Joint Distributions, Independence
Covariance and Correlation
18.05 Spring 2014

Jeremy Orloff and Jonathan Bloom

$X\backslash Y$	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

Central Limit Theorem

Setting: $X_1, X_2, \ldots$ i.i.d. with mean μ and standard dev. σ.

For each n:

$$\overline{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$$
$$S_n = X_1 + X_2 + \ldots + X_n.$$

Conclusion: For large n:

$$\overline{X}_n \approx \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$
$$S_n \approx \mathrm{N}\left(n\mu, n\sigma^2\right)$$

Standardized S_n or $\overline{X}_n \approx \mathrm{N}(0, 1)$

Board Question: CLT

1. Carefully write the statement of the central limit theorem.

2. To head the newly formed US Dept. of Statistics, suppose that 50% of the population supports Erika, 25% supports Ruthi, and the remaining 25% is split evenly between Peter, Jon and Jerry.

A poll asks 400 random people who they support. What is the probability that at least 55% of those polled prefer Erika?

3. (Not for class. Solution will be on posted slides.)
An accountant rounds to the nearest dollar. We'll assume the error in rounding is uniform on [-.5, .5]. Estimate the probability that the total error in 300 entries is more than $5.

Solution on next page

Solution 2

answer: 2. Let $\mathcal{E}$ be the number polled who support Erika.
The question asks for the probability $\mathcal{E} > .55 \cdot 400 = 220$.

$E(\mathcal{E}) = 400(.5) = 200$ and $\sigma^2_{\mathcal{E}} = 400(.5)(1 - .5) = 100 \Rightarrow \sigma_{\mathcal{E}} = 10$.

Because $\mathcal{E}$ is the sum of 400 Bernoulli(.5) variables the CLT says it is approximately normal and standardizing gives

$$\frac{\mathcal{E} - 200}{10} \approx Z \qquad \text{and} \qquad P(\mathcal{E} > 220) \approx P(Z > 2) \approx .025$$

3. Let X_j be the error in the j^{th} entry, so, $X_j \sim U(-.5, .5)$.

We have $E(X_j) = 0$ and $\text{Var}(X_j) = 1/12$.

The total error $S = X_1 + \ldots + X_{300}$ has $E(S) = 0$,
$\text{Var}(S) = 300/12 = 25$, and $\sigma_S = 5$.

Standardizing we get, by the CLT, $S/5$ is approximately standard normal.
That is, $S/5 \approx Z$.

So $P(S < 5 \text{ or } S > 5) \approx P(Z < 1 \text{ or } Z > 1) \approx \boxed{.32}$.

Joint Distributions

X and Y are *jointly distributed* random variables.

Discrete: Probability mass function (pmf):

$$p(x_i, y_j)$$

Continuous: probability density function (pdf):

$$f(x, y)$$

Both: cumulative distribution function (cdf):

$$F(x, y) = P(X \leq x, Y \leq y)$$

Discrete joint pmf: example 1

Roll two dice: $X = \#$ on first die, $Y = \#$ on second die

X takes values in 1, 2, ..., 6, Y takes values in 1, 2, ..., 6

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

pmf: $p(i, j) = 1/36$ for any i and j between 1 and 6.

Discrete joint pmf: example 2

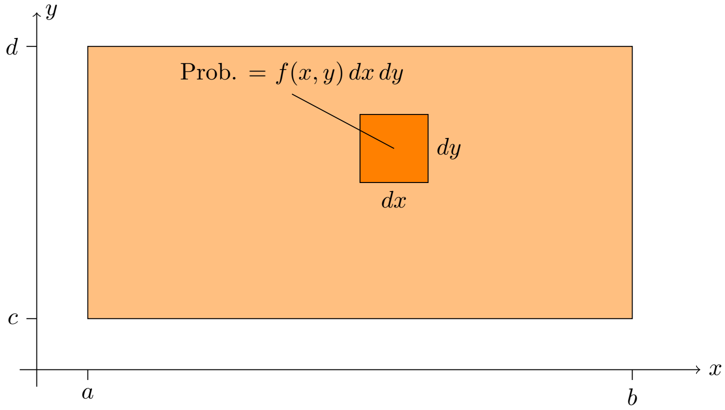
Roll two dice: $X = \#$ on first die, $T =$ total on both dice

$X\backslash T$	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

Continuous joint distributions

- X takes values in $[a, b]$, Y takes values in $[c, d]$
- (X, Y) takes values in $[a, b] \times [c, d]$.
- *Joint probability* **density** *function* (pdf) $f(x, y)$

$f(x, y)\,dx\,dy$ is the probability of being in the small square.



