Covariance

What’s Important?

# variables		
1	Expectation	Average Value
	Variance	Distance^2 from mean
2	Relationship	

Relationship between X and Y

As X ↗ What happens to Y ? ↗ ? ↘ ?

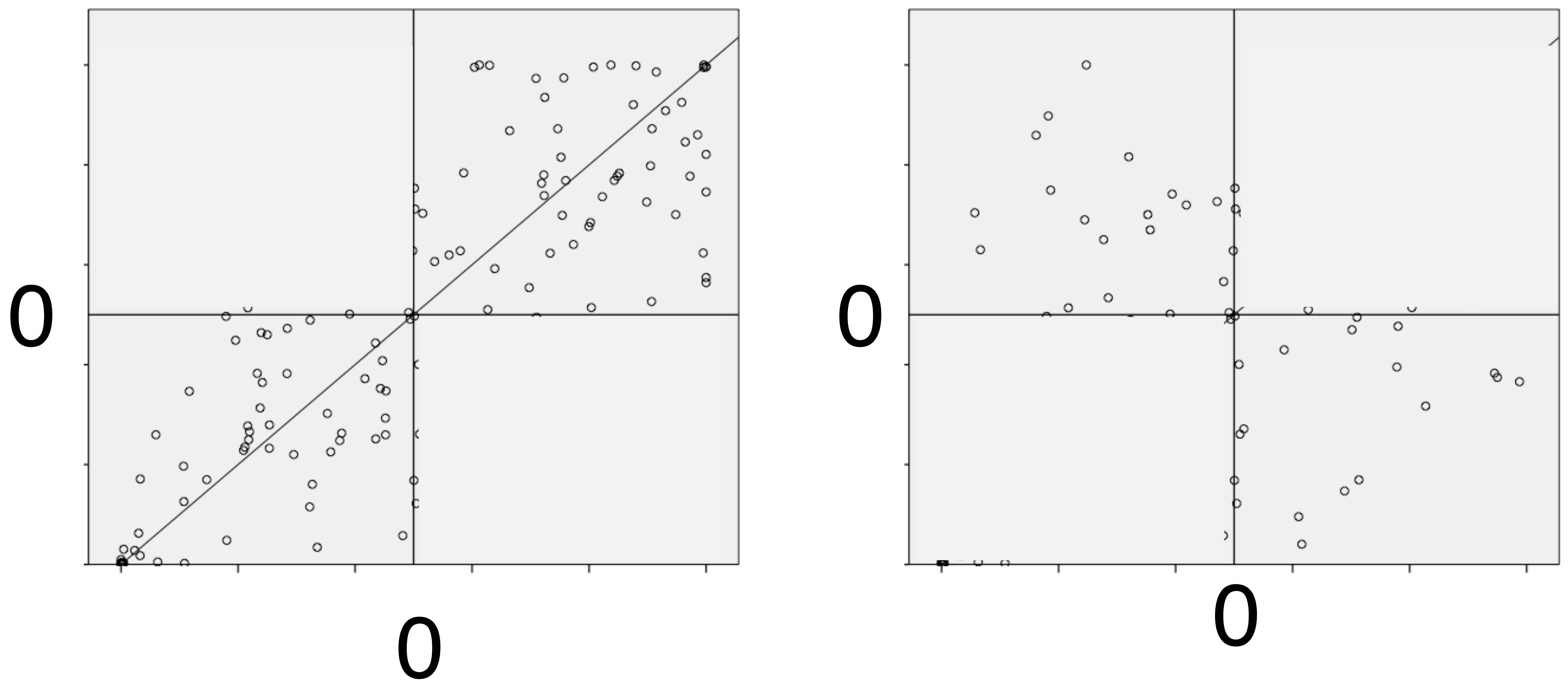
Height vs. weight ↗

Student

Absences vs. grade ↘

Simplification

X ≥ 0 Do we expect $Y \gtrless 0$?

