Unit #4: Expectation, Variance, and Covariance

3.4.1, 3.4.2

Learning Objectives

At the end of this unit, students should be able to:

1. Define the mean/expected value, variance, and standard deviation of X, where X is either a discrete or continuous random variable.

2. Compute the mean, variance, and standard deviation of a given random variable X (either where the pmf/pdf is given explicitly, or where X is a particular type of rv, e.g., binomial).

3. Compute the mean, variance, and standard deviation of a function of a random variable (i.e., g(X)).

4. Simplify the calculations in 3 when g is a linear function.

5. Define, compute, and interpret the covariance between two random variables X and Y.

6. Define,compute,and interpret the correlation between two random variables X and Y.

7. State and prove important properties about the mean (i.e., linearity), the variance, and covariance.

$\bar{x} = \frac{1}{n}\sum x_i = \sum \underbrace{\frac{1}{n}}_{\text{weight}} x_i$

Mean/Expected Value

Consider a university having 15,000 students and let *X* equal the number of courses for which a randomly selected student is registered.

The pdf of *X* is given to you as follows:

x	1	2	3	4	5	6	7
$p(x) = P(X=x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

Students pay more money when enrolled in more courses, and so the university wants to know what the average number of courses students take per semester.

Mean/Expected Value

Definition: For a discrete random variable *X* with pdf/pmf $P(X=x)$ $\cancel{f(X)}$, the *expected value* or mean value of *X* is denoted as *E(X)* and is calculated as:

$$E(X) = \sum_{x} x\,P(X=x)$$

Mean/Expected Value

What is E(X)?

x	1	2	3	4	5	6	7
$p(x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

$E(X) = (1)(0.01) + (2)(0.03) + \ldots + (7)(0.02) = 4.57$

Mean/Expected Value

Definition: For a continuous random variable *X* with pdf *f(x)*, the *expected value* or mean value of *X* is calculated as:

$$E(X) = \int_x x f(x)\,dx = \int_{-\infty}^{\infty} x f(x)\,dx$$

Reminder : $\int u\,dv = uv - \int v\,du$ (IBP) $\quad \boxed{\text{LIPET}}$ log, inv trig, powers, exp, Trig

Mean/Expected Value

<u>Example:</u> The lifetime (in years) of a certain brand of battery is exponentially distributed with $\boxed{\lambda}$ *= 0.25*.

$u = x \qquad v = -e^{-\lambda x}$

$du = dx \qquad dv = \lambda e^{-\lambda x}dx$

$-xe^{-\lambda x} + \int e^{-\lambda x}dx = \boxed{-xe^{-\lambda x} - \frac{1}{\lambda}e^{-\lambda x} + c}$

How long, on average, will the battery last?

$$E(X) = \int_0^\infty \underset{u}{x}\, \underbrace{\lambda e^{-\lambda x}dx}_{dv} = \lim_{b\to\infty}\left[-xe^{-\lambda x} - \frac{1}{\lambda}e^{-\lambda x}\right]_0^b$$

$$= \lim_{b\to\infty}\left[\cancel{-be^{-\lambda(b)}}^{\to 0} - \cancel{\frac{1}{\lambda}e^{-\lambda(b)}}^{\to 0} + 0 + \frac{1}{\lambda}e^{-\lambda(0)}\right] = \frac{1}{\lambda}.$$

Expected Value of a Function

If a *discrete* r.v. X has a a pmf $P(X = x)$, then the expected value of any function $\boxed{g(X)}$, computed as:

$$E\Big(g(x)\Big) = \sum_x g(x) P(X=x)$$

Note that $E[g(X)]$ is computed in the same way that $E(X)$ itself is, except that $g(x)$ is substituted in place of x.

Expected Value of a Function

If a *continuous* r.v. X has a a pdf $f(x)$, then the expected value of any function $g(X)$, computed as:

$$E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x)\, dx$$

Note that $E[g(X)]$ is computed in the same way that $E(X)$ itself is, except that $g(x)$ is substituted in place of x.