Properties of the joint pmf and pdf

Discrete case: probability mass function (pmf)

1. $0 \le p(x_i, y_j) \le 1$
2. Total probability is 1.

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) = 1$$

Continuous case: probability density function (pdf)

1. $0 \le f(x, y)$
2. Total probability is 1.

$$\int_c^d \int_a^b f(x, y)\, dx\, dy = 1$$

Note: $f(x, y)$ can be greater than 1: it is a density *not* a probability.

Example: discrete events

Roll two dice: $X = \#$ on first die, $\quad Y = \#$ on second die.

Consider the event: $A =$ '$Y - X \geq 2$'

Describe the event A and find its probability.

Example: discrete events

Roll two dice: $X = \#$ on first die, $Y = \#$ on second die.

Consider the event: $A =$ '$Y - X \geq 2$'

Describe the event A and find its probability.

answer: We can describe A as a set of (X, Y) pairs:

$$A = \{(1,3),(1,4),(1,5),(1,6),(2,4),(2,5),(2,6),(3,5),(3,6),(4,6)\}.$$

Or we can visualize it by shading the 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

$P(A) =$ sum of probabilities in shaded cells $= 10/36$.

Example: continuous events

Suppose (X, Y) takes values in $[0, 1] \times [0, 1]$.

Uniform density $f(x, y) = 1$.

Visualize the event '$X > Y$' and find its probability.

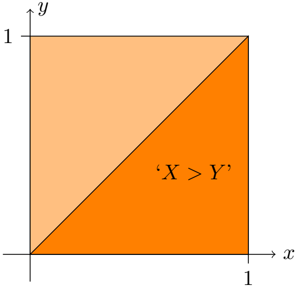
Example: continuous events

Suppose (X, Y) takes values in $[0, 1] \times [0, 1]$.

Uniform density $f(x, y) = 1$.

Visualize the event '$X > Y$' and find its probability.

answer:



The event takes up half the square. Since the density is uniform this is half the probability. That is, $P(X > Y) = .5$

Cumulative distribution function

$$F(x, y) = P(X \le x,\ Y \le y) = \int_c^y \int_a^x f(u, v)\, du\, dv.$$

$$f(x, y) = \frac{\partial^2 F}{\partial x \partial y}(x, y).$$

Properties

1. $F(x, y)$ is non-decreasing. That is, as x or y increases $F(x, y)$ increases or remains constant.
2. $F(x, y) = 0$ at the lower left of its range.
   If the lower left is $(-\infty, -\infty)$ then this means

$$\lim_{(x,y)\to(-\infty,-\infty)} F(x, y) = 0.$$

3. $F(x, y) = 1$ at the upper right of its range.

Marginal pmf

Roll two dice: $X = \#$ on first die, $T =$ total on both dice.

The marginal pmf of X is found by summing the rows. The marginal pmf of T is found by summing the columns

$X \backslash T$	2	3	4	5	6	7	8	9	10	11	12	$p(x_i)$
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
$p(t_j)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

Marginal pdf

Example. Suppose X and Y take values on the square $[0,1] \times [1,2]$ with joint pdf $f(x,y) = \frac{8}{3}x^3y$.

The marginal pdf $f_X(x)$ is found by *integrating out* the y. Likewise for $f_Y(y)$.

answer:

$$f_X(x) = \int_1^2 \frac{8}{3}x^3y\,dy = \left[\frac{4}{3}x^3y^2\right]_1^2 = \boxed{4x^3}$$

$$f_Y(y) = \int_0^1 \frac{8}{3}x^3y\,dx = \left[\frac{2}{3}x^4y^1\right]_0^1 = \boxed{\frac{2}{3}y}.$$

Board question

Suppose X and Y are random variables and

- (X, Y) takes values in $[0, 1] \times [0, 1]$.
- the pdf is $\frac{3}{2}(x^2 + y^2)$.

1. Show $f(x, y)$ is a valid pdf.
2. Visualize the event $A =$ '$X > .3$ and $Y > .5$'. Find its probability.
3. Find the cdf $F(x, y)$.
4. Find the marginal pdf $f_X(x)$. Use this to find $P(X < .5)$.
5. Use the cdf $F(x, y)$ to find the marginal cdf $F_X(x)$ and $P(X < .5)$.
6. See next slide

Board question continued

6. (New scenario) From the following table compute $F(3.5, 4)$.