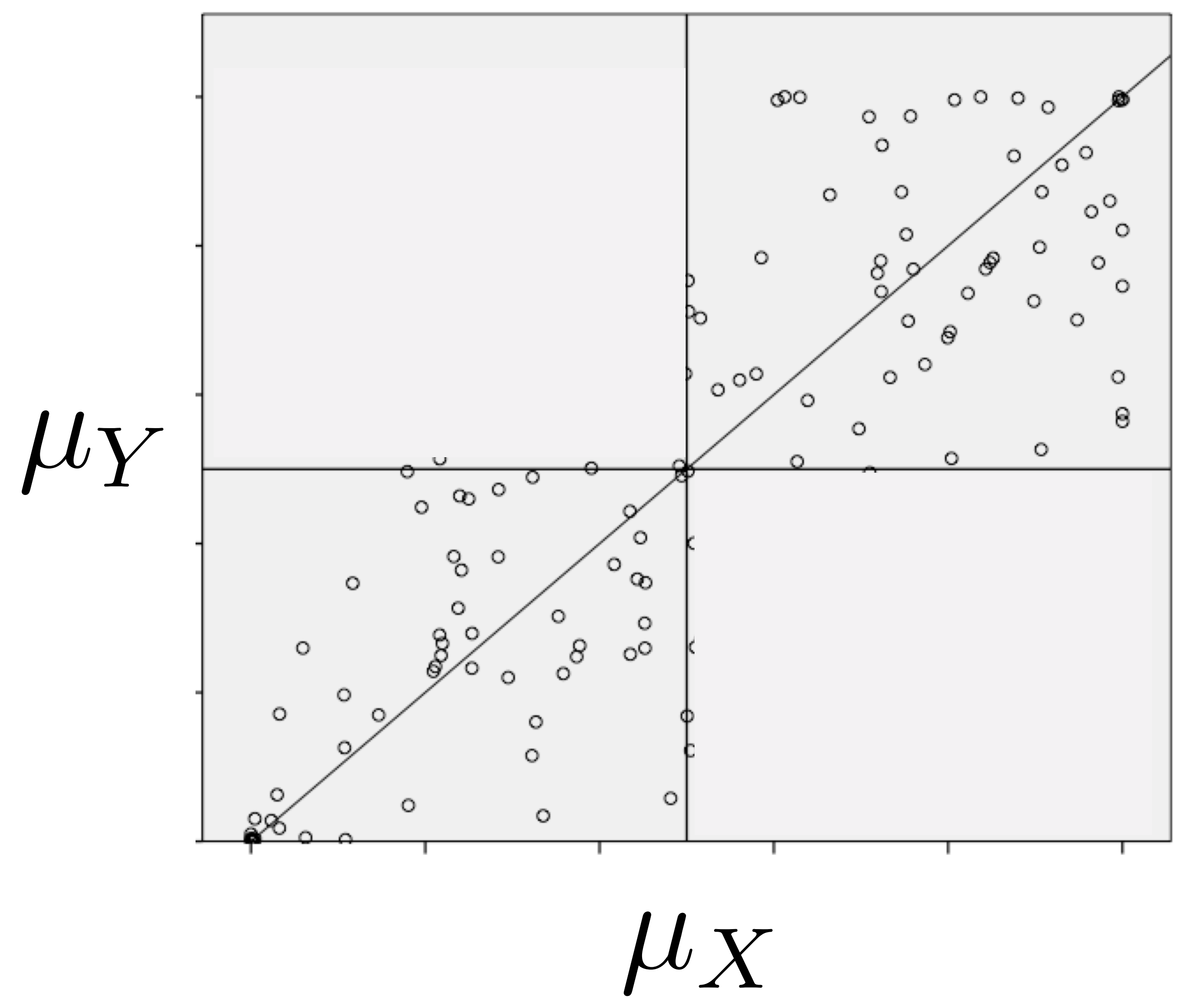


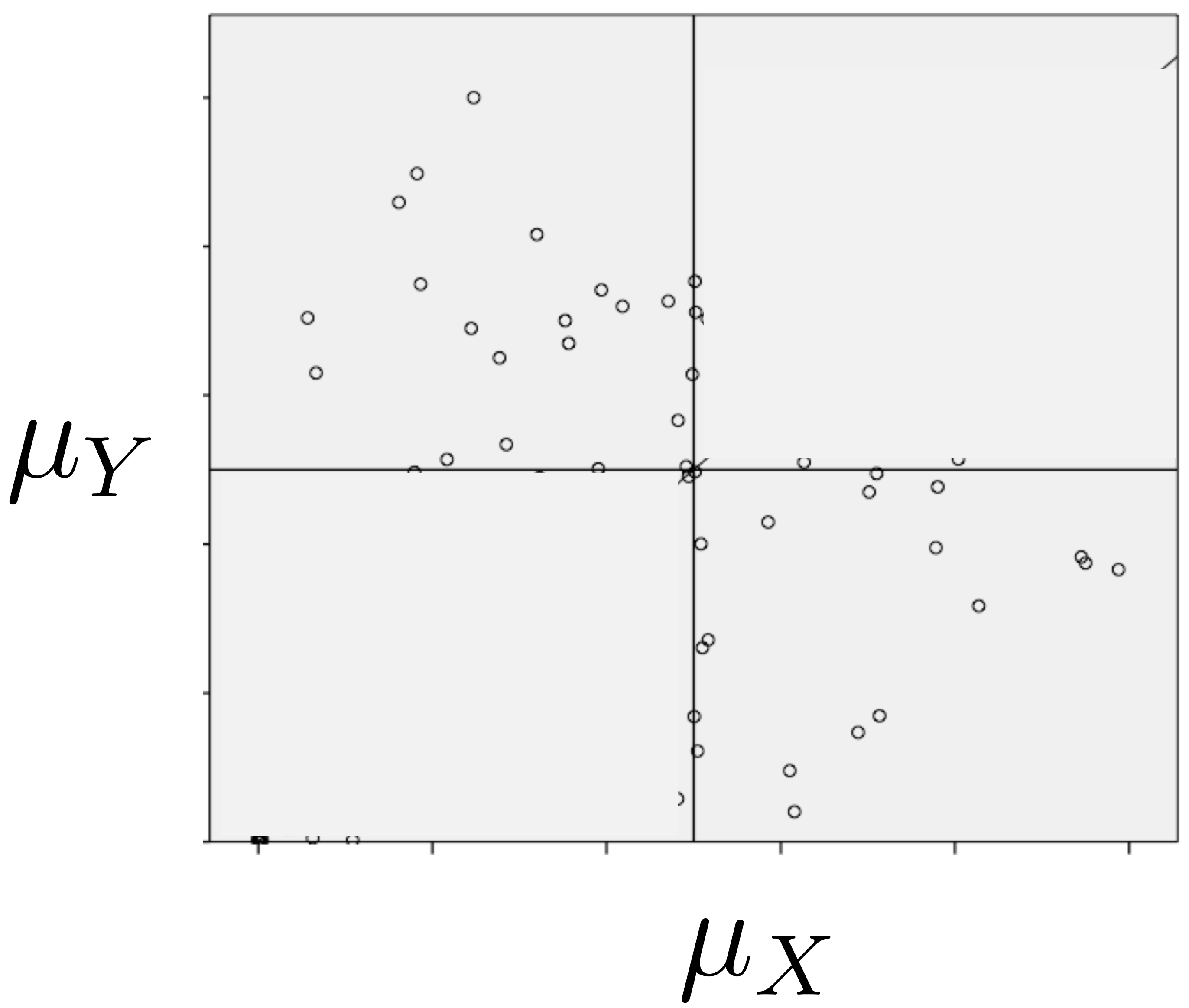
Y ≥ 0 Expect Y ≤ 0

E(XY) ≥ 0 E(XY) ≤ 0

More Relevant

Suppose $X \geq \mu_X$ Do we expect $Y \gtrless \mu_Y$?





