left((X-\mu_X)(Y-\mu_Y)\right)$$
$$= \sum_{x,y} p(x,y)(x-1)(y-1) = \frac{1}{4}$$

$$\mathrm{Cov}(X,Y) = E(XY) - \mu_X\mu_Y$$
$$= 1\cdot 1\cdot \frac{2}{8} + 1\cdot 2\cdot \frac{1}{8} + 2\cdot 1\cdot \frac{1}{8} + 2\cdot 2\cdot \frac{1}{8} - 1$$
$$= \frac{5}{4} - 1 = \frac{1}{4}$$

Properties of Covariance: Example

If Cov(X,Y)=0, X and Y are independent?

Example 2. (Zero covariance does not imply independence.) Let X be a random variable that takes values $-2, -1, 0, 1, 2$; each with probability $1/5$. Let $Y = X^2$. Show that $\text{Cov}(X, Y) = 0$ but X and Y are not independent.

$Y \backslash X$	-2	-1	0	1	2	$p(y_j)$
0	0	0	$1/5$	0	0	$1/5$
1	0	$1/5$	0	$1/5$	0	$2/5$
4	$1/5$	0	0	0	$1/5$	$2/5$
$p(x_i)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$	1

Properties of Covariance: Example

If Cov(X,Y)=0, X and Y are independent?

Example 2. (Zero covariance does not imply independence.) Let X be a random variable that takes values $-2, -1, 0, 1, 2$; each with probability $1/5$. Let $Y = X^2$. Show that $\text{Cov}(X, Y) = 0$ but X and Y are not independent.

$Y \backslash X$	-2	-1	0	1	2	$p(y_j)$
0	0	0	1/5	0	0	1/5
1	0	1/5	0	1/5	0	2/5
4	1/5	0	0	0	1/5	2/5
$p(x_i)$	1/5	1/5	1/5	1/5	1/5	1

$$E(X) = 0, \quad E(Y) = 2$$

$$\text{Cov}(X, Y) = \frac{1}{5}(-8 - 1 + 1 + 8) - \mu_X \mu_y = 0.$$

$$p(x = -2, y = 0) = 0 \neq p_X(x = -2) \cdot p_Y(y = 0) = \frac{1}{25}$$

X and Y are NOT independent

Correlation

The correlation coefficient of X and Y

$$\mathrm{Corr}(X, Y) = \rho_{X,Y} = \rho = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

- The covariance is not dimensionless.
- The correlation is dimensionless.

$\rightarrow$ If units of X and Y change, their covariance also changes.

$\rightarrow$ But, their correlation does not change.

$\rightarrow$ The correlation is a standardized covariance.

Correlation

- For constants $a(\neq 0), b, c(\neq 0), d$

$$\mathrm{Corr}(aX+b, cY+d) = \frac{a}{|a|}\frac{c}{|c|}\mathrm{Corr}(X,Y)$$

- If X and Y are independent, then $\mathrm{Corr}(X,Y) = 0$

  But, $\mathrm{Corr}(X,Y) = 0$ does not mean the independence of X and Y.

