EE2S31 Signal Processing – Stochastic Processes

Lecture 1: Multiple random variables (Ch.5)

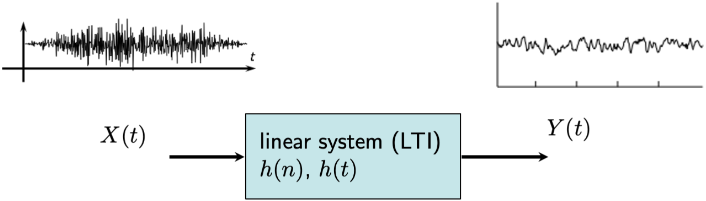
Alle-Jan van der Veen

–

20 Apr 2022

In this course

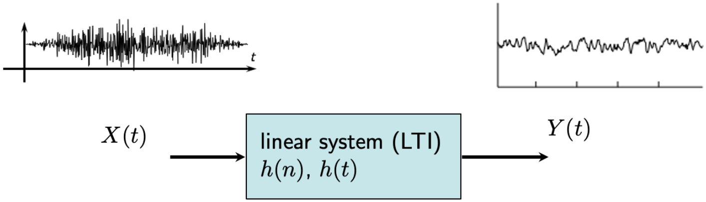
We will look at "random signals" (communication signals, speech, ECG, EEG):



For deterministic processes:

- Given $h(t)$ and $X(t)$ we know $Y(t)$: $Y(t) = h(t) * X(t)$
- What can we say about $Y(t)$ when $X(t)$ is not deterministic, but a stochastic process?
- What *is* a stochastic process?

Message of this course



Statistical descriptions of $X(t)$:

- mean μ_X
- Autocorrelation function $R_X(\tau)$.

Statistical descriptions of $Y(t)$:

- mean $\mu_Y = \mu_X \int_t h(t)dt$
- $R_Y(\tau) = h(\tau) * h(-\tau) * R_X(\tau)$.

Message of this course

Some questions to answer:

- What is a stochastic process?
- How to characterize a stochastic process?

- What is the autocorrelation function?
- Statistical descriptions of output $Y(t)$ of LTI system with WSS input $X(t)$:
  - mean $\mu_Y = \mu_X \int_t h(t)dt$
  - $R_Y(\tau) = h(\tau) * h(-\tau) * R_X(\tau)$.

- Frequency domain: power spectral density

In this course [Stochastic Processes]

Multiple random variables:

- Pairs of RVs: Ch. 5
- Random vectors: Ch. 8
- Conditional probability: Ch. 7
- Sums of RVs, moment generating function, bounds : Chs. 6, 9, 10
- Estimation of random variables (MMSE, MAP, ML): Ch. 12

Random processes:

- Stochastic processes: Ch. 13
- Auto-correlation function, power spectral density: Supplement

Prior knowledge: EE1M31 Probability and Statistics

- Probability spaces, conditional probability, Bayes' rule
  - Discrete and continuous random variables; PDF; CDF
  - Transformations of random variables
- Central Limit Theorem
- Expectation and (co)variance
- Parameter estimates
  - Maximum likelihood, linear regression
- Hypothesis tests

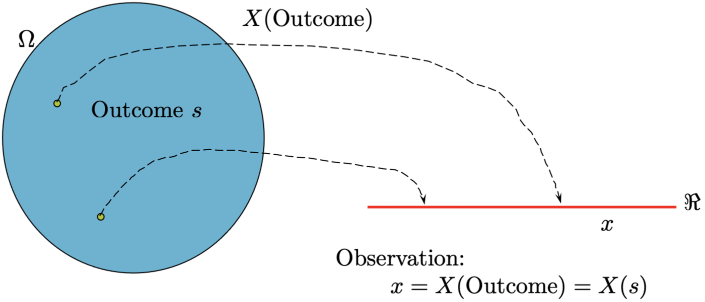
Recap EE2ML1: Definition of a Random Variable (1)

- Experiments often take place in a physical world: We observe physical quantities.
- Probability theory works in a mathematical world, with mathematical tools.
- How to map "physical observations" to numbers we can do mathematics on?
  - E.g. Flip a coin: Outcomes are head/tail
  - E.g. Observe flooding because of frequent rainfall. Outcomes are yes/no

Definition of a Random Variable (2)