$$E[(X - \mu_X)(Y - \mu_Y)] \geq 0$$

$$E[(X - \mu_X)(Y - \mu_Y)] \leq 0$$

Expect $Y \geq \mu_Y$

Expect $Y \leq \mu_Y$

Covariance

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

Amount X and Y vary together

Example

Add

Alternative Formulation

Recall Variance

$$V(X) = E[(X - \mu_X)^2] \;\; = E(X^2) - \mu_X^2$$

Expectation of square — square of expectation

Similar for Covariance

2nd Covariance Formulation

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

$$= E[XY - X\mu_Y - \mu_X Y + \mu_X\mu_Y]$$

$$= E(XY) - E(X\mu_Y) - E(\mu_X Y) + E(\mu_X\mu_Y)$$

$$= E(XY) - E(X)\mu_Y - \mu_X E(Y) + \mu_X\mu_Y$$

$$= E(XY) - \mu_X\mu_Y$$

Expectation of product — product of expectation

Properties

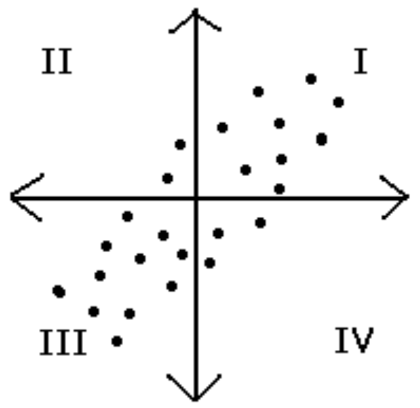
$$\mathrm{Cov}(X,X) = EX^2 - \mu_X^2 = V(X)$$

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

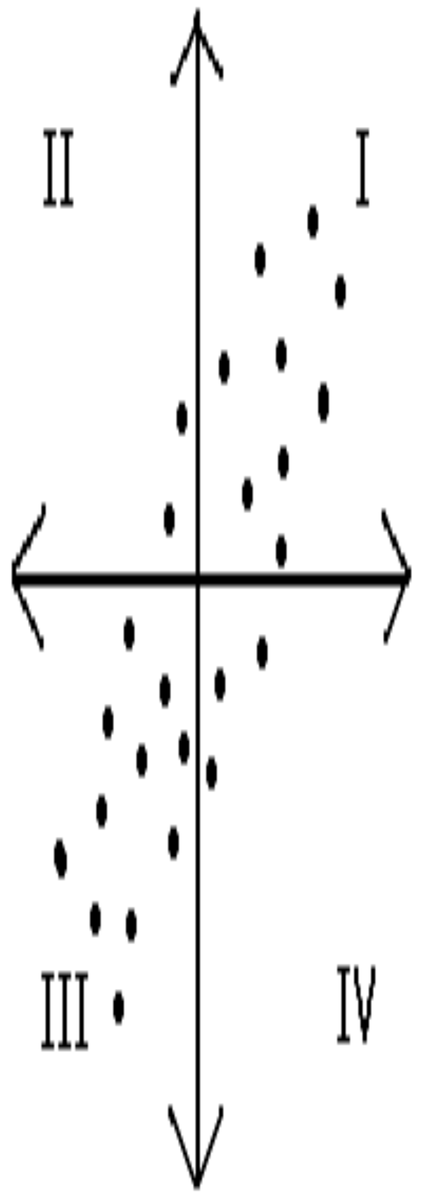
$$\mathrm{Cov}(aX,Y) = E(aXY) - \mu_{aX}\mu_Y = aE(XY) - a\mu_X\mu_Y = a\mathrm{Cov}(X,Y)$$

