Joint Distributions, Independence
Covariance and Correlation
18.05 Spring 2022

$X\backslash Y$	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

Announcements/Agenda

Announcements

- Exam 1 on Thursday March 10. Covers classes 1-7.
- Designed for 1 hour. You will have the full 80 minutes.
- Review materials will be posted tomorrow.
- Class on Tuesday 3/8 will be mostly review.
- For the exam you will be given a table of standard normal probabilities.
- You can bring in a cheat sheet: 1 side of an 8×11 sheet of paper. You'll turn it in with the exam for 5 points.

Agenda

- Joint distributions: pmf, pdf, cdf
- Marginal distributions
- Independence
- Covariance and correlation

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, $T =$ sum of both dice

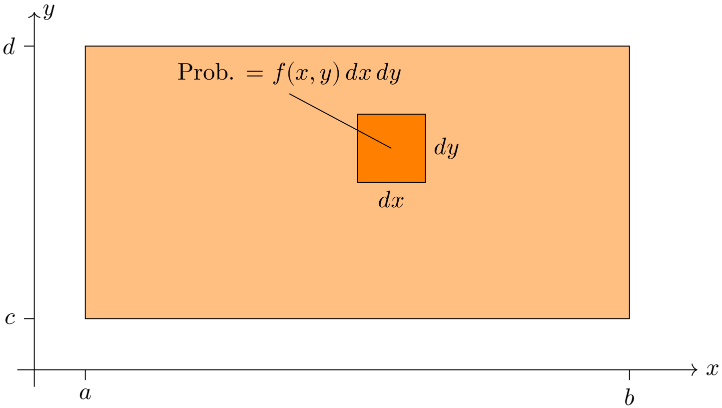
$X \backslash T$	2	3	4	5	6	7	8	9	10	11	12
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36

E.g. $p(4, 2) = 0$, $p(1, 7) = 1/36$.

Continuous joint distributions

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

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



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.

18.02 in 18.05

- You should understand double integrals conceptually as double sums.
- You should be able to compute double integrals over rectangles.
- For a non-rectangular region, when $f(x, y) = c$ is constant, you should know that the double integral is the same as $c\times$(the area of the region).
- You should be able to compute partial derivatives.

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.

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.

Solution: 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 $f(x, y) = 1$.

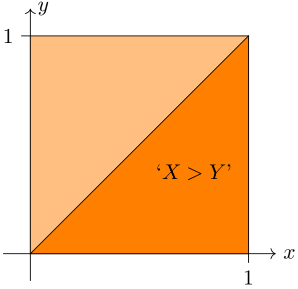
Visualize the event '$X > Y$' and find its probability.

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.
Solution:



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

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

(a and c are the bottom of the ranges of X and Y respectively.)

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

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.
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 pmf of X is found by summing the rows. The pmf of T is found by summing the columns. These are called marginal pmfs of the joint distribution.

$X \backslash T$	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: Joint distributions

Suppose X and Y are random variables and

- (X, Y) takes values in $[0, 1] \times [0, 1]$.
- the pdf is $f(x, y) = x + y$.

(a) Show $f(x, y)$ is a valid pdf.

(b) Visualize the event $A =$ '$X > 0.3$ and $Y > 0.5$'. Find its probability.

(c) Find the cdf $F(x, y)$.

(d) Use the cdf $F(x, y)$ to find the marginal cdf $F_X(x)$ and $P(X < 0.5)$.

(e) Find the marginal pdf $f_X(x)$. Use this to find $P(X < 0.5)$.

(f) See next slide

Board question: continued

(f) (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

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

Independence means probabilities multiply!

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: Independence III

Which of the following joint pdfs are the variables 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) $f(x,y) = 6e^{-3x-2y}$

(a) i **(b)** ii **(c)** iii **(d)** i, ii

(e) i, iii **(f)** ii, iii **(g)** i, ii, iii **(h)** None

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.

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

- Positive covariance: When X is bigger than μ_X then Y is 'usually' bigger than μ_Y, and vice versa
- Negative covariance: When X is bigger than μ_X then Y is 'usually' smaller than μ_Y, and vice versa
- Zero covariance: The sign of $X - \mu_X$ tells us nothing about the sign of $(Y - \mu_Y)$.

Properties of covariance

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

Table 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 Cov(X, Y).

Are X and Y independent?

Table 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 Cov(X, Y).

Are X and Y independent?

Covariance = 0. Not independent!

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

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. $-1 \le \rho \le 1$.

4. $\rho = 1$ if and only if $Y = aX + b$ with $a > 0$

5. $\rho = -1$ if and only if $Y = aX + b$ with $a < 0$.

Board question: Covariance and correlation

Flip a fair coin 11 times. (The tosses are all independent.)

Let $X =$ number of heads in the first 6 flips

Let $Y =$ number of heads on the last 6 flips.

Compute Cov(X, Y) and Cor(X, Y).

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. That is, the average age of those who die is less than any other 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:

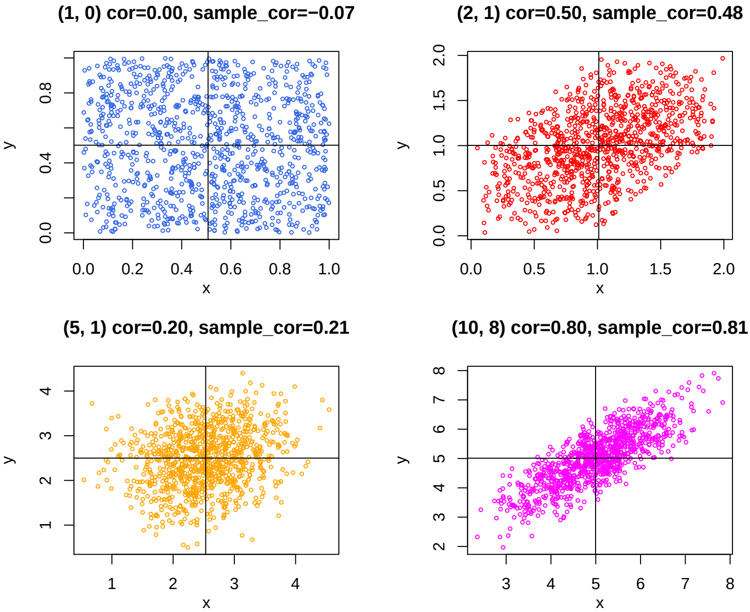
$$X = X_1 + X_2 + X_3 + X_4 + X_5$$
$$Y = X_3 + X_4 + X_5 + X_6 + X_7$$

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



Intuition check

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 Cov(X, Y) is:

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

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

(This is computed in the answer to the next board question.)

Board question: Even more tosses

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