4 The mean, variance and covariance

(Chs 3.4.1, 3.4.2)

Mean (Expected Value) of X

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
$p(x)$	.01	.03	.13	.25	.39	.17	.02

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) of *X*

For a discrete random variable *X* with pdf f(*x*), the **expected value** or mean value of *X* is denoted as *E(X)* and is calculated as:

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

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
$p(x)$	.01	.03	.13	.25	.39	.17	.02

What is E(X)?

Mean (Expected Value) of *X*

For a discrete random variable *X* with pdf *f(x)*, the expected value or mean value of *X* is denoted as as *E(X)* and is calculated as:

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

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

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

Example

The lifetime (in years) of a certain brand of battery is exponentially distributed with $\lambda = 0.25$.

How long, on average, will the battery last?

Example

The lifetime (in years) of a certain brand of battery is exponentially distributed with $\lambda = 0.25$.

How long, on average, will the battery last?

Review: How long will it be before 50% of the batteries fail?

The Expected Value of a Function

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

$$E[g(X)] = \sum_{x} g(x) * P(X = x)$$

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*.

The Expected Value of a Function

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

$$E[g(X)] = \sum_x g(x) * P(X = x)$$

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

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) \cdot f(x)\, dx$$

The Expected Value of a Function

A random variable X is distributed $f(x)$ such that

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

What is $E(X^3)$?

The Expected Value of a Function

A random variable *X* is distributed *f(x)* such that

$$f(x) = \tfrac{3}{4}(1-x^2),\ -1 <= x <= 1.$$

What is $E(X^3)$?

Review: What is *F(x)*?

Expected Value of $g(X) = aX+b$

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

Proposition

$$E(aX + b) = a \cdot E(X) + b$$

This works for continuous and discrete random variables.

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
$p(x)$	.01	.03	.13	.25	.39	.17	.02

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 to make from a student each a semester?

The Variance of *X*

For a discrete random variable *X* with pdf *f(x)*, the variance of *X* is denoted as Var(X) = σ^2_X and is calculated as:

$$Var[X] = E[(X-\mu)^2] = \sum_x (x-\mu)^2 * P(X=x)$$

The standard deviation (SD) of *X* is

$$\sigma_X = \sqrt{\sigma_X^2}$$

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
$p(x)$	.01	.03	.13	.25	.39	.17	.02

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

The Variance of *X*

For a discrete random variable *X* with pdf *f(x)*, the variance of *X* is denoted as Var(X) = σ^2_X and is calculated as:

$$Var[X] = E[(X-\mu)^2] = \sum_x (x-\mu)^2 * P(X=x)$$

Similarly, for a continuous random variable *X* with pdf *f(x)*, the variance of *X* is denoted as Var(X) = σ^2_X and is calculated as:

$$Var[X] = \int_x (x-\mu)^2 f(x)dx$$

A Shortcut Formula for σ^2

The variance can also be calculated using an alternative formula:

$$V(x) = \sigma^2 = E(X^2) - E(X)^2$$

Why would we use this equation instead?

The Variance of a Function

The variance of *g*(*X*) is calculated similarly to what we have seen:

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

with a similar form for the discrete case.

The Variance of a Function

The variance of *g*(*X*) is calculated similarly to what we have seen:

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

with a similar form for the discrete case.

As with the expected value, there is also a short-cut formula if *g(X)* is a linear function of *X*.

$$Var[aX + b] = a^2 Var[X]$$

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

Classwork #1

A computer store has purchased 3 computers of a certain type at $500 apiece. It will sell them for $1000 apiece. The manufacturer has agreed to repurchase any computers still unsold after a specified period at $200 apiece.

Let X denote the number of computers sold, and suppose that

$$p(0) = .1, \quad p(1) = .2, \quad p(2) = .3 \quad \text{and} \quad p(3) = .4.$$

What is the expected profit?

Classwork #2

Show that if $E(X) = \sum_x x * P(X = x)$

then

$$E(aX + b) = a \cdot E(X) + b$$

Show that

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

Binomial Mean and Variance

If $X \sim \text{Bin}(n, p)$, then

Expectation: $E(X) = np$ (let's prove this one)

Variance: $V(X) = np(1 - p) = npq$

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?

Example

cont'd

NOTE: Even though X can take on only integer values, $E(X)$ need not be an integer.

If we perform a large number of independent binomial experiments, each with $n = 10$ trials and $p = .75$, then the average number of successes per experiment will be close to 7.5.

What is the probability that X is within 1 standard deviation of its mean value?

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)	½(a+b)	$1/12*(b-a)^2$
Exp(λ)	$1/\lambda$	$1/\lambda^2$
Weib(α,β)	$\beta\Gamma(1+1/\alpha)$	$\beta^2\{\Gamma(1+2/\alpha)-[\Gamma(1+1/\alpha)]^2\}$
Beta(α,β)	$\alpha/(\alpha+\beta)$	$\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$

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.

The **covariance** between two rv's *X* and *Y* is

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

$$= \begin{cases} \sum_x \sum_y (x - \mu_X)(y - \mu_Y) p(x, y) & X, Y \text{ discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y)\, dx\, dy & X, Y \text{ continuous} \end{cases}$$

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.