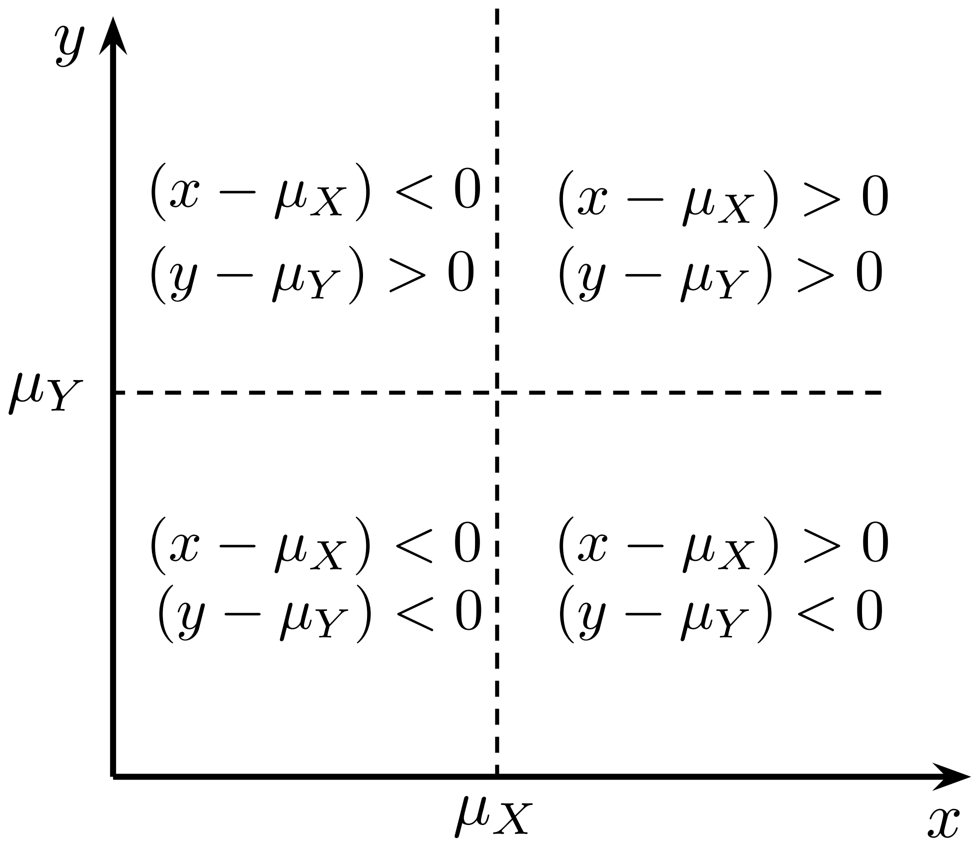
- $-1 \leq \mathrm{Corr}(X,Y) \leq 1$

$$\mathrm{Corr}(X,Y) = +1 \text{ if and only if } Y = aX + b \text{ with } a > 0$$

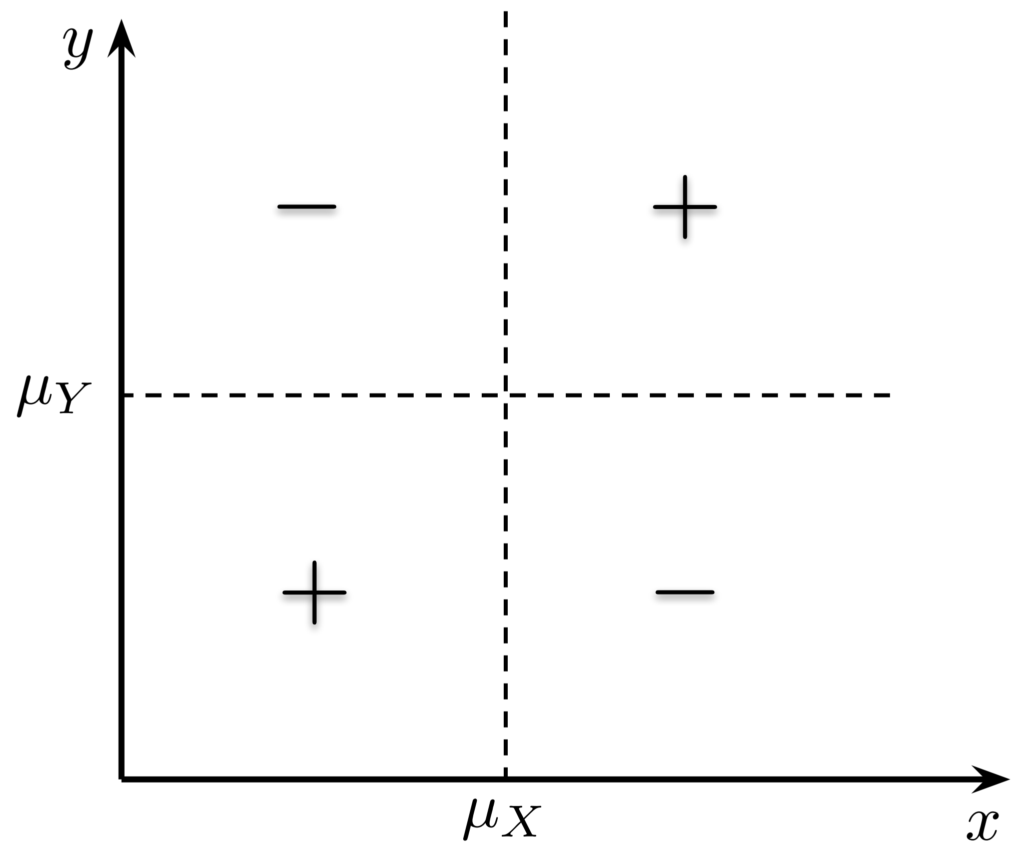
$$\mathrm{Corr}(X,Y) = -1 \text{ if and only if } Y = aX + b \text{ with } a < 0$$