Expected Value of a Function

Example: A random variable *X* has pdf:

$$f(x) = \frac{3}{4}(1-x^2), \quad -1 \le x \le 1$$

① What is $E(X^3)$? $g(x) = x^3$

$$E(X^3) = \int_{-1}^{1} x^3\left(\frac{3}{4}(1-x^2)\right)dx = \frac{3}{4}\left(\frac{x^4}{4} - \frac{x^6}{6}\right)\Bigg|_{-1}^{1} = 0$$

Review: What is $F(x)$?

$$P(X \le x) = F(x) = \int_{-1}^{x} \frac{3}{4}(1-t^2)\,dt = \frac{3}{4}x - \frac{1}{4}x^3 + \frac{1}{2} = F(x)$$

Expected Value of a *Linear* Function

a, b constants

If $g(X)$ is a *linear function* of X *(i.e.,* $\boxed{g(X) = aX + b}$*)* then $E[g(X)]$ can be easily computed from $E(X)$.

Theorem: $\boxed{E(g(X)) = E(aX+b) = aE(X)+b}$

Proof:

$$E(g(X)) = \sum_x (ax+b)P(X=x) = \sum_x \left[axP(X=x) + bP(X=x)\right]$$

$$= \sum_x axP(X=x) + \sum_x bP(X=x) = a\underbrace{\sum_x xP(X=x)}_{E(X)} + b\underbrace{\sum_x P(X=x)}_{=1} = aE(X)+b$$

Note: This works for continuous and discrete random variables.

Expected Value of a *Linear* Function

<u>Example:</u> Consider a university having 15,000 students and let *X* equal the number of courses for which a randomly selected student is registered.

The pdf of *X* is given to you as follows:

x	1	2	3	4	5	6	7
$p(x) = P(X=x)$	.01	.03	.13	.25	.39	.17	.02
~~Number registered~~	150	450	1950	3750	5850	2550	300

$g(x) = 500X + 100$

Earlier, we calculated that *E(X)* was 4.57. If students pay \$500 per course plus a \$100 per-semester registration fee, what is the average amount of money the university can expect a student to pay each a semester?

$E(g(X)) = 500E(X) + 100 = 500(4.57) + 100 = \$2,385.$

$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

Variance of a Random Variable

Definition: For a discrete random variable *X* with pdf/pmf $P(X=x) =$ *f(x)*, the *variance* of *X* is denoted as $Var(X) = \sigma^2$ or σ_X^2 and is calculated as: Let $\mu = E(X)$

$$\sigma^2 = Var(X) = \sum_x (x-\mu)^2 P(X=x)$$

The standard deviation (SD) of *X* is

$$\sigma = \sqrt{Var(X)} = \sqrt{\sigma^2}$$

Variance of a Random Variable

Example: Consider a university having 15,000 students and let *X* equal the number of courses for which a randomly selected student is registered.

The pdf of *X* is given to you as follows:

x	1	2	3	4	5	6	7
$p(x) = P(X=x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

Earlier, we calculated that *E(X)* was 4.57. What is *Var(X)*? What about sd(X)?

$$Var(X) = (1-4.57)^2(0.01) + (2-4.57)^2(0.03) + \dots (7-4.57)^2(0.02) = 1.26\ldots$$

$$sd(X) = \sigma \approx \sqrt{1.26} \approx 1.12\ldots$$

Variance of a Random Variable

Definition: For a continuous random variable *X* with pdf *f(x)*, the variance of *X* is denoted as $Var(X) = \sigma^2$ and is calculated as:

$$Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

A Shortcut Formula for Variance

The variance can also be calculated using an alternative formula:

$$Var(X) = E(X^2) - [E(X)]^2 = \boxed{E(X^2) - \mu^2}$$

Proof? $Var(X) = \int (x-\mu)^2 f(x)\,dx = \int (x^2 - 2x\mu + \mu^2) f(x)\,dx$

$$= \underbrace{\int x^2 f(x)\,dx}_{E(X^2)} - 2\mu \underbrace{\int x f(x)\,dx}_{=E(X)=\mu} + \mu^2 \underbrace{\int f(x)\,dx}_{=1} = E(X^2) - 2\mu^2 + \mu^2$$

$$= E(X^2) - \mu^2$$

Why would we use this equation instead?