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

Size

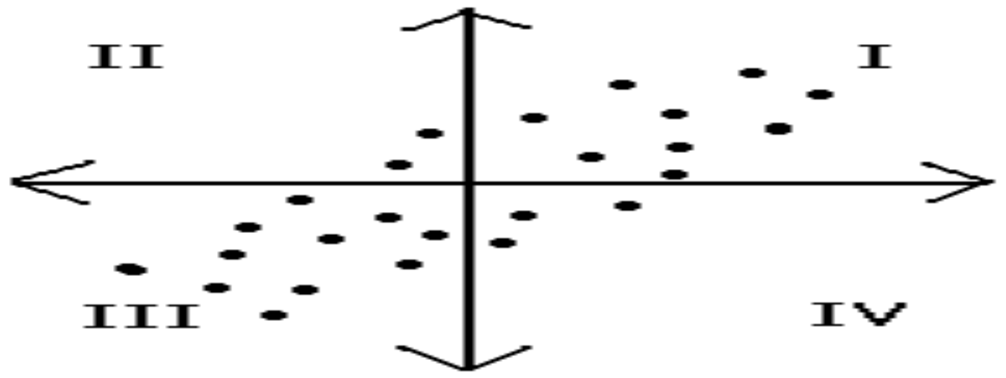


Small



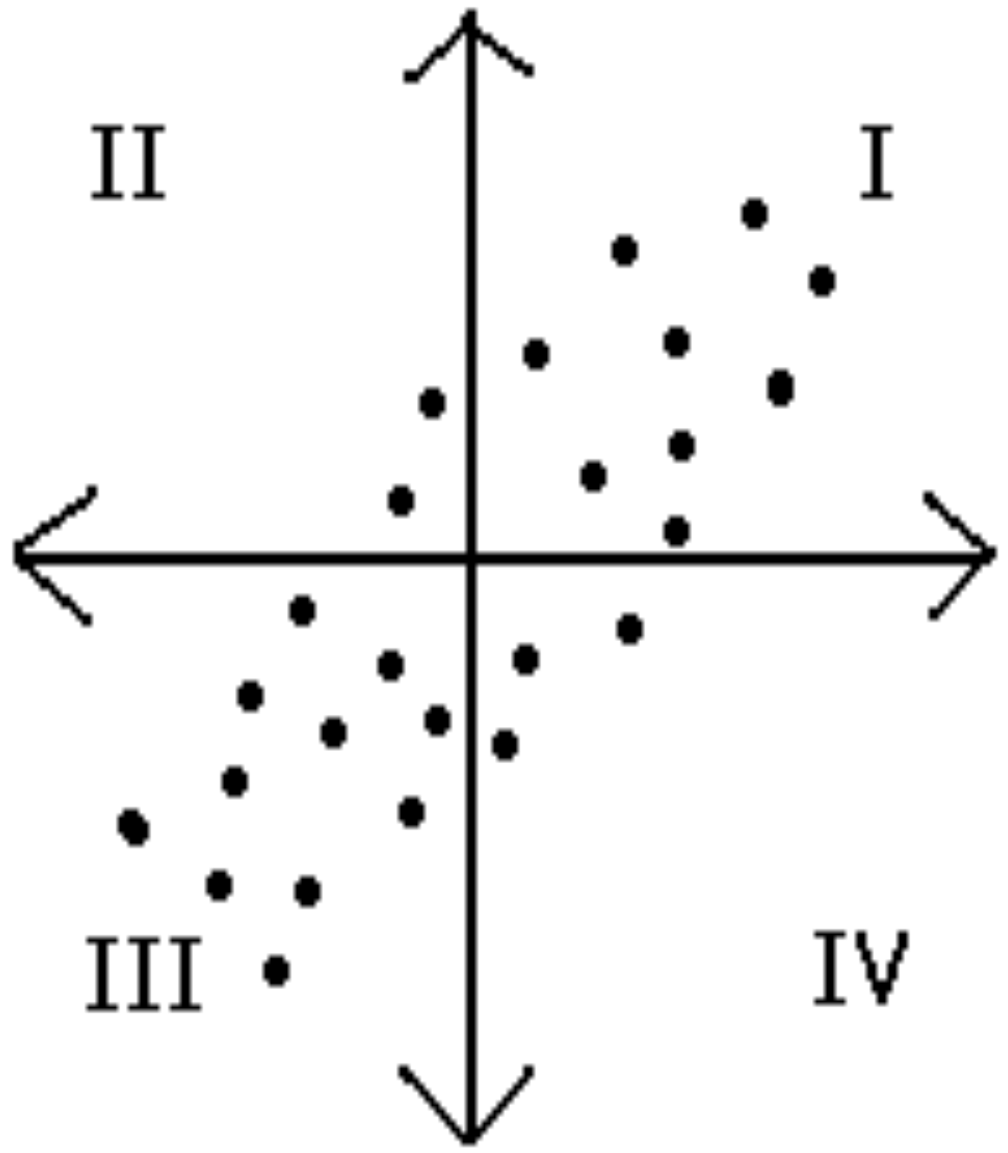
Larger

x a



Larger

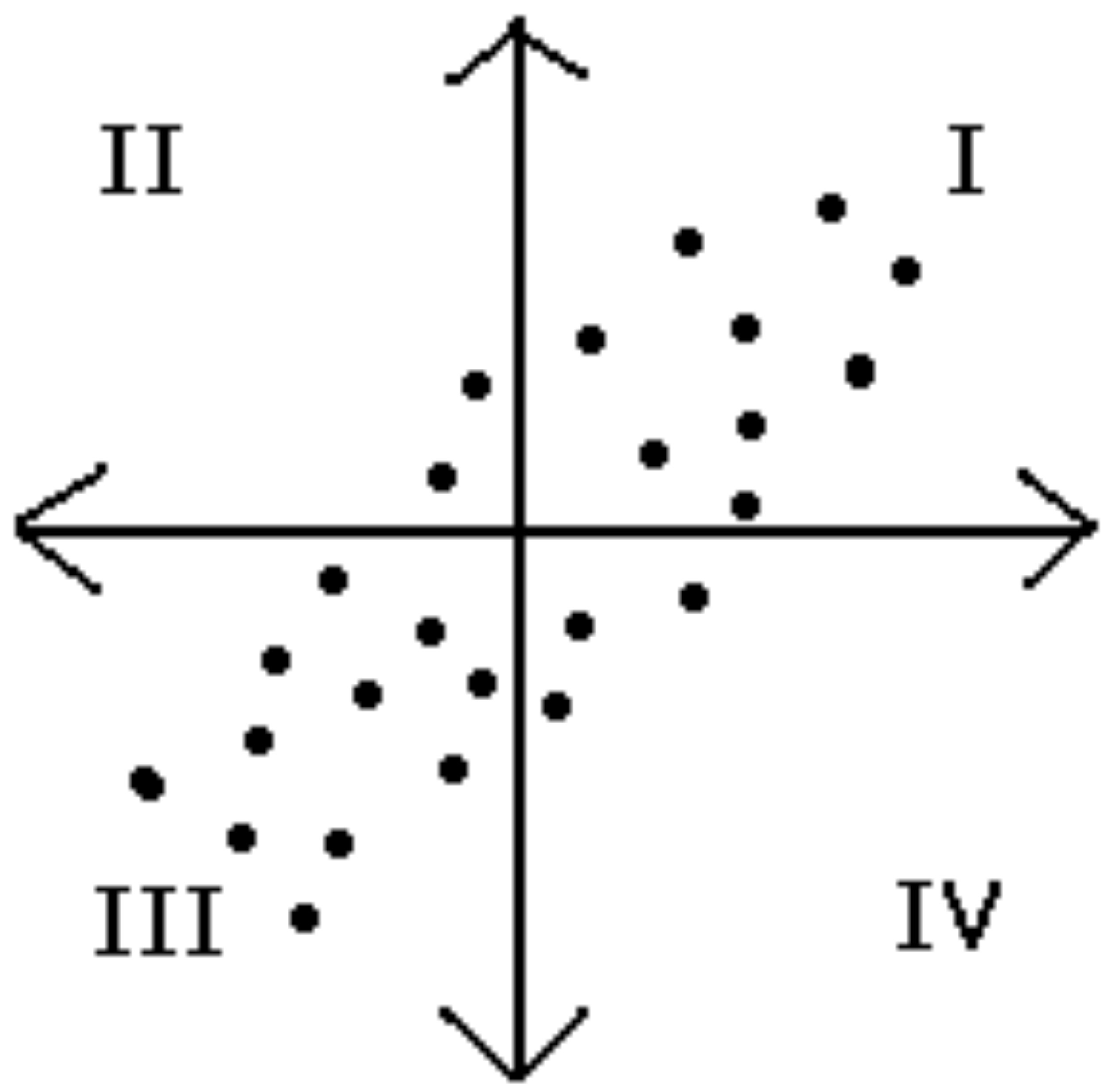
x a



Largest

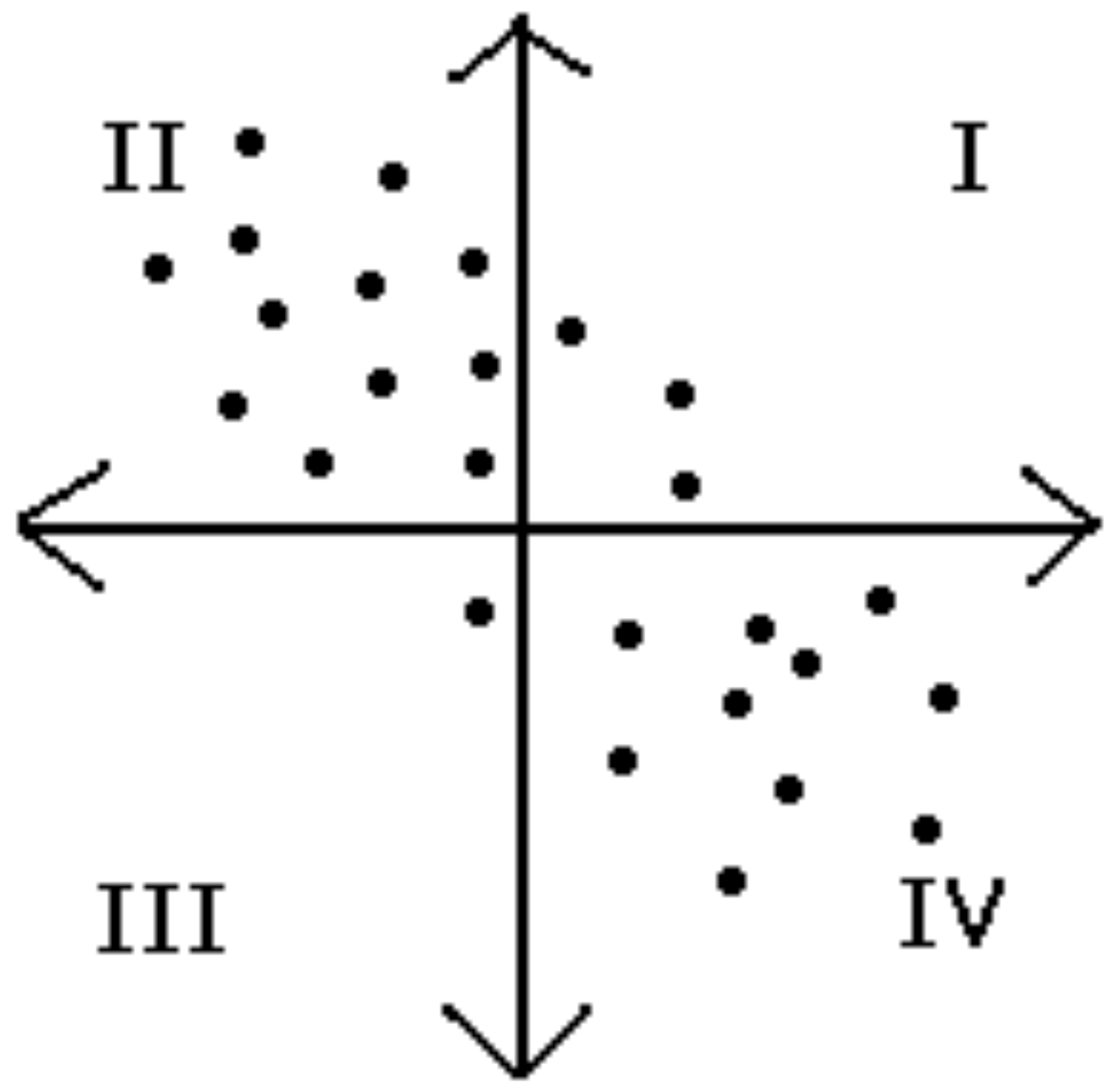
x a^2

Correlation



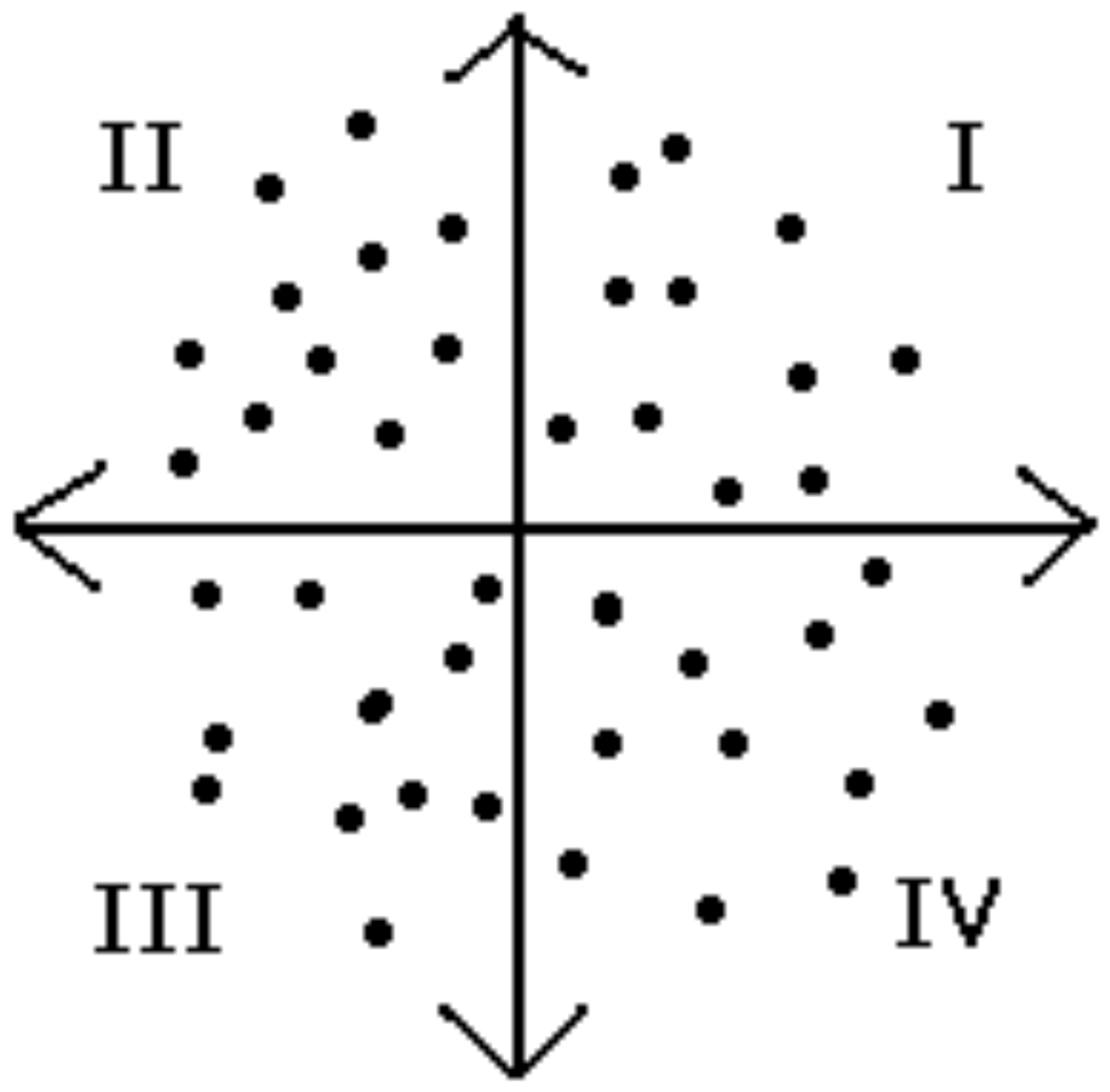
$\mathrm{Cov}(X,Y) > 0$

Positively correlated



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

Negatively correlated



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

Uncorrelated

Correlation Examples

$X, Y \sim B(\frac{1}{2})$

Correlation			
Positive	$X, X+Y$	$X, 2X-Y$	$\min(X,Y), \max(X,Y)$
Uncorrelated	X, Y	$3X, 4X+1$	$X+Y, X-Y$
Negative	$X, Y-X$	$X-2Y, Y-2X$	$\lvert X-Y \rvert, \min(X,Y)$

Interpretation

Very roughly speaking

X, Y, linearly dependent   Y = aX + b

Observed value of	Increases by
X	1