Outcome space: Ω
Outcome: elements of Ω
Event: subset of Ω

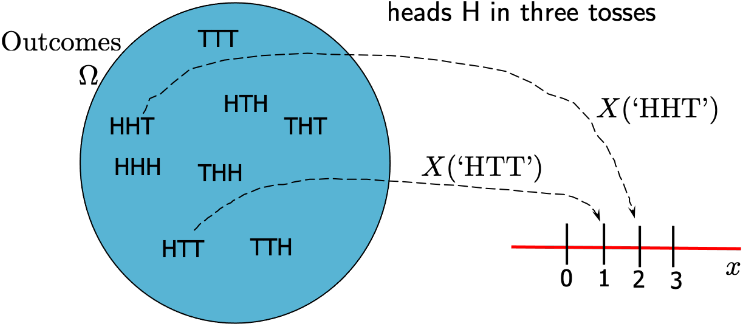
Random variable: function of outcome, maps to real number (enabling ordering!)



Definition of a Random Variable (3)

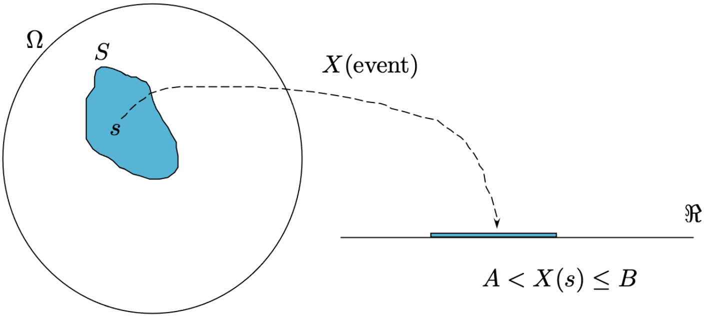
Experiment: Head or tail in 3 tosses

Random variable $X(\text{outcome})$: number of heads H in three tosses



Random variable is simply a rule to assign a real number to each outcome

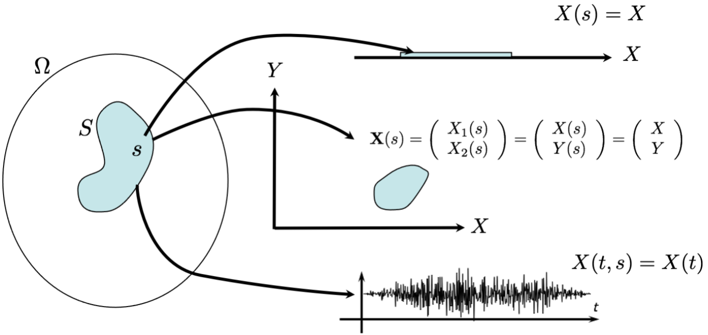
Definition of a Random Variable (4)



Event: set S of outcomes

- Roll even; grown-up
- X(event): $\{2, 4, 6\}$; $[18, \infty)$

Later: extend to stochastic processes



One experiment $\Rightarrow$ one **realization** (observation) of the stochastic/random process

Scalar random variables

A random variable is a real number that represents the outcome of an experiment.

- **Discrete random variable:** countable number of outcomes; described by a **probability mass function** (PMF)

$$P_X(x) = \mathrm{P}[X = x]$$

- **Continuous random variable:** continuous sample space; described by a **probability density function** (PDF) $f_X(x)$.

A random variable is described by a **cumulative distribution function** (CDF) $F_X(x)$

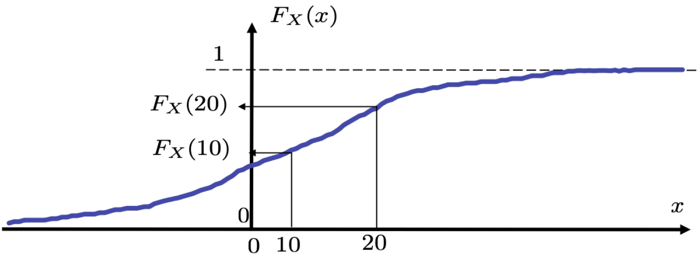
$$F_X(x) = \mathrm{P}[X \leq x]$$

Properties of the CDF

$F_X(x) = \mathrm{P}[X \leq x]$

Properties

- $F_X(-\infty) = 0, \quad F_X(\infty) = 1$
- $\mathrm{P}[x_1 < X \leq x_2] = F_X(x_2) - F_X(x_1)$
- For all $x' \geq x$: $F_X(x') \geq F_X(x)$



Probability mass function (PMF)

Discrete RV's are usually described by their **Probability Mass Function** (PMF),

$$P_X(x) = \mathrm{P}[X = x]$$

Properties

For a discrete random variable X with PMF $P_X(x)$ and range S_X:

- $P_X(x) \geq 0, \qquad \forall x$
- $\sum_{x \in S_X} P_X(x) = 1$
- For an event $B \subset S_X$, the probability that X is in the set B is

$$\mathrm{P}[B] = \sum_{x \in B} P_X(x)$$

- Hence: $F_X(x) = \mathrm{P}[X \leq x] = \sum_{u \leq x} P_X(u)$

Example Scalar RV: from PMF to CDF

Y	$y = 0$	$y = 1$	$y = 2$
$P_Y(y)$	0.1	0.09	0.81

- Plot the CDF $F_Y(y)$.
- What is $F_Y(1)$?