less computation Time

Variance of a Function

The variance of $g(X)$ is calculated as follows:

Discrete:

$$Var(g(X)) = \sum_x (g(x) - E(g(x)))^2 P(X=x)$$

Continuous:

$$Var(g(X)) = \int_x (g(x) - E(g(x)))^2 f(x)dx$$

Variance of a Function

$ax+b$

$E(X) = \mu$

As with the expected value, there is also a shortcut formula if *g(X)* is a linear function of *X*:

$E(g(x)) = \mu_g$

$$\boxed{Var(g(X)) = a^2 Var(X)}$$

Proof: $Var(g(X)) = \int (g(x) - \mu_g)^2 f(x)\,dx = \int (ax+b - a\mu - b)^2 f(x)\,dx$

$$= \int [a(x-\mu)]^2 f(x)\,dx = a^2 \underbrace{\int (x-\mu)^2 f(x)\,dx}_{Var(x)} = a^2 Var(x).$$

Can we do a simple proof to show this is true?

$Var(X) = 1.26$

$\boxed{sd(Y) = 500\, sd(X)}$

Variance of a Function

Example: Consider a university having 15,000 students and let *X* equal the number of courses for which a randomly selected student is registered.

The pdf of *X* is given to you as follows:

x	1	2	3	4	5	6	7
$p(x) = P(X = x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

Cost for a student $= g(X) = Y = 500X + 100$

Earlier, we calculated that *E(X)* was 4.57. If students pay $500 per course plus a $100 per-semester registration fee, what is the expected **standard deviation** of the amount of money students to pay each a semester?

$$Var(Y) = Var(500X + 100) = 500^2 Var(X) = 500^2(1.26) = 315,000$$

$$sd(Y) = \sqrt{Var(Y)} \approx 561.25$$

$g(x) = 1000x + 200(3-x) - 1500$

$\boxed{g(x) = 800x - 900}$

$Var(x) = E(x^2) - [E(x)]^2$

$= 0^2(0.1) + 1^2(0.2) + 2^2(0.3) + 3^2(0.4) - 4 = 1$

Classwork

$\Rightarrow sd(x) = 1$

A computer store has purchased 3 computers of a certain type at \$500 each. It will sell them for \$1000 each. The manufacturer has agreed to repurchase any computers still unsold after a specified period at \$200 each. Let *X* denote the number of computers sold, and suppose that:

P(X = 0) = 0.1,   P(X = 1) = 0.2,   P(X = 2) = 0.3, P(X = 3) = 0.4.

$E(x) = (0)(0.1) + (1)(0.2) + (2)(0.3) + (3)(.4) = 2$

What is the expected profit? What is the standard deviation of the profit?

$E(g(x)) = 800E(x) - 900 = 800(2) - 900 = \$700.$

$sd(g(x)) = 800\,sd(x) = \$800$

Classwork

1. For a discrete rv X, show that $E(aX + b) = a\ E(X) + b$.

Done earlier

2. **Show** that $\underbrace{E(X - E(X))^2}_{Var(X)} = E(X^2) - (E(X))^2$

Done earlier

Binomial Mean and Variance

Let $X \sim \text{Bin}(n, p)$. Then:

1. $E(X) = np$
2. $\text{Var}(X) = np(1 - p)$.

Proof?

1.) Let $Y_1, \ldots, Y_n \sim \text{Bern}(p)$. Then $X = \sum_{i=1}^{n} Y_i \sim \text{Bin}(n,p)$

$$E(Y_i) = (0)P(Y_i = 0) + (1)P(Y_i = 1) = P(Y_i = 1) = p$$

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

2.) ...