$X \backslash Y$	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

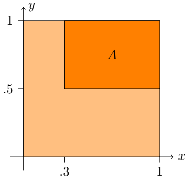
answer: See next slide

Solution

answer: 1. Validity: Clearly $f(x,y)$ is positive. Next we must show that total probability $= 1$:

$$\int_0^1 \int_0^1 \frac{3}{2}(x^2+y^2)\,dx\,dy = \int_0^1 \left[\frac{1}{2}x^3 + \frac{3}{2}xy^2\right]_0^1 dy = \int_0^1 \frac{1}{2} + \frac{3}{2}y^2\,dy = 1.$$

2. Here's the visualization



The pdf is not constant so we must compute an integral

$$P(A) = \int_{.3}^1 \int_{.5}^1 4xy\,dy\,dx = \int_{.3}^1 \left[2xy^2\right]_{.5}^1 dx = \int_{.3}^1 \frac{3x}{2}\,dx = \boxed{0.525}.$$

(continued)

Solutions 3, 4, 5

3.

$$F(x,y) = \int_0^y \int_0^x \frac{3}{2}(u^2+v^2)\,du\,dv = \boxed{\frac{x^3y}{3} + \frac{xy^3}{3}}.$$

4.

$$f_X(x) = \int_0^1 \frac{3}{2}(x^2+y^2)\,dy = \left[\frac{3}{2}x^2y + \frac{y^3}{2}\right]_0^1 = \boxed{\frac{3}{2}x^2 + \frac{1}{2}}$$

$$P(X < .5) = \int_0^{.5} f_X(x)\,dx = \int_0^{.5} \frac{3}{2}x^2 + \frac{1}{2}\,dx = \left[\frac{1}{2}x^3 + \frac{1}{2}x\right]_0^{.5} = \boxed{\frac{5}{16}}.$$

5. To find the marginal cdf $F_X(x)$ we simply take y to be the top of the y-range and evalute F:

$$F_X(x) = F(x,1) = \frac{1}{2}(x^3 + x).$$

Therefore $\quad P(X < .5) = F(.5) = \frac{1}{2}(\frac{1}{8} + \frac{1}{2}) = \boxed{\frac{5}{16}}.$

6. On next slide

Solution 6

6. $F(3.5, 4) = P(X \leq 3.5, Y \leq 4)$.

$X \backslash Y$	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

Add the probability in the shaded squares: $F(3.5, 4) = 12/36 = 1/3$.

Independence

Events A and B are independent if

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

Random variables X and Y are independent if

$$F(x, y) = F_X(x)F_Y(y).$$

Discrete random variables X and Y are independent if

$$p(x_i, y_j) = p_X(x_i)p_Y(y_j).$$

Continuous random variables X and Y are independent if

$$f(x, y) = f_X(x)f_Y(y).$$

Concept question: independence I

Roll two dice: X = value on first, Y = value on second

$X\backslash Y$	1	2	3	4	5	6	$p(x_i)$
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
$p(y_j)$	1/6	1/6	1/6	1/6	1/6	1/6	1

Are X and Y independent?   1. Yes   2. No

answer: 1. Yes. Every cell probability is the product of the marginal probabilities.

Concept question: independence II

Roll two dice: X = value on first, T = sum

$X\backslash T$	2	3	4	5	6	7	8	9	10	11	12	$p(x_i)$
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
$p(y_j)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

Are X and Y independent? 1. Yes 2. No

answer: 2. No. The cells with probability zero are clearly not the product of the marginal probabilities.

Concept Question

Among the following pdf's which are independent? (Each of the ranges is a rectangle chosen so that $\int\int f(x, y)\,dx\,dy = 1$.)

i) $f(x, y) = 4x^2y^3$.

ii) $f(x, y) = \frac{1}{2}(x^3y + xy^3)$.

iii) $f(x, y) = 6e^{-3x-2y}$

Put a 1 for independent and a 0 for not-independent.

(a) 111 (b) 110 (c) 101 (d) 100

(e) 011 (f) 010 (g) 001 (h) 000

answer: *(c). Explanation on next slide.*

Solution

(i) Independent. The variables can be separated: the marginal densities are $f_X(x) = ax^2$ and $f_Y(y) = by^3$ for some constants a and b with $ab = 4$.

(ii) Not independent. X and Y are not independent because there is no way to factor $f(x, y)$ into a product $f_X(x)f_Y(y)$.

(iii) Independent. The variables can be separated: the marginal densities are $f_X(x) = a\mathrm{e}^{-3x}$ and $f_Y(y) = b\mathrm{e}^{-2y}$ for some constants a and b with $ab = 6$.

Covariance

Measures the degree to which two random variables vary together, e.g. height and weight of people.