Probability density function (PDF)

Continuous RV's are usually described by their **Probability Density Function** (PDF)

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$$

Note: it can be larger than 1! ($f_X(x)$ is *not* a probability).

Properties

- $f_X(x) \geq 0\,, \quad \forall x$
- $F_X(x) = \int_{-\infty}^{x} f_X(x)\mathrm{d}x$
- $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$

Example:
exponential distribution

$$f_X(x) = \lambda e^{-\lambda x}$$
$$F_X(x) = \int_0^x \lambda e^{-\lambda u}\mathrm{d}u = 1 - e^{-\lambda x}$$

For discrete RV's the PDF also exists, using delta functions

Expected value; mean and variance

The expected value of a continuous RV is

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

More general, if $g(x)$ is a function, then the expected value of $g(X)$ is

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x) dx$$

Mean value

$$\mu_X = E[X]$$

Variance

$$\sigma_X^2 = \mathrm{var}[X] = E[(X - E[X])^2] \;=\; E[X^2] - (E[X])^2$$

Example Scalar RV: from PDF to CDF

Given is the exponential PDF

$$f_T(t) = \begin{cases} \frac{1}{3}e^{-t/3} & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Calculate the CDF $F_T(t)$
- Calculate the expected value $\mathrm{E}[T]$

Example Scalar RV: from PDF to CDF

The exponential PDF is

$$f_T(t) = \begin{cases} \frac{1}{3} e^{-t/3} & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

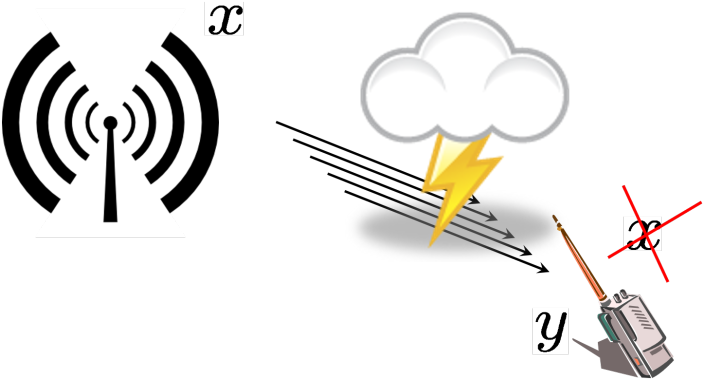
Its CDF is

$$F_T(t) = \int_{-\infty}^{t} f_T(u)\, \mathrm{d}u = \int_0^t \frac{1}{3} e^{-u/3} \mathrm{d}u = [-e^{-u/3}]_0^t = 1 - e^{-t/3}$$

The expected value of T is

$$\begin{aligned} \mathrm{E}[T] &= \int_0^\infty \frac{t}{3} e^{-t/3} \mathrm{d}t = \int_0^\infty -t\, \mathrm{d}e^{-t/3} \\ &= \left[-te^{-t/3}\right]_0^\infty + \int_0^\infty e^{-t/3} \mathrm{d}t = 0 + \left[-3e^{-t/3}\right]_0^\infty = 3 \end{aligned}$$

Ch.5: From Scalar RVs to Multiple RVs



Model: $Y = X + N$

How to model this scenario?

$\Rightarrow$ Using a pair of random variables: (X, Y) or (X, N)

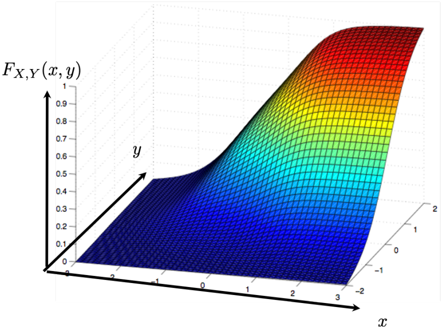
How to describe multiple RVs?