Binomial Mean and Variance

Example: A biased coin is tossed 10 times, so that the odds of heads are 3:1.

What notation do we use to describe X?

What is the mean of X? The variance?

Mean and Variance for Other Distributions

Distribution	E(X)	Var(X)
Geom(π)	$1/\pi$	$(1-\pi)/\pi$
NB(r, π)	$r*\pi/(1-\pi)$	$r*\pi/(1-\pi)^2$
Poisson(λ)	λ	λ
Uniform(a,b)	$\frac{1}{2}(a+b)$	$1/12*(b-a)^2$
Exp(λ)	$1/\lambda$	$1/\lambda^2$
~~Weib(α,β)~~	$\cancel{\beta\Gamma(1+1/\alpha)}$	$\cancel{\beta^2\{\Gamma(1+2/\alpha)-[\Gamma(1+1/\alpha)]^2\}}$
Beta(α,β)	$\alpha/(\alpha+\beta)$	$\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$

STAT 2600 18-19

Sp. 18: APPM 2720

Covariance

When two random variables *X* and *Y* are not independent, it is frequently of interest to assess how strongly they are related to one another.

$E(X) = \mu_x$

$E(Y) = \mu_y$

Definition: The *covariance* between two rv's *X* and *Y* is defined as:

$$\mathrm{Cov}(X,Y) = \begin{cases} \sum_{x,y} (x-\mu_x)(y-\mu_y)P(X=x, Y=y) & \text{(discrete)} \\ \int_x \int_y (x-\mu_x)(y-\mu_y) f(x,y)\,dx\,dy & \text{(cont.)} \end{cases}$$

Covariance

If both variables tend to deviate in the same direction (both go above their means or below their means at the same time), then the covariance will be positive.

If the opposite is true, the covariance will be negative.

linearly

If X and Y are not strongly related, the covariance will be near 0.

Covariance

The following shortcut formula for Cov(X, Y) simplifies the computations.

<u>Theorem</u>:

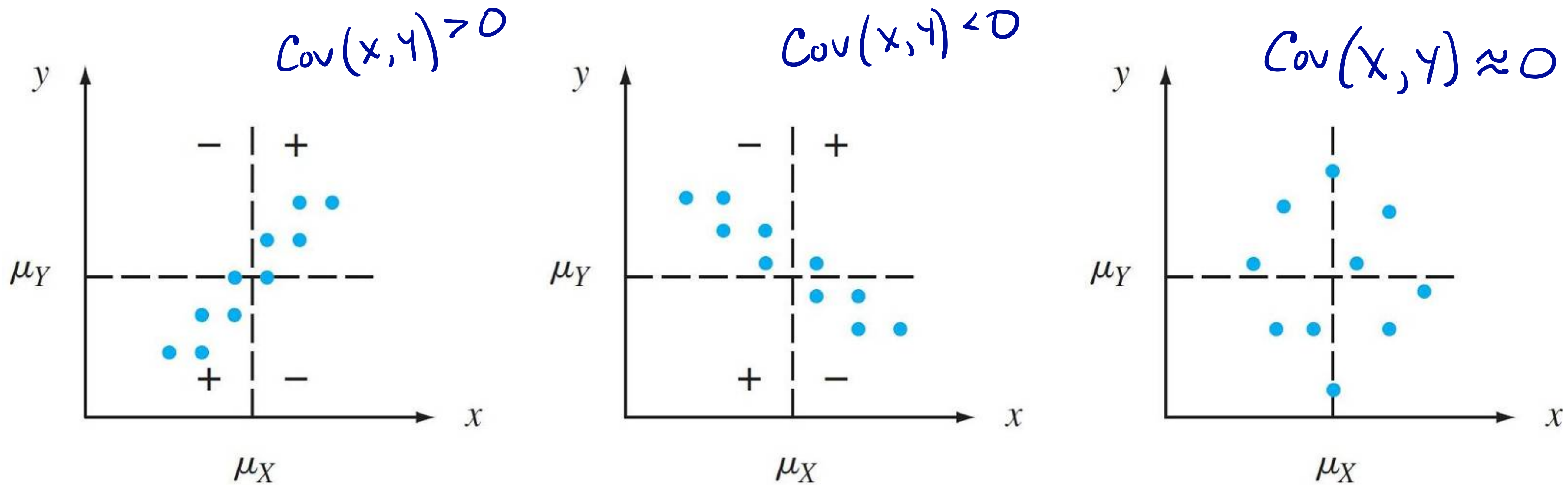
$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

