Random vectors - cont'd

Probability and Statistics
(EPRST)

Lecture 9

Continuous random vectors

Definition

*A random vector $\boldsymbol{X} = (X, Y)$ has a **continuous joint distribution**, if there exists a function $f : \mathbb{R}^2 \to [0, \infty)$ such that*

$$\mathbb{P}(\boldsymbol{X} \in B) = \mathbb{P}((X, Y) \in B) = \iint_B f(x, y) dx dy$$

*for $B \subset \mathbb{R}^2$. Function f is called **joint density**.*

As in dimension one, function $f : \mathbb{R}^2 \to \mathbb{R}$ is joint density of a certain random vector, if f takes non-negative values and

$$\iint_{\mathbb{R}^2} f(x, y) \mathrm{d}x \mathrm{d}y = 1.$$

Joint density - an example

Example

Set $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = \begin{cases} cxy, & (x,y) \in [0,1]^2, \\ 0, & (x,y) \notin [0,1]^2. \end{cases}$$

- *For what values of* c *is this function a two-dimensional density?*
- *If* f *is a joint density of* $\boldsymbol{X} = (X, Y)$*, what is* $\mathbb{P}(\boldsymbol{X} \in [0, 1/2]^2)$*?*

Marginal distributions

Regardless of whether a random vector $\boldsymbol{X} = (X, Y)$ is discrete or continuous, distributions of random variables X and Y can be derived from its joint distribution, because for $A, B \subset \mathbb{R}$

$$\begin{aligned} \mathbb{P}(X \in A) &= \mathbb{P}(\boldsymbol{X} = (X, Y) \in A \times \mathbb{R}), \\ \mathbb{P}(Y \in B) &= \mathbb{P}(\boldsymbol{X} = (X, Y) \in \mathbb{R} \times B). \end{aligned}$$

The distributions of (one-dimensional) random variables X and Y are called **marginal distributions of random vector $\boldsymbol{X}$**.

How to determine marginal distributions of discrete vectors?

Suppose X, Y are (one-dimensional) discrete random variables taking values in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Given the joint distribution of the discrete random vector $\boldsymbol{X} = (X, Y)$, that is, probabilities

$$\mathbb{P}(X = x, Y = y)$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$, one determines the marginal distributions from

$$\mathbb{P}(X = x) = \sum_{y \in \mathcal{Y}} \mathbb{P}(X = x, Y = y) \ \forall x$$

and

$$\mathbb{P}(Y = y) = \sum_{x \in \mathcal{X}} \mathbb{P}(X = x, Y = y) \ \forall y.$$

If $\boldsymbol{X}$ has a discrete joint distribution, then its marginal distributions are discrete as well.

Marginal distributions of a discrete random vector - an example

Example

Find the marginal distributions for random vector $\boldsymbol{X} = (X, Y)$ with the joint distribution given by the table:

-

$X \backslash Y$	0	1
0	1/4	1/4
1	1/4	1/4

-

$X \backslash Y$	0	1
0	1/8	3/8
1	3/8	1/8

Joint vs marginal distributions

The above example illustrates an important fact (which is true regardless of distribution types):

- knowledge of the joint distribution allows to determine the marginal distributions,
- in general (without some additional assumptions), knowledge of the marginal distributions does not allow to determine the joint distribution.

The joint distribution contains more information than both marginal distributions together.

How to determine marginal distributions of continuous vectors?

If function f (defined on $\mathbb{R}^2$) is the joint density of random vector $\boldsymbol{X} = (X, Y)$, then the marginal distributions of $\boldsymbol{X}$ are continuous, and the marginal densities can be computed from the following formulas

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y)\mathrm{d}y, \\ f_Y(y) &= \int_{-\infty}^{\infty} f(x, y)\mathrm{d}x. \end{aligned}$$

Example

Compute the marginal densities of random vector $\boldsymbol{X} = (X, Y)$, uniformly distributed on a set D, $\boldsymbol{X} \sim \mathcal{U}(D)$, where

- $D = [0, 1]^2$,
- $D = \{(x, y) \in \mathbb{R}^2 : x, y \geq 0, x + y \leq 1\}$.