For multiple RVs we define the **joint cumulative distribution function** (j-CDF) as

$$F_{X,Y}(x,y) = \mathrm{P}[X \leq x, Y \leq y]$$

- **Joint PMF** (for discrete RV): $P_{X,Y}(x,y) = \mathrm{P}[X = x, Y = y]$
- **Joint PDF** (for continuous RV): $f_{X,Y}(x,y) = \dfrac{\partial^2 F_{X,Y}(x,y)}{\partial x \, \partial y}$

The joint cumulative distribution function



Properties

- $0 \leq F_{X,Y}(x,y) \leq 1$
- $F_{X,Y}(x,-\infty) = 0, \quad F_{X,Y}(-\infty,y) = 0, \quad F_{X,Y}(\infty,\infty) = 1$
- If $x \leq x_1$ and $y \leq y_1$ then $F_{X,Y}(x,y) \leq F_{X,Y}(x_1,y_1)$
- **Marginals:** $F_X(x) = F_{X,Y}(x,\infty), \quad F_Y(y) = F_{X,Y}(\infty,y)$

Example: joint CDF

Given

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 5 \\ 0 & y < 6 \\ (x-5)(y-6) & 5 \leq x < 6, 6 \leq y < 7 \\ y-6 & x \geq 6, 6 \leq y < 7 \\ x-5 & 5 \leq x < 6, y \geq 7 \\ 1 & \text{otherwise} \end{cases}$$

Give the marginal CDFs $F_X(x)$ and $F_Y(y)$.

The joint probability mass function

Using the j-CDF can be complex. It is often easier to work with the joint PMF and joint PDF.

The **joint PMF** is:

$$P_{X,Y}(x,y) = \mathrm{P}[X = x, Y = y]$$

Let $S_{X,Y} = \{(x,y) \,|\, P_{X,Y}(x,y) > 0\}$ be the set of possible (X,Y) pairs.

Example

$P_{X,Y}(x,y)$	$y = 0$	$y = 1$	$y = 2$
$x = 0$	0.01	0	0
$x = 1$	0.09	0.09	0
$x = 2$	0	0	0.81

$$\sum_{(x,y)\in S_{X,Y}} P_{X,Y}(x,y) = 1$$

The joint PMF for events

For discrete RVs X and Y and any set B in the X, Y-plane, the probability of the event $\{(X, Y) \in B\}$ is

$$\mathrm{P}[B] = \sum_{(x,y)\in B} P_{X,Y}(x, y).$$

Example

Let B be the event $\{X = Y\}$. Give $\mathrm{P}[B]$ for the joint PMF $P_{X,Y}(x, y)$

$P_{X,Y}(x, y)$	$y = 0$	$y = 1$	$y = 2$
$x = 0$	0.01	0	0
$x = 1$	0.09	0.09	0
$x = 2$	0	0	0.81

The marginal PMF

The marginal PMF $P_Y(y)$ is obtained by summing over all points with the property that $Y = y$:

$$P_Y(y) = \sum_{x \in S_X} P_{X,Y}(x,y), \qquad P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x,y)$$

Example

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$	$P_X(x)$
$x=0$	0.01	0	0	0.01
$x=1$	0.09	0.09	0	0.18
$x=2$	0	0	0.81	0.81
$P_Y(y)$:	0.1	0.09	0.81	

Example: From joint PMF to joint CDF

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81

$$F_{X,Y}(x,y) = P[X \leq x, Y \leq y]$$

$F_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0.01	0.01
$x=1$	0.1	0.19	0.19
$x=2$	0.1	0.19	1

From marginal PMF to joint PMF?

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$P_X(x)$
$x=0$	?	?	0.5
$x=1$	?	?	0.5
$P_Y(y)$	0.5	0.5	

From marginal PMF to joint PMF?

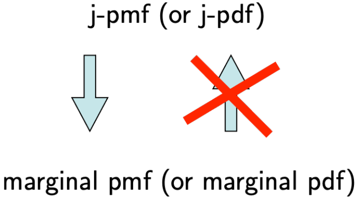
$P_{X,Y}(x,y)$	$y=0$	$y=1$	$P_X(x)$
$x=0$	0.5	0	0.5
$x=1$	0	0.5	0.5
$P_Y(y)$	0.5	0.5	

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$P_X(x)$
$x=0$	0.25	0.25	0.5
$x=1$	0.25	0.25	0.5
$P_Y(y)$	0.5	0.5	

$P_{X,Y}(x,y)$	$y=0$	$y=1$	$P_X(x)$
$x=0$	0.1	0.4	0.5
$x=1$	0.4	0.1	0.5
$P_Y(y)$	0.5	0.5	

From marginal PMF to joint PMF?