According to this formula, no intermediate subtractions are necessary. This is analogous to the shortcut for the variance computation we saw earlier.

Covariance

The covariance depends on *both* the set of possible pairs and the probabilities of those pairs.

Below are examples of 3 types of “co-varying”:



Covariance

Example: An insurance agency services customers who have both a homeowner's policy and an automobile policy. For each type of policy, a deductible amount must be specified. For an automobile policy, the choices are \$100 and \$250, whereas for a homeowner's policy, the choices are \$0, \$100, and \$200.

Suppose an individual--Bob--is selected at random from the agency's files. Let X = his deductible amount on the auto policy and Y = his deductible amount on the homeowner's policy.

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \sum_x \sum_y (xy)P(X=x, Y=y) - (175)(125)$$

$$= (100)(0)(0.2) + (100)^2(0.1) + (100)(200)(0.2) + (250)(100)(0.15) + (250)(200)(0.3) - (175)(125) = 1875$$

Covariance

$E(X) = 175$; $E(Y) = 125$

Suppose the joint pmf is given by the insurance company in the accompanying **joint probability table:**

$p(x, y) = P(X=x, Y=y)$		y: 0	100	200	$P(X=x)$
x	100	.20	.10	.20	0.5
	250	.05	.15	.30	0.5
	$P(Y=y)$	0.25	0.25	0.5	1

What is the covariance between X and Y?

Important Properties

a, b are constants

1. $E(aX+bY) = aE(X) + bE(Y)$
2. $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2ab*Cov(X, Y)$
3. $Var(aX – bY) = a^2Var(X) + b^2Var(Y) - 2ab*Cov(X, Y)$

Correlation

Definition: The *correlation coefficient* of X and Y, denoted by $\underline{\text{Corr}(X,Y)}$ or just $\underline{\rho_{X,Y}}$, is defined by:

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

It represents a "scaled" covariance – correlation ranges between -1 and 1.

$Cov(X,Y) = 1875$; $Var(X) = E(X^2) - [E(X)]^2 = (100)^2(0.5) + (250)^2(0.5) - 175^2 = 5625$

$Var(Y) = E(Y^2) - [E(Y)]^2 = 0^2(0.25) + (100)^2(0.25) + (200)^2(0.5) - 125^2$
$= 6875$

Correlation

$\sigma_X = 75$

$\sigma_Y \approx 83$

In the insurance example, what is the correlation between *X* and *Y*?

$p(x, y)$		y: 0	100	200	$P(X=x)$
x	100	.20	.10	.20	0.5
	250	.05	.15	.30	0.5
	$P(Y=y)$	0.25	0.25	0.5	

$$\rho_{X,Y} = \frac{1875}{(75)(83)} \approx 0.3$$

Correlation

Important Properties of Covariance and Correlation:

1. Cov(aX + b, cY + d) = a c Cov (X, Y)

2. Corr(aX + b, cY + d) = sgn(ac) Corr(X, Y)

3. For any two rv's X and Y, $-1 \leq$ Corr(X, Y) ≤ 1

4. $\rho_{x,y}$ = 1 or –1 iff Y = aX + b for some numbers a and b with
   a 0.