X, Y random variables with means μ_X and μ_Y

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

Properties of covariance

Properties

1. $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$ for constants a, b, c, d.
2. $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$.
3. $\text{Cov}(X, X) = \text{Var}(X)$
4. $\text{Cov}(X, Y) = E(XY) - \mu_X \mu_Y$.
5. If X and Y are independent then $\text{Cov}(X, Y) = 0$.
6. **Warning:** The converse is not true, if covariance is 0 the variables might not be independent.

Concept question

Suppose we have the following joint probability table.

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

We compute the covariance using the following steps.

1. $E(X) = 0 \quad E(Y) = 1/2$
2. $E(XY) = -1/4 + 1/4 = 0$
3. Therefore $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$.

Because the covariance is 0 we know that X and Y are independent

1. True 2. False

Concept question

Suppose we have the following joint probability table.

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

We compute the covariance using the following steps.

1. $E(X) = 0 \quad E(Y) = 1/2$
2. $E(XY) = -1/4 + 1/4 = 0$
3. Therefore $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$.

Because the covariance is 0 we know that X and Y are independent

1. True 2. False

Key point: covariance measures the linear relationship between X and Y. It can completely miss a quadratic or higher order relationship.

Board question: computing covariance

Flip a fair coin 3 times.

Let X = number of heads in the first 2 flips

Let Y = number of heads on the last 2 flips.

Compute $\text{Cov}(X, Y)$,

Solution

$X\backslash Y$	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

From the marginals compute $E(X) = 1 = E(Y)$. By the table compute

$$E(XY) = 1 \cdot \frac{2}{8} + 2\frac{1}{8} + 2\frac{1}{8} + 4\frac{1}{8} = \frac{5}{4}.$$

So $\text{Cov}(X, Y) = \frac{5}{4} - 1 = \boxed{\frac{1}{4}}$.

A more conceptual solution is on the next slide.

Alternative Solution

Use the properties of covariance.
X_i = the number of heads on the i^{th} flip. (So $X_i \sim \text{Bernoulli}(.5)$.)

$$X = X_1 + X_2 \quad \text{and} \quad Y = X_2 + X_3.$$

Know $E(X_i) = 1/2$ and $\text{Var}(X_i) = 1/4$. Therefore $\mu_X = 1 = \mu_Y$.
Use Property 2 (linearity) of covariance

$$\begin{aligned} \text{Cov}(X, Y) &= \text{Cov}(X_1 + X_2, X_2 + X_3) \\ &= \text{Cov}(X_1, X_2) + \text{Cov}(X_1, X_3) + \text{Cov}(X_2, X_2) + \text{Cov}(X_2, X_3). \end{aligned}$$

Since the different tosses are independent we know

$$\text{Cov}(X_1, X_2) = \text{Cov}(X_1, X_3) = \text{Cov}(X_2, X_3) = 0.$$

Looking at the expression for $\text{Cov}(X, Y)$ there is only one non-zero term

$$\text{Cov}(X, Y) = \text{Cov}(X_2, X_2) = \text{Var}(X_2) = \boxed{\frac{1}{4}}.$$

Stop and rest

We'll stop here and finish the remainder of these slides in the next class.

Correlation

Like covariance, but removes scale.
The *correlation coefficient* between X and Y is defined by

$$\text{Cor}(X, Y) = \rho = \frac{\text{Cov}(X, Y)}{\sigma_X \, \sigma_Y}.$$

Properties:
1. ρ is the covariance of the standardized versions of X and Y.
2. ρ is dimensionless (it's a ratio).
3. $-1 \leq \rho \leq 1$. $\rho = 1$ if and only if $Y = aX + b$ with $a > 0$ and $\rho = -1$ if and only if $Y = aX + b$ with $a < 0$.

Real-life correlations

- Over time, amount of Ice cream consumption is correlated with number of pool drownings.
- In 1685 (and today) being a student is the most dangerous profession.
- In 90% of bar fights ending in a death the person who started the fight died.
- Hormone replacement therapy (HRT) is correlated with a lower rate of coronary heart disease (CHD).

Discussion is on the next slides.

Real-life correlations discussion

- Ice cream does not cause drownings. Both are correlated with summer weather.

- In a study in 1685 of the ages and professions of deceased men, it was found that the profession with the lowest average age of death was "student." But, being a student does not cause you to die at an early age. Being a student means you *are* young. This is what makes the average of those that die so low.

- A study of fights in bars in which someone was killed found that, in 90% of the cases, the person who started the fight was the one who died.

  Of course, it's the person who survived telling the story.