Knowing the marginal PMF (or PDF), does NOT mean that you can derive the j-PMF (or j-PDF)!



The joint probability density function

For continuous multiple random variables we use the *joint PDF*

The joint PDF is related to the joint CDF by:

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\mathrm{d}u\mathrm{d}v$$

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

Properties

- $f_{X,Y}(x,y) \geq 0, \quad \forall\,(x,y)$
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 1$

Example j-PDF

Given

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 5 \\ 0 & y < 6 \\ (x-5)(y-6) & 5 \leq x < 6, 6 \leq y < 7 \\ y-6 & x \geq 6, 6 \leq y < 7 \\ x-5 & 5 \leq x < 6, y \geq 7 \\ 1 & \text{otherwise} \end{cases}$$

Give the j-PDF $f_{X,Y}(x,y)$.

Example j-PDF

Given

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 5 \\ 0 & y < 6 \\ (x-5)(y-6) & 5 \leq x < 6, 6 \leq y < 7 \\ y-6 & x \geq 6, 6 \leq y < 7 \\ x-5 & 5 \leq x < 6, y \geq 7 \\ 1 & \text{otherwise} \end{cases}$$

Give the j-PDF $f_{X,Y}(x,y)$.

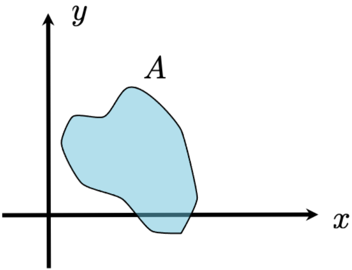
$$\begin{aligned} f_{X,Y}(x,y) &= \begin{cases} \frac{\partial^2}{\partial x \partial y}(x-5)(y-6) & 5 \leq x < 6, 6 \leq y < 7 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & 5 \leq x < 6, 6 \leq y < 7 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

The joint probability density function

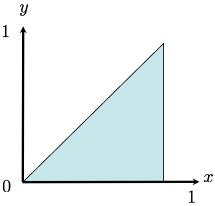
The j-PDF and j-PMF thus simply extend what we already know from scalar RVs.

However: Finding the right integration area for these higher dimensional integrals can be complicated!

$$\mathrm{P}[A] = \int_A \int f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$$



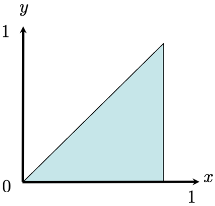
Example



$$f_{X,Y}(x,y) \quad = \quad \begin{cases} c & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Value of c?

Example



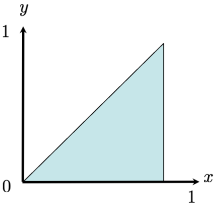
$$f_{X,Y}(x,y) = \begin{cases} c & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Value of c?

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1\int_0^x c\,\mathrm{d}y\,\mathrm{d}x = 1$$

$$\Rightarrow c = 2$$

Example



$$f_{X,Y}(x,y) = \begin{cases} c & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

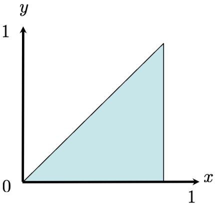
Value of c?

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1\int_0^x c\,\mathrm{d}y\,\mathrm{d}x = 1$$

$$\Rightarrow c = 2$$

$\mathrm{P}[X > 1/2]$?

Example



$$f_{X,Y}(x,y) = \begin{cases} c & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Value of c?

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1\int_0^x c\,\mathrm{d}y\,\mathrm{d}x = 1$$

$$\Rightarrow c = 2$$

$\mathrm{P}[X > 1/2]$?

$$\mathrm{P}[X > 1/2] = \int_{1/2}^1\int_0^x 2\,\mathrm{d}y\,\mathrm{d}x = 3/4$$

The marginal PDF

How to obtain $f_X(x)$ and $f_Y(y)$ from $f_{X,Y}(x,y)$?

