Joint Distributions, Independence
Covariance and Correlation
18.05 Spring 2014

$X\backslash Y$	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

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

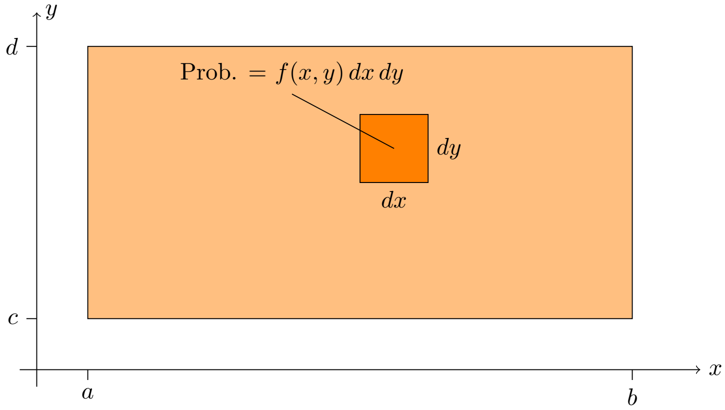
Roll two dice: $X = \#$ on first die, $T =$ total on both dice

$X \backslash T$	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 $[a, b]$, $\qquad$ 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. $0 \leq p(x_i, y_j) \leq 1$
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. $0 \leq f(x, y)$
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.

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 $= 10/36$.

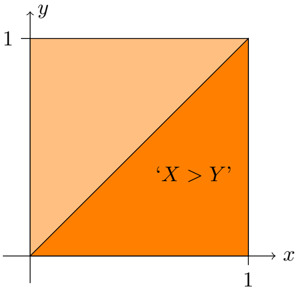
Example: continuous events

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.

answer:



The event takes up half the square. Since the density is uniform this is half the probability. That is, $P(X > Y) = 0.5$

Cumulative distribution function

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

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

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 and pdf

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

For continuous distributions the marginal pdf $f_X(x)$ is found by integrating out the y. Likewise for $f_Y(y)$.

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 $A =$ '$X > 0.3$ and $Y > 0.5$'. Find its probability.
3. Find the cdf $F(x, y)$.
4. Find the marginal pdf $f_X(x)$. Use this to find $P(X < 0.5)$.
5. Use the cdf $F(x, y)$ to find the marginal cdf $F_X(x)$ and $P(X < 0.5)$.
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

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

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

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

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

At your table work out the covariance $\text{Cov}(X, Y)$.

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 12 times.

Let $X =$ number of heads in the first 7 flips

Let $Y =$ number of heads on the last 7 flips.

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

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).

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:

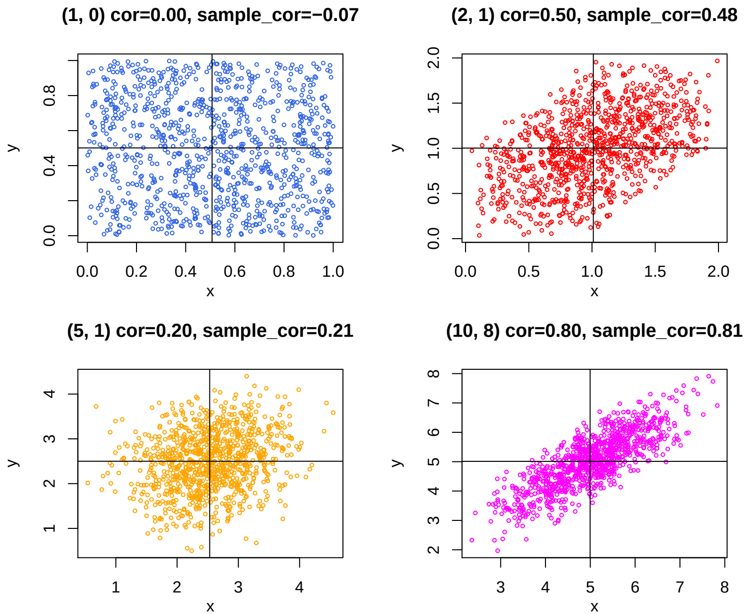
$$X = X_1 + X_2 + X_3 + X_4 + X_5$$
$$Y = X_3 + X_4 + X_5 + X_6 + X_7$$

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

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