Covariance and Correlation

A measure that implies whether X and Y have a strong relationship

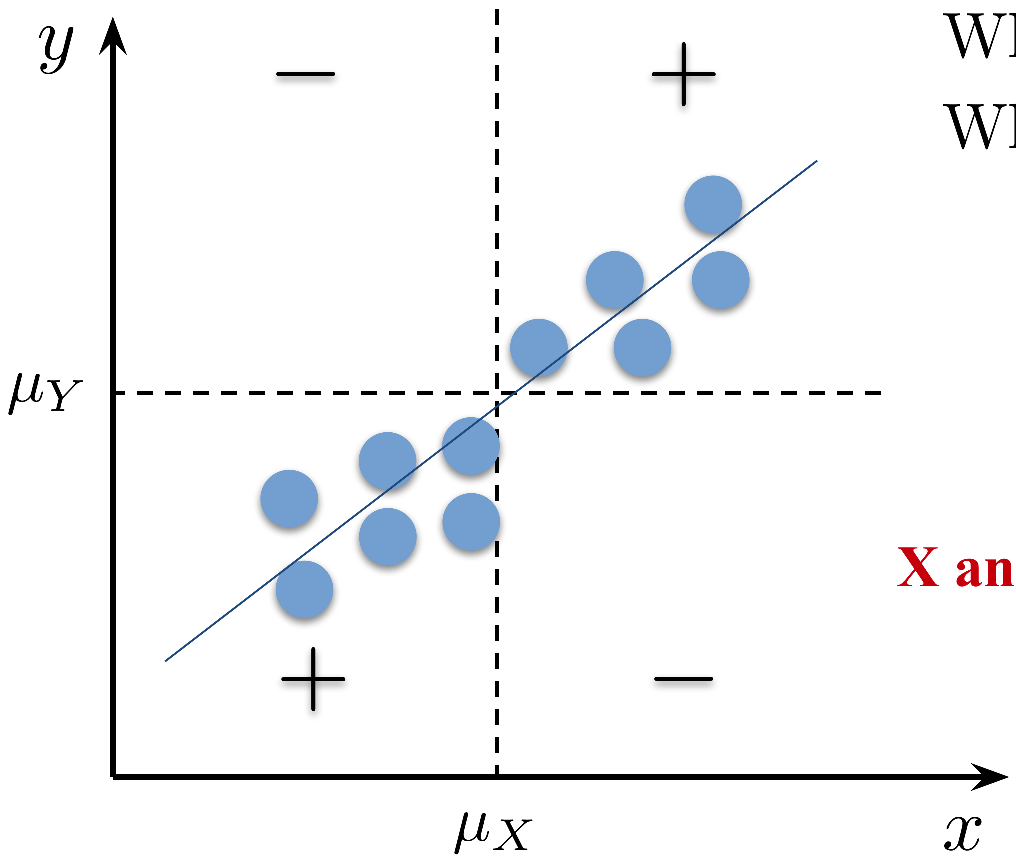


sign of $(x - \mu_X)(y - \mu_Y)$



Covariance and Correlation

A measure that implies whether X and Y have a strong relationship



$$\text{When } (x - \mu_X) > 0, (y - \mu_Y) > 0$$

$$\text{When } (x - \mu_X) < 0, (y - \mu_Y) < 0$$

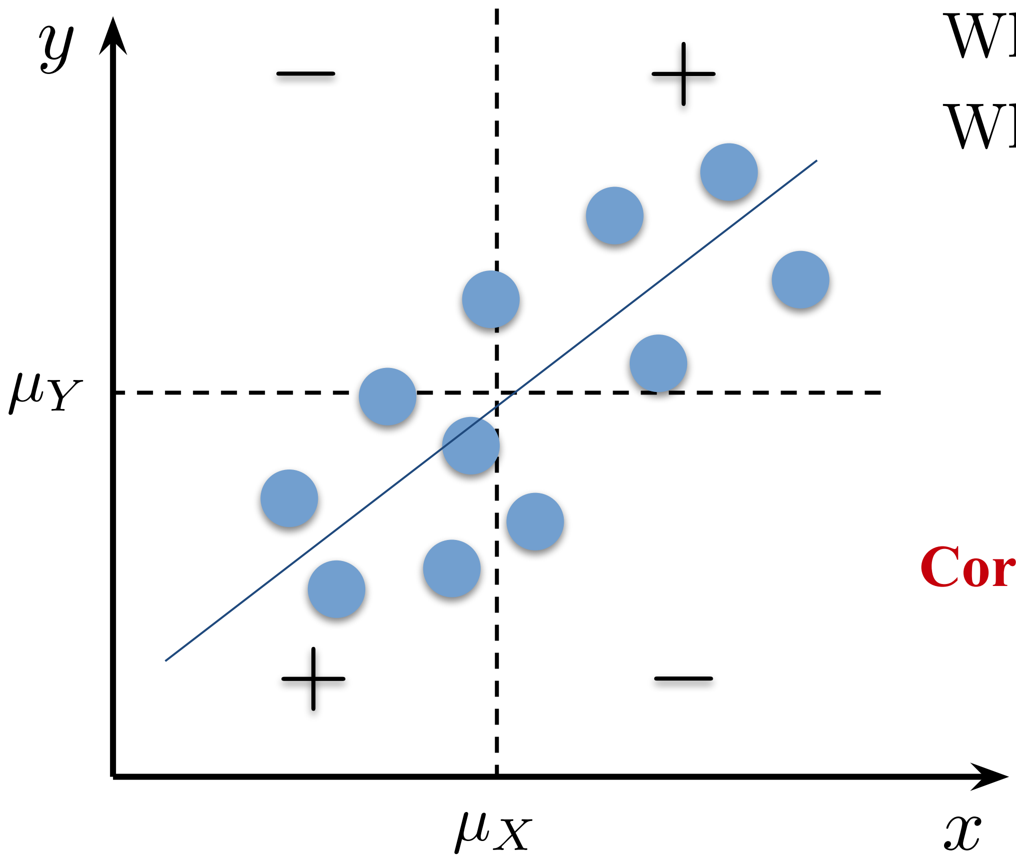
$$(x - \mu_X)(y - \mu_Y) > 0$$

$$\mathrm{Cov}(X, Y) > 0$$

X and Y have a strong positive relationship.