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

Remember for PMFs:

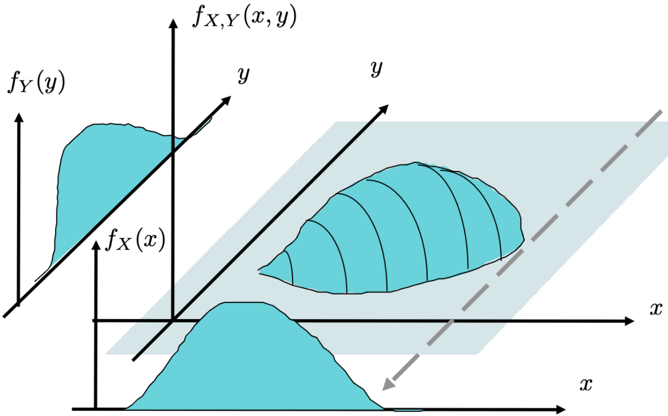
$$
\begin{aligned}
P_X(x) &= \sum_{y \in S_Y} P_{X,Y}(x,y) \\
P_Y(y) &= \sum_{x \in S_X} P_{X,Y}(x,y)
\end{aligned}
$$

Proof

Using the CDF:

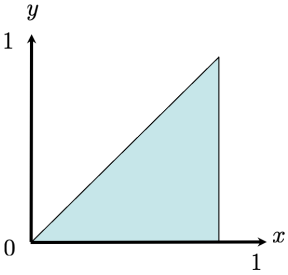
$$
\begin{aligned}
F_X(x) &= \mathrm{P}[X \leq x] = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{X,Y}(u,y)\mathrm{d}y \right) \mathrm{d}u \\
f_X(x) &= \frac{\partial F_X(x)}{\partial x} = \int_{-\infty}^{\infty} f_{X,Y}(u,y)\mathrm{d}y
\end{aligned}
$$

Illustration



Example marginal PDF

$$f_{X,Y} = \begin{cases} 6y & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$



What are the marginal PDFs?

$$\begin{aligned} f_X(x) &= \int_0^x 6y \, \mathrm{d}y = 3x^2 \, , \quad 0 \leq x \leq 1 \\ f_Y(y) &= \int_y^1 6y \, \mathrm{d}x = 6y(1-y) \, , \quad 0 \leq y \leq 1 \end{aligned}$$

Independent RVs

Two RVs are independent if and only if:

- Discrete RVs: $P_{X,Y}(x,y) = P_X(x)\,P_Y(y)$
- Continuous RVs: $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$

Also: $F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$

Example

$$f_{X,Y}(x,y) = \begin{cases} 6y & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent?

Independent RVs

Two RVs are independent if and only if:

- Discrete RVs: $P_{X,Y}(x,y) = P_X(x)\,P_Y(y)$
- Continuous RVs: $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$

Also: $F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$

Example

$$f_{X,Y}(x,y) = \begin{cases} 6y & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent?

No, as $f_X(x)f_Y(y) \neq f_{X,Y}(x,y)$

Expected values of functions of RVs

Let $g(X, Y)$ be a function of the variables X and Y.

Expected value of $W = g(X, Y)$, i.e., $E\left[g(X, Y)\right]$?

- Determine the PDF of W and calculate $\int_{-\infty}^{\infty} w\, f_W(w)\, \mathrm{d}w$.

  However, often it is difficult to find the PDF of $W = g(X, Y)$.

- Much easier: $\mathrm{E}[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\, \mathrm{d}x\, \mathrm{d}y$

The sum of two RVs, $W = X + Y$, then results in

- $\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$
- $\mathrm{var}[X + Y] = \mathrm{var}[X] + \mathrm{var}[Y] + 2\underbrace{\mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])]}_{\text{covariance}}$

Covariance

Covariance: $\text{cov}[X, Y] = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])]$

Correlation: $r_{X,Y} = \text{E}[X\,Y]$

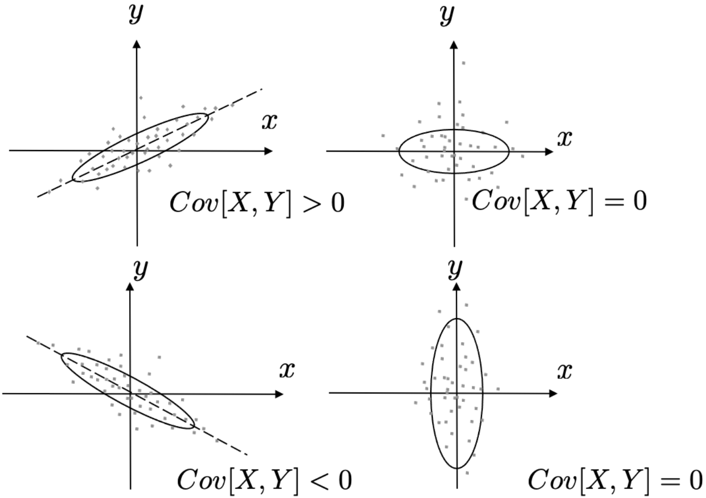
- $\text{cov}[X, Y] = r_{X,Y} - \text{E}[X]\,\text{E}[Y]$.
- If X or Y is zero mean, then the covariance is equal to the correlation.