Independence of random variables

Recall - random events A, B, defined on the common sample space, are **independent**, if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Extending this definition to random variables is simple:

Definition

*Random variables X and Y (defined on the common sample space, with the values from the sets $\mathcal{X}$ and $\mathcal{Y}$, respectively) are **independent**, if any two random events of the form*

$$\{X \in A\} \quad and \quad \{Y \in B\}$$

($A \subset \mathcal{X}$, $B \subset \mathcal{Y}$) are independent (as random events), so if

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B)$$

for all $A \subset \mathcal{X}$, $B \subset \mathcal{Y}$.

Independence of discrete r.v.s

If the joint distribution of $\boldsymbol{X} = (X, Y)$ is discrete, it is more convenient to check independence from the following condition (equivalent with the definition)

Theorem

R.v.s X and Y, taking values in $\mathcal{X}$ and $\mathcal{Y}$, are independent iff

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Example

We flip two symmetric coins (heads $\to$ 1, tails $\to$ 0). Let X denote sum of the outcomes , and Y - the absolute value of their difference. Are X and Y independent?

Independence of continuous r.v.s

Theorem

If the joint distibution of X and Y is continuous, with density $f_{X,Y}$, then X and Y are independent iff

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y).$$

Example

Random vector $\boldsymbol{X} = (X, Y)$ is distributed uniformly over $D \subset \mathbb{R}^2$. Are X and Y independent if

- $D = [0, 1]^2$,
- $D = \{(x, y) \in \mathbb{R}^2 : x, y \geq 0, x + y \leq 1\}$.

Expectation of a function of a random vector

If $h : \mathbb{R}^2 \to \mathbb{R}$, and (X, Y) is a random vector, then the expectation $\mathbb{E}h(X, Y)$ can be computed directly from the formula

$$\mathbb{E}h(X, Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} h(x, y)\mathbb{P}(X = x, Y = y)$$

when X and Y are discrete, or from the formula

$$\mathbb{E}h(X, Y) = \iint_{\mathbb{R}^2} h(x, y) f(x, y) \mathrm{d}x\mathrm{d}y,$$

when X and Y are continuous and f is the joint density of (X, Y).

Example

Compute $\mathbb{E}XY$ if (X, Y) is distributed uniformly over the unit disk.

Expectation of the product of r.v.s

Theorem

If X and Y are independent then

$$\mathbb{E}(XY) = \mathbb{E}X \cdot \mathbb{E}Y.$$

Variance of the sum of independent r.v.s

Theorem

If X and Y are independent (and $\mathbb{E}X^2 < \infty$, $\mathbb{E}Y^2 < \infty$), then

$$\operatorname{Var}(X+Y) = \operatorname{Var} X + \operatorname{Var} Y.$$

Variance of binomial r.v.s

Let X denote the number of successes in n Bernoulli trials, so

$$X \sim \mathrm{bin}(n, p).$$

If

$$X_i = \begin{cases} 1, & \text{if a success at } i\text{-th trial}, \\ 0, & \text{if a failure at } i\text{-th trial}, \end{cases} \qquad i = 1, \ldots, n,$$

then $X = X_1 + \ldots + X_n$. Since Bernoulli trials are independent, so are r.v.s $X_1, \ldots, X_n$. Therefore

$$\operatorname{Var} X = \operatorname{Var}(X_1 + \ldots + X_n) = \operatorname{Var} X_1 + \ldots + \operatorname{Var} X_n.$$

Since

$$\operatorname{Var} X_i = p(1 - p),$$

we conclude that

$$\operatorname{Var} X = np(1 - p).$$

Covariance

Definition

Covariance between r.v.s X ***and*** Y *is*

$$\mathrm{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\right].$$

Covariance is defined for r.v.s satisfying $\mathbb{E}X^2 < \infty$ and $\mathbb{E}Y^2 < \infty$.

Equivalent formula for covariance:

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y).$$

Definition

R.v.s X *and* Y *with* $\mathrm{Cov}(X, Y) = 0$, *are* ***uncorrelated***.