Covariance and Correlation

A measure that implies whether X and Y have a strong relationship



$$\text{When } (x - \mu_X) > 0, (y - \mu_Y) > 0$$

$$\text{When } (x - \mu_X) < 0, (y - \mu_Y) < 0$$

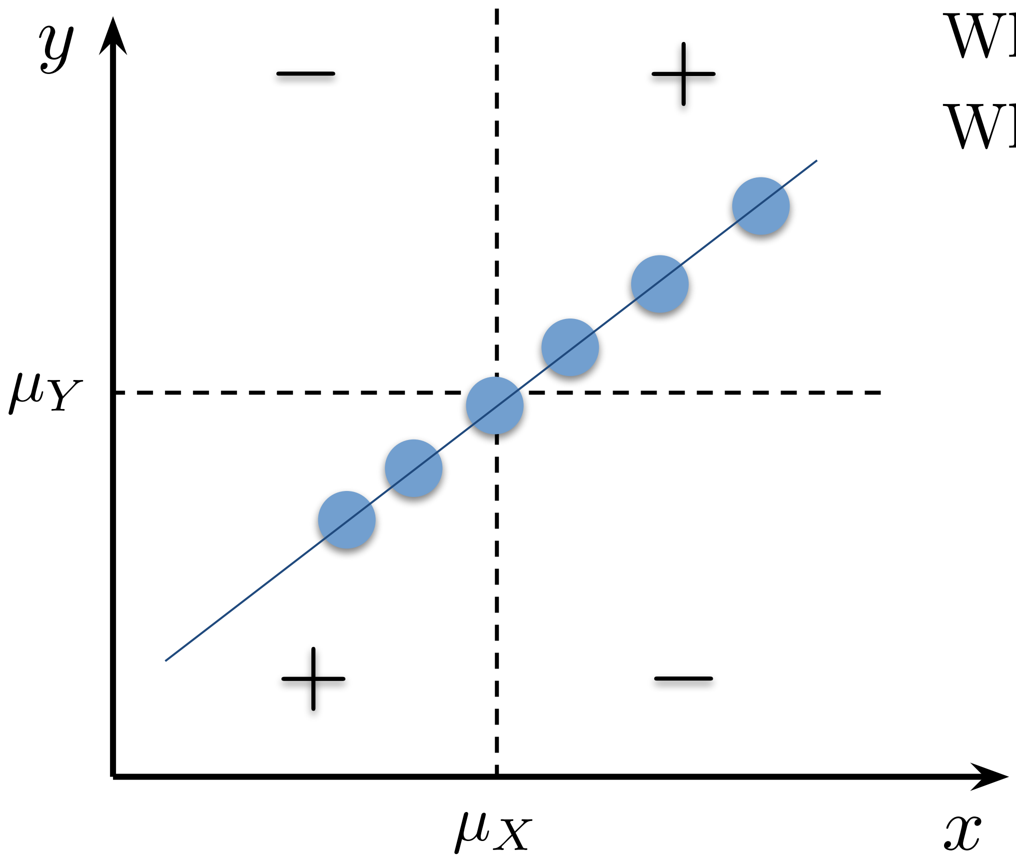
$$(x - \mu_X)(y - \mu_Y) > 0$$

$$\mathrm{Cov}(X, Y) > 0$$

Correlation might reduce, even though its sign is still positive.

Positive Unit Correlation