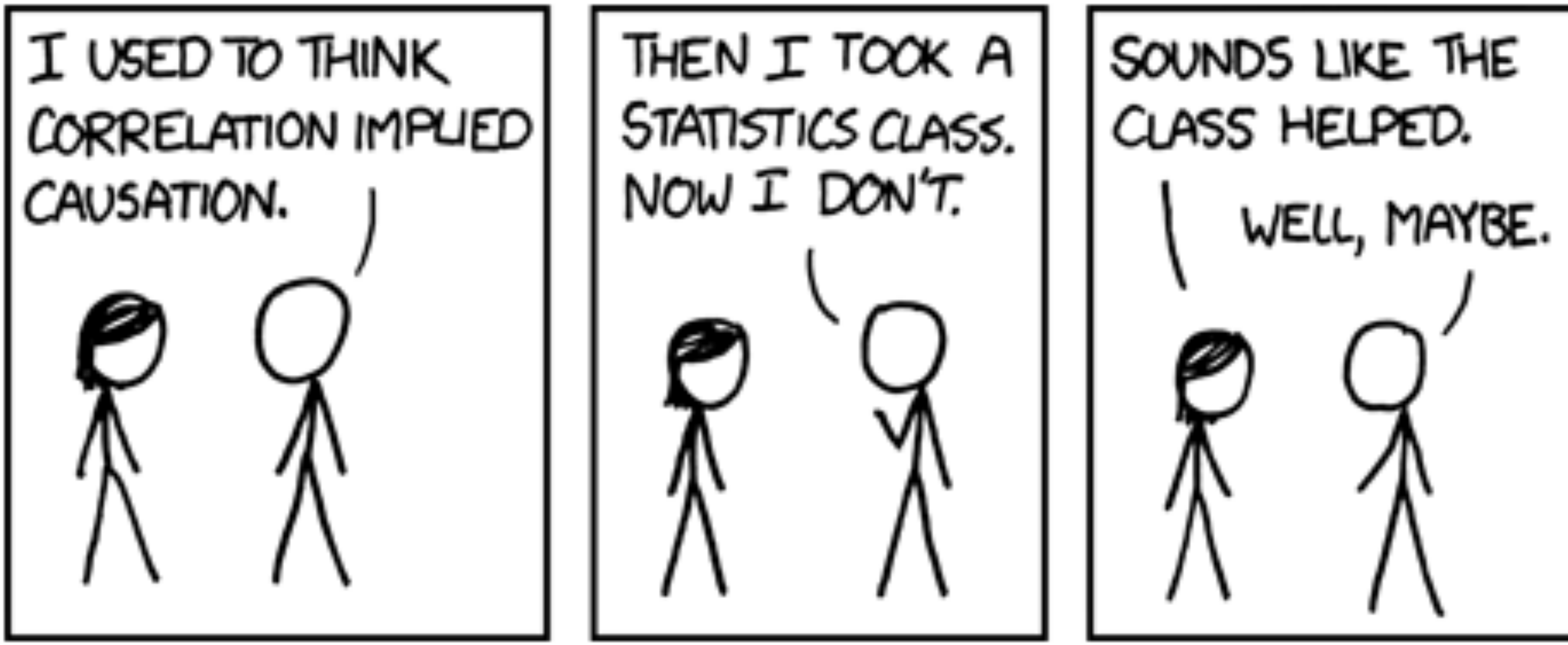
Correlation

If X and Y are independent, then $\rho_{x,y} = 0$, but $\rho_{x,y} = 0$ does not imply independence.

The correlation coefficient is a measure of the **linear relationship** between X and Y, and only when the two variables are perfectly related in a linear manner will be as positive or negative as it can be.

Two variables could be uncorrelated yet highly dependent because there is a strong nonlinear relationship, so be careful not to conclude too much from knowing that $\rho_{x,y} = 0$.

David Hume – Causality



xkcd.com/552/