Covariance - some properties

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- $\mathrm{Cov}(X, X) = \mathrm{Var}\, X$
- $\mathrm{Cov}(aX, bY) = ab\,\mathrm{Cov}(X, Y)$
- $\mathrm{Cov}(X, c) = \mathrm{Cov}(c, X) = 0$
- $$\begin{aligned}\mathrm{Cov}(X_1 + X_2, Y_1 + Y_2) = \mathrm{Cov}(X_1, Y_1) + \mathrm{Cov}(X_1, Y_2) + \\ + \mathrm{Cov}(X_2, Y_1) + \mathrm{Cov}(X_2, Y_2)\end{aligned}$$
- **Cauchy-Schwarz inequality**:

$$|\mathrm{Cov}(X, Y)| \leq \sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}$$

Variance of $X + Y$ - general formula

Using some properties of covariance, we can easily derive an important formula for the variance of the sum of random variables in the general case (that is, *without assuming independence*).

Theorem

If X and Y are r.v.s (for which the variances exist) then

$$\mathrm{Var}(X + Y) = \mathrm{Var}\,X + 2\,\mathrm{Cov}(X, Y) + \mathrm{Var}(Y).$$

Correlation coefficient

Definition

*The **correlation** between r.v.s X and Y us*

$$\rho_{X,Y} := \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var} X}\sqrt{\operatorname{Var} Y}}.$$

(This makes sense only if the covariance makes sense and is undefined in the degenerate cases $\operatorname{Var} X = 0$ *or* $\operatorname{Var} Y = 0$*.)*

Some properties:

- $-1 \le \rho_{X,Y} \le 1$ (correlation, unlike covariance, is bounded),
- $\rho_{X,Y} = 0$ iff X and Y are uncorrelated.

Covariance/correlation - some examples

Example

Compute $\mathrm{Cov}(X, Y)$ *and* $\rho_{X,Y}$ *when*

- (X, Y) *has the joint distribution given by the table*

$X \backslash Y$	*0*	*1*
0	$1/4 + \varepsilon$	$1/4 - \varepsilon$
1	$1/4 - \varepsilon$	$1/4 + \varepsilon$

- (X, Y) *is uniformly distributed over the unit disk,*
- $(X, Y) \sim U([0, 1]^2)$.

Independence vs uncorrelatedness

Recall:

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y).$$

On the other hand, if X and Y are independent, then $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}Y$.

Corollary

If X and Y are independent (and $\mathbb{E}X^2 < \infty$, $\mathbb{E}Y^2 < \infty$), then X and Y are uncorrelated (so $\mathrm{Cov}(X, Y) = 0$, equivalently $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$).

Corollary

If X and Y are uncorrelated, then $\mathrm{Var}(X + Y) = \mathrm{Var}\,X + \mathrm{Var}\,Y$.

Independence vs uncorrelatedness - cont'd

So

if X and Y are independent (and their covariance is well defined), then they are uncorrelated.

The converse is false:

if X and Y are uncorrelated then they are *not* necessarily independent.

Covariance matrix

Definition

If (X, Y) is a random vector with $\mathbb{E}X^2 < \infty$ and $\mathbb{E}Y^2 < \infty$, then the ***covariance matrix*** *of (X, Y) is*

$$\boldsymbol{C}_{(X,Y)} := \begin{bmatrix} \operatorname{Cov}(X, X) & \operatorname{Cov}(X, Y) \\ \operatorname{Cov}(Y, X) & \operatorname{Cov}(Y, Y) \end{bmatrix} = \begin{bmatrix} \operatorname{Var} X & \operatorname{Cov}(X, Y) \\ \operatorname{Cov}(X, Y) & \operatorname{Var} Y \end{bmatrix}.$$

A straightforward generalization to n-dimensions:

Definition

If $\boldsymbol{X} = (X_1, \ldots, X_n)$ is a random vector with $\mathbb{E}X_i^2 < \infty$ for all $i = 1, \ldots, n$ then the ***covariance matrix*** *of vector $\boldsymbol{X}$ is the $n \times n$ matrix*

$$\boldsymbol{C}_{\boldsymbol{X}} = \left[\operatorname{Cov}(X_i, X_j)\right]_{i,j=1}^{n}.$$

Covariance matrix - some properties

- covariance matrices are symmetric and non-negative definite,
- the covariance matrix of random vector $\mathbf{AX} + \mathbf{b}$ is

$$\mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}^T.$$

($\mathbf{A}$ is a matrix of numbers, $\mathbf{b}$ is a vector of numbers).