The covariance describes how X and Y vary together:

Let $W = (X - \text{E}[X])(Y - \text{E}[Y])$.

- $\text{cov}[X, Y] = \text{E}[W] > 0$? Typical values of W positive.
  $\Rightarrow X - \text{E}[X]$ and $Y - \text{E}[Y]$ have the same sign.
- $\text{cov}[X, Y] = \text{E}[W] < 0$? Typical values of W negative.
  $\Rightarrow X - \text{E}[X]$ and $Y - \text{E}[Y]$ have opposite sign.

Covariance



The correlation coefficient

The **correlation coefficient** between two RVs is defined as

$$\rho_{X,Y} = \frac{\text{cov}[X, Y]}{\sqrt{\text{var}[X]\text{var}[Y]}}$$

- $-1 \leq \rho_{X,Y} \leq 1$
- If $|\rho_{X,Y}| = 1$: perfectly correlated; if $\rho_{X,Y} = 0$: uncorrelated.

The correlation coefficient $\rho_{X,Y}$ does not change under linear transformations:

- Let: $X_1 = aX + b$ and $Y_1 = cY + d$, then

$$\rho_{X_1,Y_1} = \rho_{X,Y}$$

- However, $\text{cov}[X_1, Y_1] = a\,c\ \text{cov}[X, Y]$

Independent RVs

If two RVs are independent, $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, then:

- $\mathrm{E}[g(X)h(Y)] = \mathrm{E}[g(X)]\,\mathrm{E}[h(Y)]$
- $r_{X,Y} = \mathrm{E}[X\ Y] = \mathrm{E}[X]\,E[Y]$
- $\mathrm{cov}[X,Y] = \rho_{X,Y} = 0$
- $\mathrm{var}[X+Y] = \mathrm{var}[X] + \mathrm{var}[Y]$

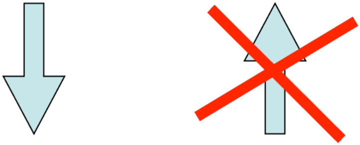
Proof

$$\begin{aligned}
\mathrm{E}[g(X)h(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(X)h(Y)f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(X)h(Y)f_X(x)f_Y(y)\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty} g(X)f_X(x)\mathrm{d}x \int_{-\infty}^{\infty} h(Y)f_Y(y)\mathrm{d}y \\
&= \mathrm{E}[g(X)]\,\mathrm{E}[h(Y)]
\end{aligned}$$

Independent vs. uncorrelated

Two RVs X and Y are independent:

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

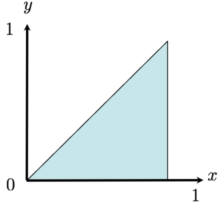


X and Y are uncorrelated:

$$\text{cov}[X, Y] = 0$$

Example: covariance

$$f_{X,Y}(x,y) = \begin{cases} 6y & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

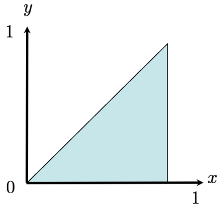


What are

- $\mathrm{E}[X]$ and $\mathrm{E}[Y]$?
- $\mathrm{E}[X\,Y]$
- $\mathrm{cov}(X, Y)$?

Example: covariance

$$f_{X,Y}(x,y) = \begin{cases} 6y & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$



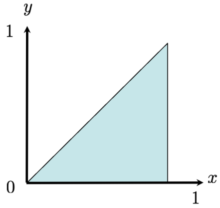
$$f_X(x) = \int_0^x 6y\,\mathrm{d}y = 3x^2\,, \quad 0 \le x \le 1$$

$$f_Y(y) = \int_y^1 6y\,\mathrm{d}x = 6y(1-y)\,, \quad 0 \le y \le 1$$

$$\mathrm{E}[X] = \int_0^1 x\,3x^2\,\mathrm{d}x = 0.75\,, \qquad \mathrm{E}[Y] = \int_0^1 y\,6y(1-y)\,\mathrm{d}y = 0.5$$

Example: covariance

$$f_{X,Y}(x,y) = \begin{cases} 6y & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$