Continued on next slide

(continued)

- In a widely studied example, numerous epidemiological studies showed that women who were taking combined hormone replacement therapy (HRT) also had a lower-than-average incidence of coronary heart disease (CHD), leading doctors to propose that HRT was protective against CHD. But randomized controlled trials showed that HRT caused a small but statistically significant increase in risk of CHD. Re-analysis of the data from the epidemiological studies showed that women undertaking HRT were more likely to be from higher socio-economic groups (ABC1), with better-than-average diet and exercise regimens. The use of HRT and decreased incidence of coronary heart disease were coincident effects of a common cause (i.e. the benefits associated with a higher socioeconomic status), rather than cause and effect, as had been supposed.

Correlation is not causation

Edward Tufte: "Empirically observed covariation is a necessary but not sufficient condition for causality."

Overlapping sums of uniform random variables

We made two random variables X and Y from overlapping sums of uniform random variables

For example:

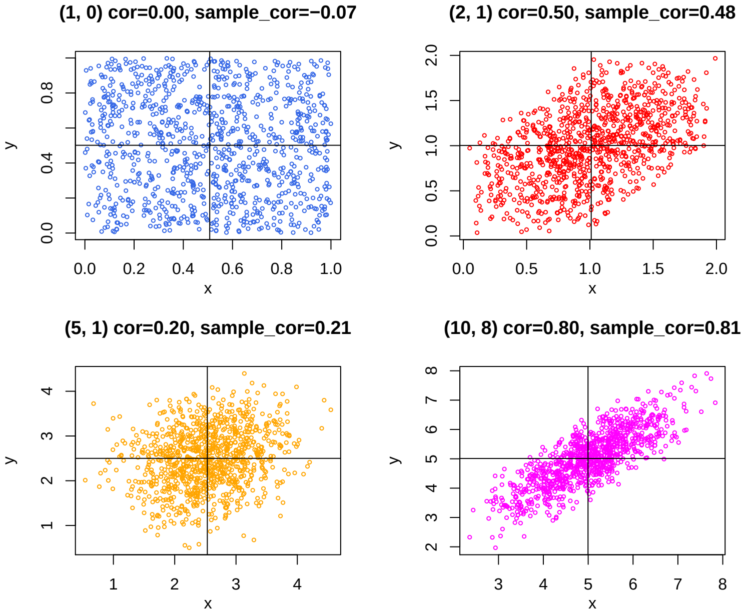
$$X = X_1 + X_2 + X_3 + X_4 + X_5$$
$$Y = X_3 + X_4 + X_5 + X_6 + X_7$$

These are sums of 5 of the X_i with 3 in common.

If we sum r of the X_i with s in common we name it (r, s).

Below are a series of scatterplots produced using R.

Scatter plots



Concept question

Toss a fair coin $2n + 1$ times. Let X be the number of heads on the first $n + 1$ tosses and Y the number on the last $n + 1$ tosses.

If $n = 1000$ then $\text{Cov}(X, Y)$ is:

(a) 0 (b) $1/4$ (c) $1/2$ (d) 1

(e) More than 1 (f) tiny but not 0

answer: 2. $1/4$. This is computed in the answer to the next table question.

Board question

Toss a fair coin $2n+1$ times. Let X be the number of heads on the first $n+1$ tosses and Y the number on the last $n+1$ tosses.

Compute $\text{Cov}(X, Y)$ and $\text{Cor}(X, Y)$.

As usual let X_i = the number of heads on the i^{th} flip, i.e. 0 or 1. Then

$$X = \sum_{1}^{n+1} X_i, \qquad Y = \sum_{n+1}^{2n+1} X_i$$

X is the sum of $n+1$ independent Bernoulli(1/2) random variables, so

$$\mu_X = E(X) = \frac{n+1}{2}, \quad \text{and} \quad \text{Var}(X) = \frac{n+1}{4}.$$

Likewise, $\mu_Y = E(Y) = \dfrac{n+1}{2}$, and $\text{Var}(Y) = \dfrac{n+1}{4}$.

Continued on next slide.

Solution continued

Now,

$$\text{Cov}(X,Y) = \text{Cov}\left(\sum_{1}^{n+1} X_i \sum_{n+1}^{2n+1} X_j\right) = \sum_{i=1}^{n+1}\sum_{j=n+1}^{2n+1} \text{Cov}(X_i X_j).$$

Because the X_i are independent the only non-zero term in the above sum is $\text{Cov}(X_{n+1}X_{n+1}) = \text{Var}(X_{n+1}) = \frac{1}{4}$ Therefore,

$$\text{Cov}(X,Y) = \frac{1}{4}.$$

We get the correlation by dividing by the standard deviations.

$$\text{Cor}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{1/4}{(n+1)/4} = \frac{1}{n+1}.$$

This makes sense: as n increases the correlation should decrease since the contribution of the one flip they have in common becomes less important.