X and Y have a perfect linear relationship.



When $(x - \mu_X) > 0, (y - \mu_Y) > 0$

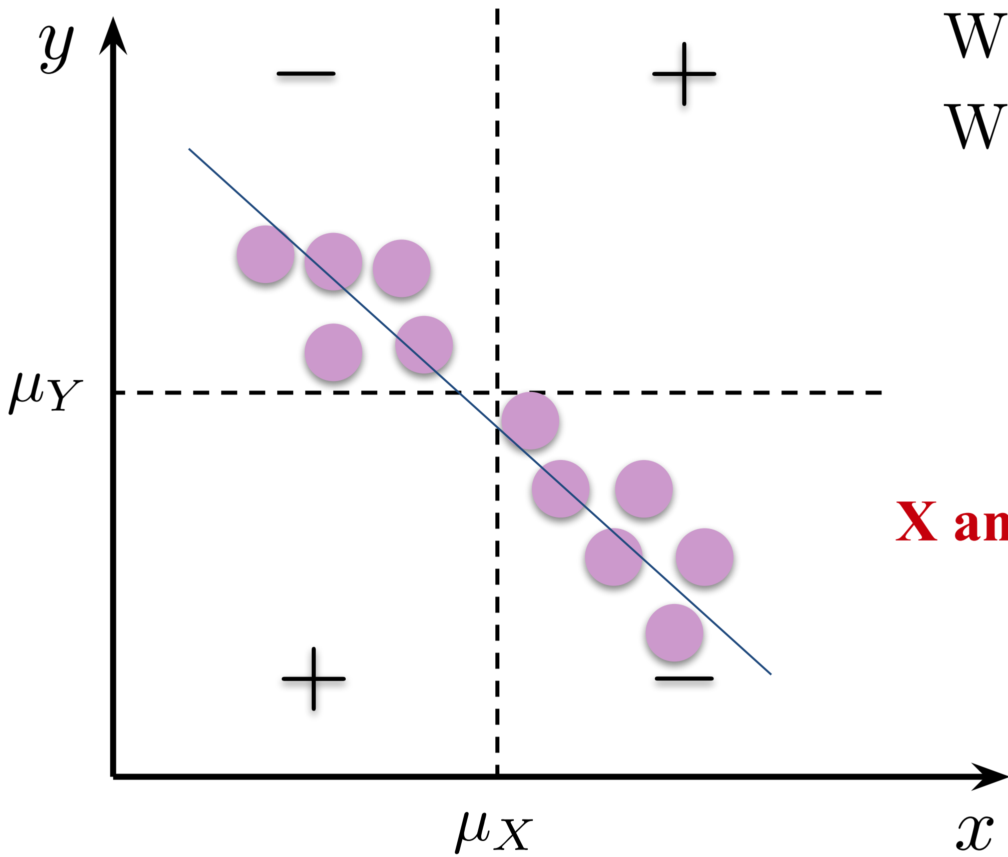
When $(x - \mu_X) < 0, (y - \mu_Y) < 0$

$$(x - \mu_X)(y - \mu_Y) > 0$$

$$\mathrm{Cov}(X, Y) > 0$$

Covariance and Correlation

A measure that implies whether X and Y have a strong relationship



$$\text{When } (x-\mu_X) > 0, (y-\mu_Y) < 0$$
$$\text{When } (x-\mu_X) < 0, (y-\mu_Y) > 0$$