Y	$\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}$

$$X = x \qquad Y = ax + b$$

$$X = x + 1 \qquad Y = a(x+1) + b = ax + b + a$$

+ 1 $+a$

$$\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} = \frac{a\mathrm{Var}(X)}{\mathrm{Var}(X)} = a$$

Back to Variance

$$V(X+Y) \stackrel{?}{=} V(X) + V(Y)$$

$$\begin{aligned}
V(X+Y) &= E(X+Y)^2 - (E(X+Y))^2 \\
&= E(X^2 + 2XY + Y^2) - (EX + EY)^2 \\
&= EX^2 + 2E(XY) + EY^2 - (E^2X + 2EX \cdot EY + E^2Y) \\
&= EX^2 - \; E^2X \; + EY^2 - \; E^2Y \; + 2(E(XY) - EX \cdot EY) \\
&= V(X) + V(Y) + 2(E(XY) - EX \cdot EY) \\
&= V(X) + V(Y) + 2\mathrm{Cov}(X,Y)
\end{aligned}$$

$$= \text{iff } \mathrm{Cov}(X,Y) = 0 \quad \text{Uncorrelated}$$

Cauchy-Schwarz Inequality

$E(X \cdot Y)$ can't take all possible values

$$|E(XY)| \leq \sqrt{EX^2} \cdot \sqrt{EY^2}$$

For any α

$$0 \leq E(\alpha X + Y)^2 = \alpha^2 EX^2 + 2\alpha E(XY) + EY^2$$