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

Covariance shortcut

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

Proposition

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

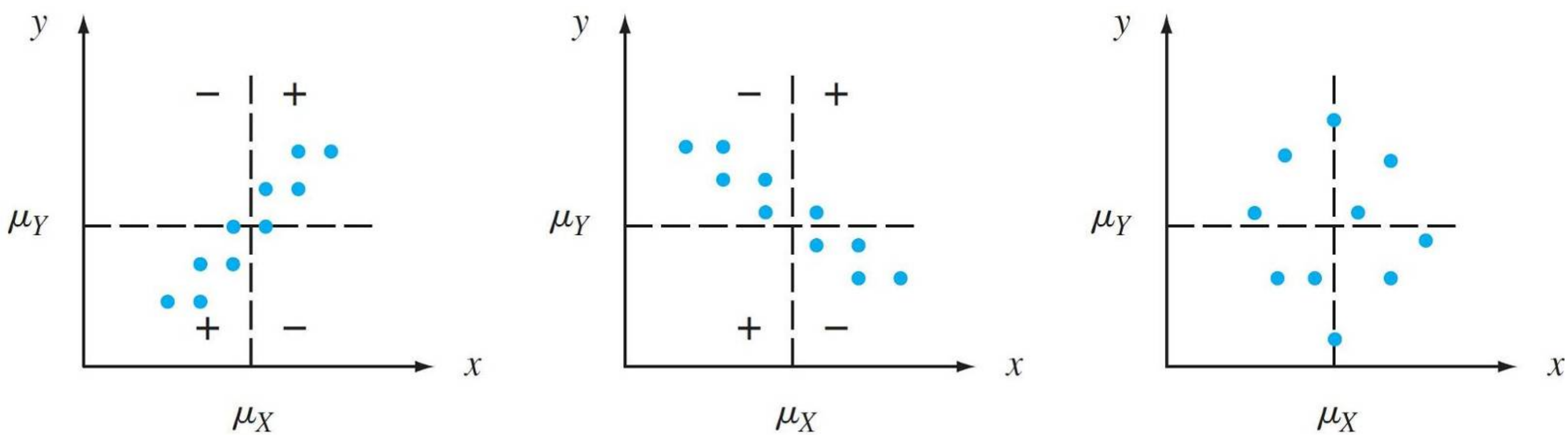
According to this formula, no intermediate subtractions are necessary; only at the end of the computation is $\mu_X \cdot \mu_Y$ subtracted from $E(XY)$.

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":



(a) positive covariance; (b) negative covariance; (c) covariance near zero

Figure 5.4

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.

Example

cont' d

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

$p(x, y)$		y: 0	100	200
x	100	.20	.10	.20
	250	.05	.15	.30

What is the covariance between X and Y?

Correlation

Definition

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

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

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

Example

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

Correlation

Propositions

1. $\text{Cov}(aX + b, cY + d) = a\,c\,\text{Cov}(X, Y)$

2. $\text{Corr}(aX + b, cY + d) = sgn(ac)\,\text{Corr}(X, Y)$

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

4. $\rho = 1$ or -1 iff $Y = aX + b$ for some numbers a and b with $a \neq 0$.

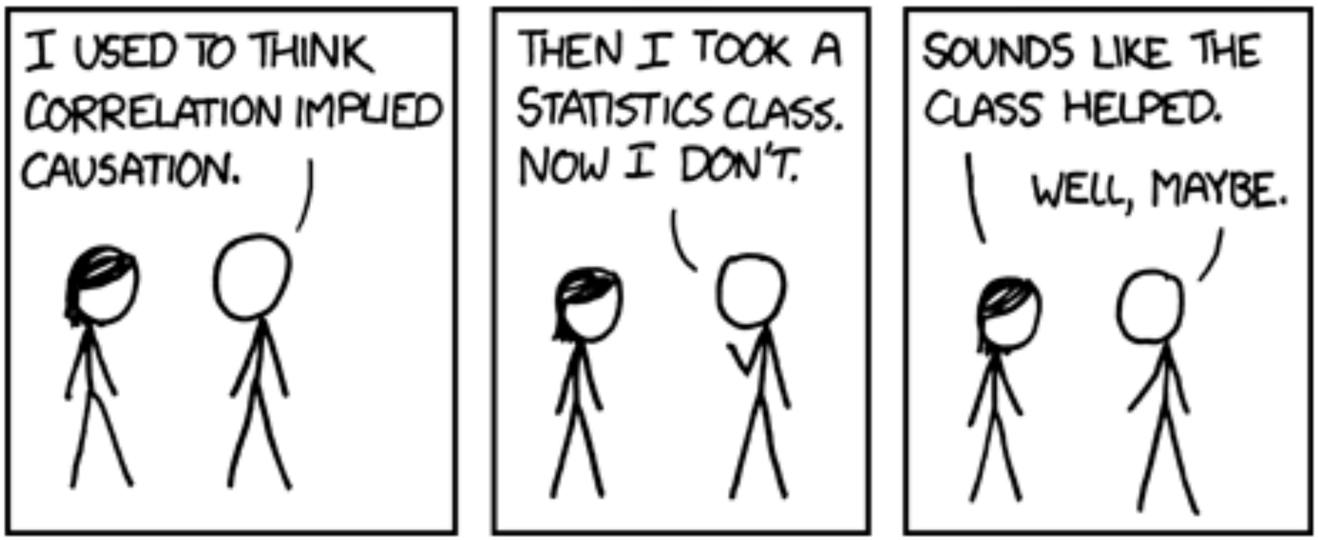
Correlation

If X and Y are independent, then $\rho = 0$, but $\rho = 0$ does not imply independence.

The correlation coefficient ρ 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 ρ 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 $\rho = 0$.

Interpreting Correlation



xkcd.com/552/

Other Useful Rules

1. E(aX+bY) = aE(X) + bE(Y)

2. Var(aX + bY) = a2Var(X) + b2Var(Y) + 2abcov(X, Y)
3. Var(aX – bY) = a2Var(X) + b2Var(Y) - 2abcov(X, Y)