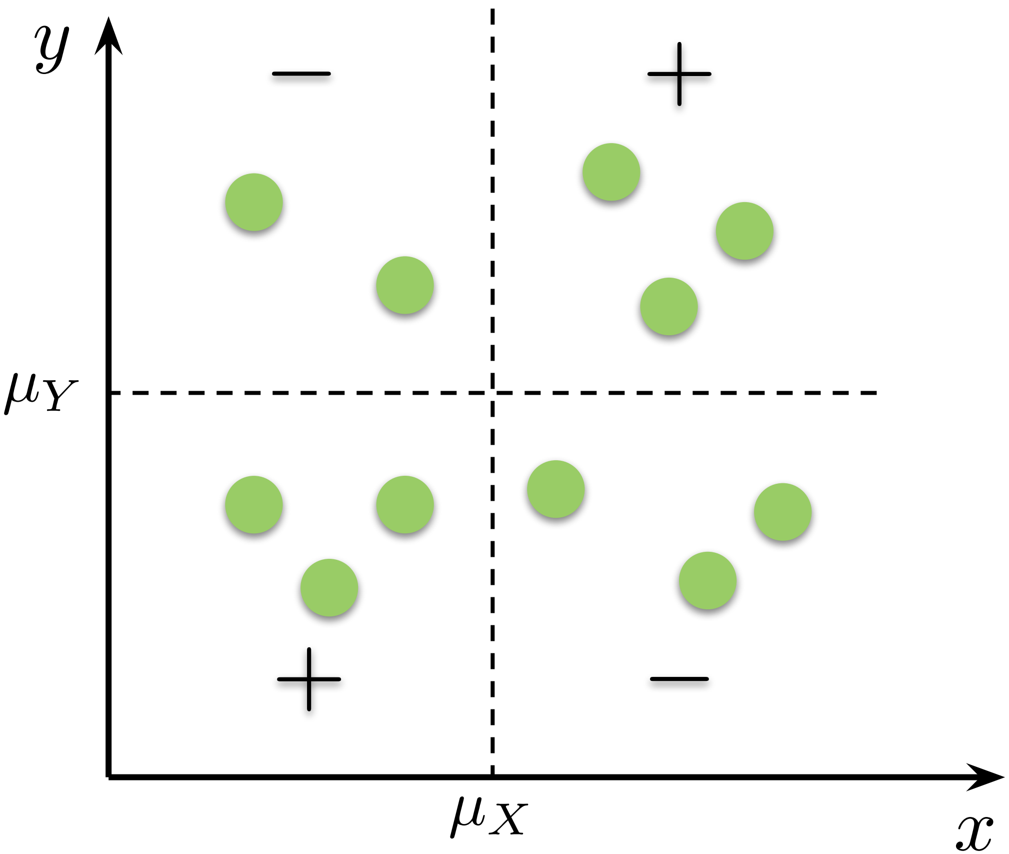
$$(x-\mu_X)(y-\mu_Y) < 0$$

$$\mathrm{Cov}(X,Y) < 0$$

X and Y have a strong negative relationship.

Covariance and Correlation

A measure that implies whether X and Y have a strong relationship



positive $(x - \mu_X)(y - \mu_Y)$
negative $(x - \mu_X)(y - \mu_Y)$
almost cancel out each other.

$$\mathrm{Cov}(X, Y) \approx 0$$

X and Y are not strongly related.

Covariance and Correlation

A measure that implies whether X and Y have a strong relationship

- $-1 \leq \mathrm{Corr}(X, Y) \leq 1$

$$\mathrm{Corr}(X,Y) = +1 \text{ if and only if } Y = aX + b \text{ with } a > 0$$

$$\mathrm{Corr}(X,Y) = -1 \text{ if and only if } Y = aX + b \text{ with } a < 0$$

- Maximum(1) and Minimum(-1) of correlation
  = perfect linear relationship between X and Y

- $|\rho| < 1$ implies that there is some relationship between X and Y but, not completely linear.
  Possibly there may be a strong nonlinear relationship.