True for all α, so discriminant must be negative

$$4(EXY)^2 - 4EX^2 \cdot EY^2 \leq 0$$

$$(EXY)^2 \leq EX^2 \cdot EY^2$$

$$V(X+Y) \text{ may be } \quad \geq \quad = \quad \text{or} \quad \leq \quad V(X)+V(Y)$$

$$X \perp Y \ \underset{\leftrightarrow}{\rightarrow} \ \sigma_{X,Y} = 0 \ \underset{\leftrightarrow}{\rightarrow} \ V(X+Y) = V(X)+V(Y)$$

$$X \perp\!\!\!\perp Y \rightarrow X \perp Y \rightarrow V(X+Y) = V(X)+V(Y)$$

$$X, Y \sim B(\tfrac{1}{2})$$

$$Y = X \qquad\qquad \sigma_{X,Y} = V(X)$$

$$V(X+Y) = V(2X) = 4V(X) = V(X)+V(Y)+2V(X)$$

$$Y = -X$$

$$V(X+Y) = V(0) - 0 = V(X)+V(X) - 2V(X)$$

Correlation Coefficient

Dependence on Units

$\text{Cov}(aX, bY) = ab\,\text{Cov}(X,Y)$

Depends on units

Height & Weight France (meters, kg)

US (feet, lb)

French More covariance than US

Relate quantities regardless of units?