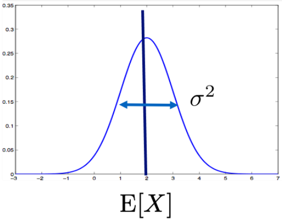
$$\mathrm{E}[XY] = \int_0^1 \int_y^1 xy\, 6y \,\mathrm{d}x\mathrm{d}y = 0.4$$

$$\mathrm{cov}[X,Y] = 0.4 - 0.75 \cdot 0.5 = 0.025$$

Gaussian variables

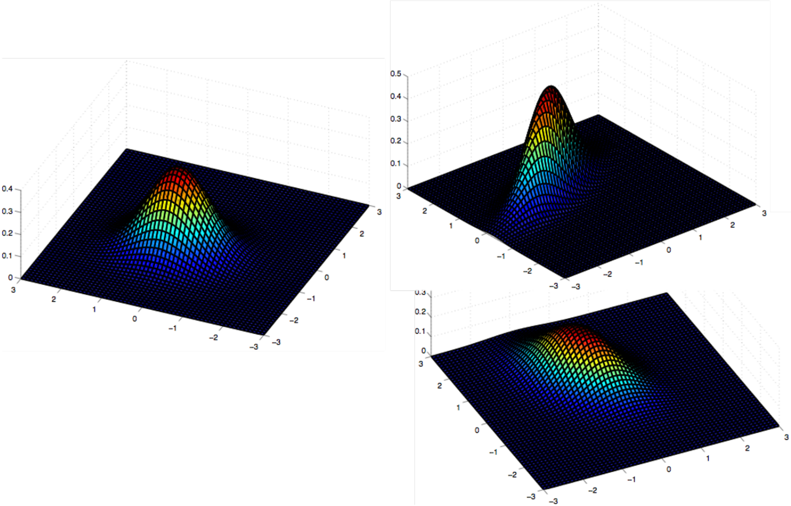
We know the marginal PDFs of the Gaussian scalar RVs X and Y:

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mathrm{E}[X])^2}{2\sigma^2}} \\
f_Y(y) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mathrm{E}[Y])^2}{2\sigma^2}}
\end{aligned}
$$



What is the joint PDF of X and Y (the bivariate Gaussian PDF)?

Joint Gaussian PDF



Exact shape depends on the
covariance/correlation coefficient between X and Y

Gaussian variables

$$f_{X,Y}(x,y) = \frac{\exp\left[-\frac{\left(\frac{x-\mathrm{E}[X]}{\sigma_X}\right)^2 - \frac{2\rho_{X,Y}(x-\mathrm{E}[X])(y-\mathrm{E}[Y])}{\sigma_X\sigma_Y} + \left(\frac{y-\mathrm{E}[Y]}{\sigma_Y}\right)^2}{2(1-\rho_{X,Y}^2)}\right]}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}$$

[No need to memorize this!]

For uncorrelated variables X and Y ($\mathrm{cov}[X,Y] = 0 \Rightarrow \rho_{X,Y} = 0$):

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-\mathrm{E}[X])^2}{2\sigma_X^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-\frac{(y-\mathrm{E}[Y])^2}{2\sigma_Y^2}}$$

Gaussian variables

Hence, for Gaussian variables it does hold that

$$X \text{ and } Y \text{ are uncorrelated: } \operatorname{cov}[X, Y] = 0$$

$$\Leftrightarrow$$

$$X \text{ and } Y \text{ are independent: } f_{X,Y}(x, y) = f_X(x)\, f_Y(y)$$

Linear transformation of Gaussian RVs

Let X and Y be bivariate Gaussian, and define

$$W_1 = a_1 X + b_1 Y\,, \qquad W_2 = a_2 X + b_2 Y$$

(assume linear independence).

Then W_1 and W_2 are bivariate Gaussian, and

- $\mathrm{E}[W_i] = a_i \mu_X + b_i \mu_Y$
- $\mathrm{var}[W_i] = a_i^2 \sigma_X^2 + b_i^2 \sigma_Y^2 + 2 a_i b_i \rho_{X,Y} \sigma_X \sigma_Y$
- $\mathrm{cov}[W_1, W_2] = a_1 a_2 \sigma_X^2 + b_1 b_2 \sigma_Y^2 + (a_1 b_2 + a_2 b_1) \rho_{X,Y} \sigma_X \sigma_Y$

[Convenient derivations using matrices will come later!]

Before the next lecture

- Read chapter 5: Multiple random variables
- Make the suggested exercises for this chapter!
  3rd ed: 5.1.1, 5.2.1, 5.2.2, 5.3.2, 5.4.1, 5.5.3, 5.5.9, 5.5.8