(Pearson's) Correlation Coefficient

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Pearson's correlation coefficient

Properties:

$$\rho_{X,X} = 1 \qquad \rho_{X,-X} = -1$$

$$\rho_{X,Y} = \rho_{Y,X}$$

$$\rho_{aX+b,cY+d} = \text{sign}(ac) \cdot \rho_{X,Y}$$

$$\text{sign}(x) = \begin{cases} 1 & x > 0 \\ 0 & 0 \\ -1 & x < 0 \end{cases}$$

If X $\nearrow$ by σ_X, by how many σ_Y do we expect Y to $\nearrow$

Bounds on $\rho_{X,Y}$?

Correlation Coefficient

$$|E(X-\mu_X)(Y-\mu_Y)| \leq \sqrt{E(X-\mu_X)^2 \cdot E(Y-\mu_Y)^2}$$

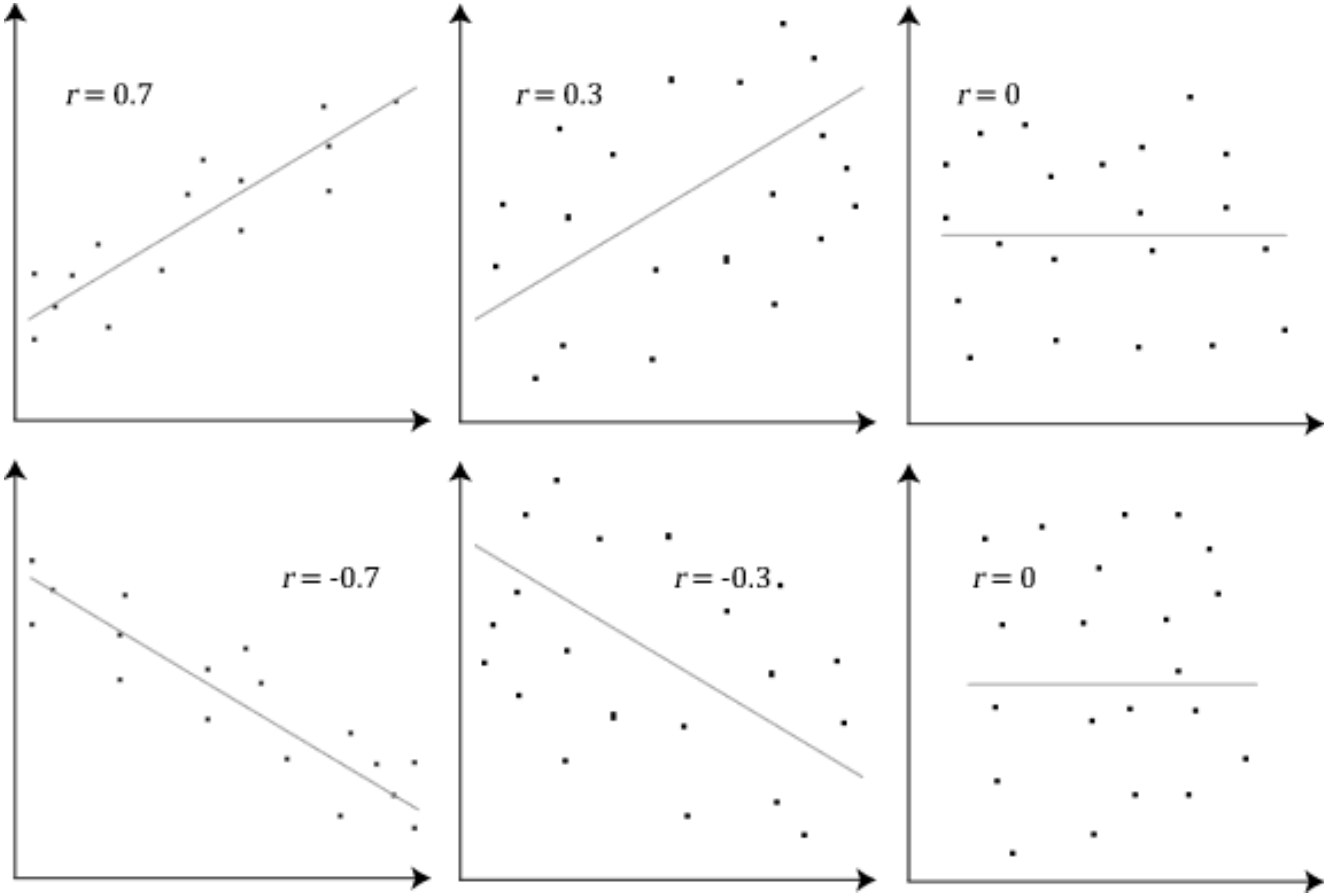
Namely

$$|\sigma_{X,Y}| \leq \sigma_X \cdot \sigma_Y$$

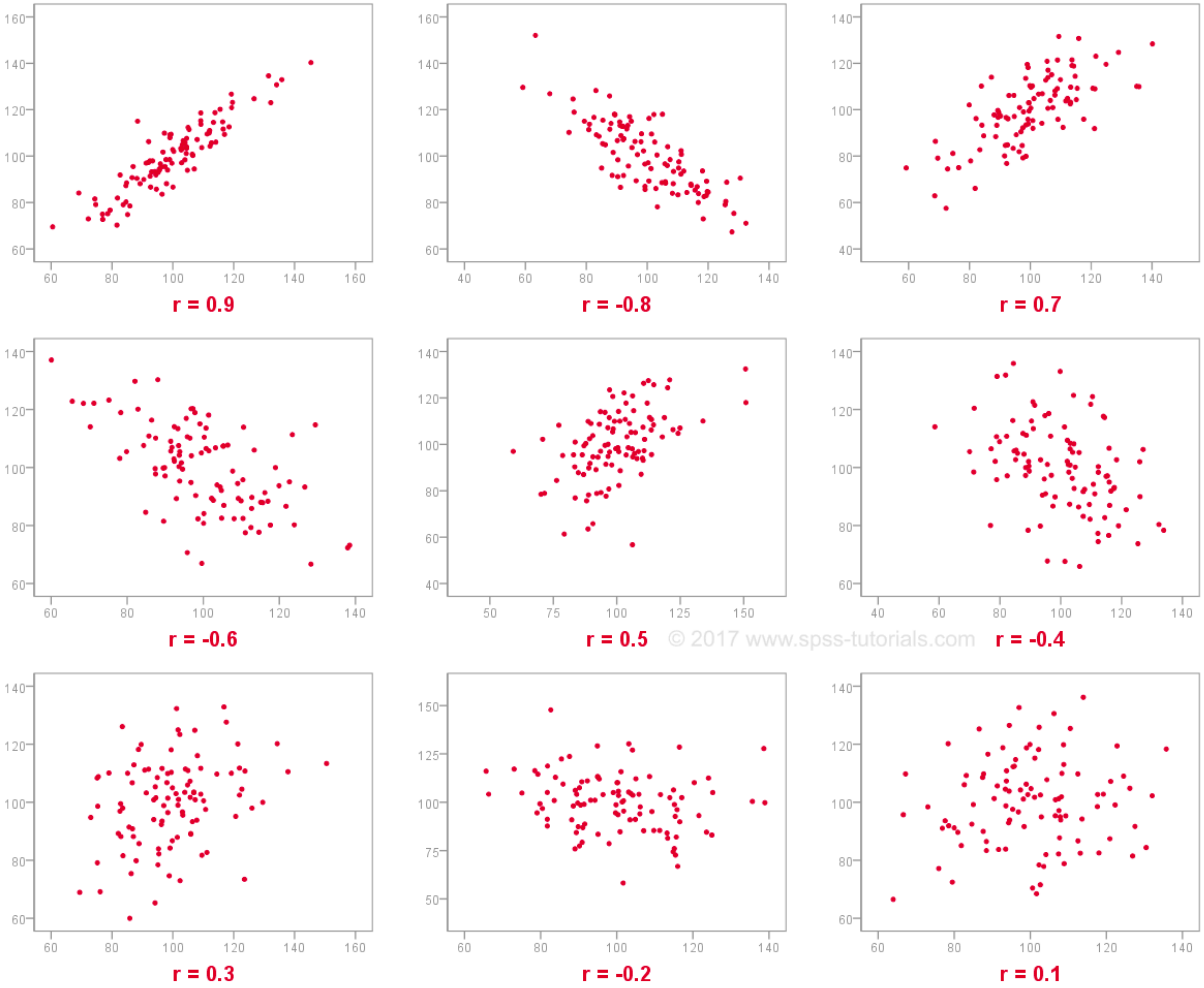
$$\rho_{X,Y} \triangleq \frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y}$$

$$|\rho_{X,Y}| \leq 1$$

Visualization



PEARSON CORRELATION (r) VISUALIZED AS SCATTERPLOT



Covariance v. Correlation Coeff.

Same	$\geq$ if increase together	$\leq$ if decrease together
	Higher absolute value if relation close to linear	
Different	Covariance	Changes with size of variables
	Correlation	Insensitive to variable size

$$\perp\!\!\!\perp \rightarrow \perp$$

Independent implies uncorrelated

$$\begin{aligned}
E(XY) &= \sum_x \sum_y xy \cdot p(x,y) \\
&= \sum_x \sum_y xy \cdot p(x)p(y) \\
&= \sum_x x \cdot p(x) \sum_y y \cdot p(y) \\
&= E(X) \cdot E(Y)
\end{aligned}$$

$$\perp \not\leftrightarrow \perp\!\!\!\perp$$

Independent $\rightarrow$ uncorrelated

Uncorrelated $\xrightarrow{?}$ independent

$$X = \begin{cases} -1 & \frac{1}{2} \\ +1 & \frac{1}{2} \end{cases}$$

$$X = -1 \rightarrow \ Y = 0$$

$$X = +1 \rightarrow \ Y = \begin{cases} +1 \\ -1 \end{cases}$$

X \ Y	-1	0	1
-1		½	
1	¼		¼

Uncorrelated

$$EX = 0 \qquad EY = 0$$

$$E(XY) = \frac{1}{4} \cdot -1 + \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 1 = 0 = EX \cdot EY$$

Clearly dependent

Note: Uncorrelated binary random pairs are independent

Do Expectations Multiply?

$$E(XY) = \sum_{x,y} xy \cdot p(x,y)$$

$$E(XY) \stackrel{?}{=} EX \cdot EY$$

$$X = Y = \begin{cases} -1 & \frac{1}{2} \\ 1 & \frac{1}{2} \end{cases}$$

	-1	1	y
-1	½	0	½
1	0	½	½
x	½	½	

$$EX = EY = 0$$

$$EX \cdot EY = 0$$

$$E(XY) = EX^2 = E(1) = 1$$

$$E(XY) \neq EX \cdot EY$$

Expectations do not always multiply!

Satisfy any relation?

Wild World of Product Expectations

For any α, β, γ ∃ X, Y with: $EX = \alpha$ $EY = \beta$ $E(XY) = \gamma$

$$Y' = X' = \begin{cases} -1 & \frac{1}{2} \\ +1 & \frac{1}{2} \end{cases}$$

$$EX' = EY' = 0 \qquad E(X'Y') = E[(X')^2] = 1$$

$$X = (\gamma - \alpha\beta)X' + \alpha \qquad Y = Y' + \beta$$

$$EX = \alpha \qquad EY = \beta$$

$$\begin{aligned} E(XY) &= E((\gamma - \alpha\beta)X' + \alpha)(Y' + \beta) \\ &= (\gamma - \alpha\beta)\underbrace{E(X'Y')}_{1} + \alpha \underbrace{EY'}_{0} + (\gamma - \alpha\beta)\beta \underbrace{EX'}_{0} + \alpha\beta \\ &= \gamma \end{aligned}$$

Can we still say something about E(XY)?