

#### One experiment, multiple observations

• Experiment S: select student

• Observations: H(s): length of student

A(s): age of student W(s): weight of student

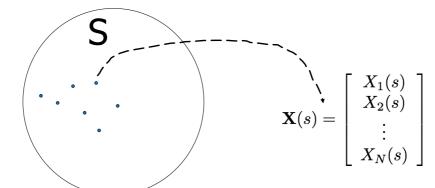
• Obviously these observations are "related"

 Put the observations in a vector (H(s), A(s), W(s))

• Outcomes: (180,23,68), (195,21,75), ...



#### **Multiple Random Variables**



- Chapter 4: N=2
- Pairs of RV; Bivariate
- Chapter 5: N>2
- Random vectors; Multivariate



#### **Pairs of random variables**

#### Discrete RVs:

- (joint) Probability Mass Function (jpdf)
- Marginal PMF can be obtained from the jpdf
- Joint CDF

#### Continuous RVs:

- joint probability density function, from the joint cumulative distribution function
- Marginal PMF can be obtaine from the jpdf
- Joint CDF



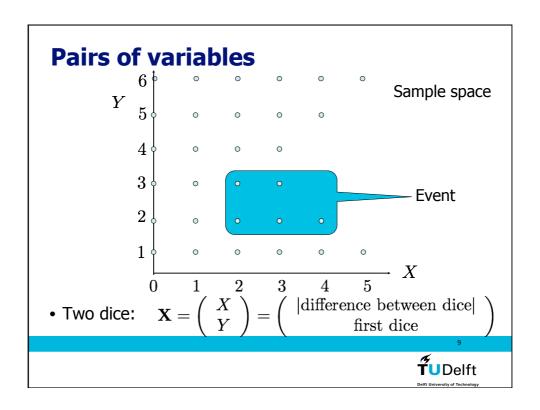
#### **Pairs of (Discrete) variables**

• Example: two dice:

$$\mathbf{X} = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} |\text{difference between dice}| \\ \text{first dice} \end{pmatrix}$$

• The components of a vector random variable are random variables themselves





#### **Notation**

• 1-D case: P(X = X)

 $P(X \in B)$ 

• 2-D case: (B is an event in  $\Re$  )

 $P(X = X,Y = Y) = P_{X,Y}(X,Y)$ 

 $P(X \in B)$ 

(B is an event in  $\Re^2$ )



## **Pairs of variables: joint PMF**

	X = 1	X = 2	X = 3
Y = 1	0,1	0,2	0,05
Y = 2	0,15	0,1	0,2
Y = 3	0,05	0,1	0,05

• For discrete RVs, joint-PMF is fully specified with a table

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

**T**UDelft

## **Pairs of variables: joint PMF**

	X = 1	X = 2	X = 3
Y = 1	0,1	0,2	0,05
Y = 2	0,15	0,1	0,2
Y = 3	0,05	0,1	0,05

• Probability of events can easily be computed:

$$P[X > Y] = ?$$



#### **Pairs of variables: joint PMF**

	X = 1	X = 2	X = 3
Y = 1	0,1	0,2	0,05
Y = 2	0,15	0,1	0,2
Y = 3	0,05	0,1	0,05

• Probability of events can easily be computed:

$$P[X > Y] = P[\{(2,1), (3,1), (3,2)\}]$$

$$= P_{X,Y}[2,1] + P_{X,Y}[3,1] + P_{X,Y}[3,2]$$

$$= 0.45$$



#### **Joint and marginal PMF**

Joint probability mass function

$$P_{X,Y}(X = x, Y = y) = P_{X,Y}(x, y)$$

• Marginal PMF:  $P_X(x)$   $P_Y(y)$ 

$$P_X(x) = P_{X,Y}(x, \text{any } y)$$
$$= \sum_{\text{all } y} P_{X,Y}(x, y)$$





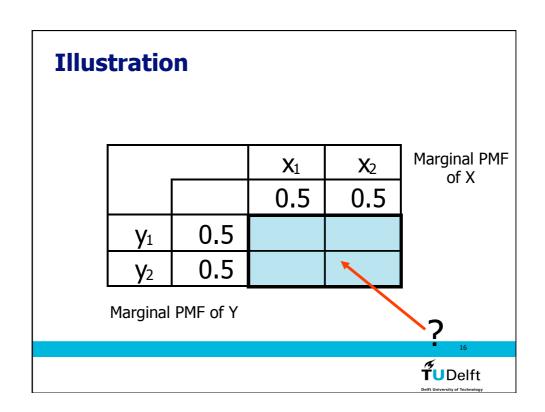
# **Marginal probability**

	X = 1	X = 2	X = 3
Y = 1	0,1	0,2	0,05
Y = 2	0,15	0,1	0,2
Y = 3	0,05	0,1	0,05
	0,30	0,40	0,30

$$P_X(2) = 0.4$$

15





## **Illustration**

		$X_1$	<b>X</b> 2
		0.5	0.5
<b>y</b> <sub>1</sub>	0.5	0.5	0
<b>y</b> 2	0.5	0	0.5

Marginal PMF of X

Marginal PMF of Y

**T**∪Delft

## Illustration

		<b>X</b> <sub>1</sub>	<b>X</b> <sub>2</sub>
		0.5	0.5
<b>y</b> <sub>1</sub>	0.5	0.25	0.25
<b>y</b> <sub>2</sub>	0.5	0.25	0.25

Marginal PMF of X

Marginal PMF of Y



#### **Illustration**



Marginal PMF of X

Marginal PMF of Y



#### **Marginal and joint probabilities**

• Note, when you know the marginal prob., it does NOT mean that you can derive the joint probabilities!!

Joint probability mass function





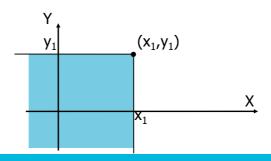
Marginal probability mass functions



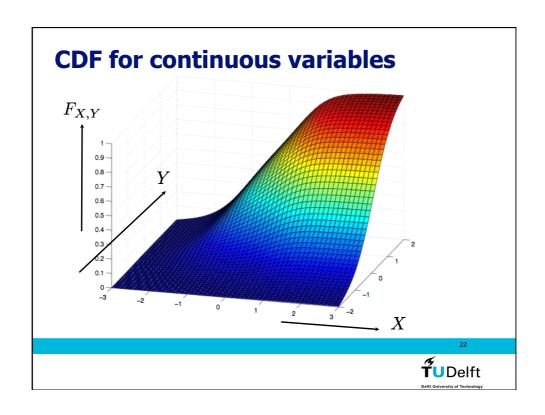
#### **Continuous random variables**

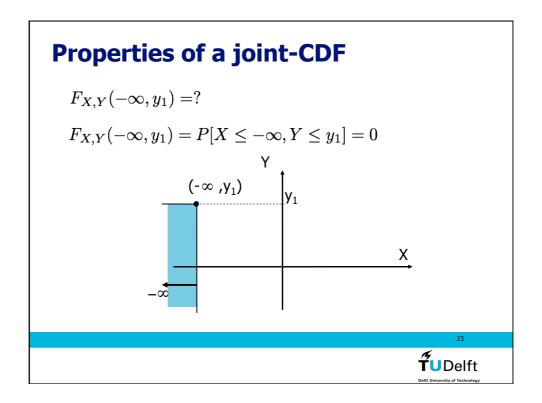
- How to generalize to 2D?
- Again, introduce the joint cumulative distribution function (cdf):

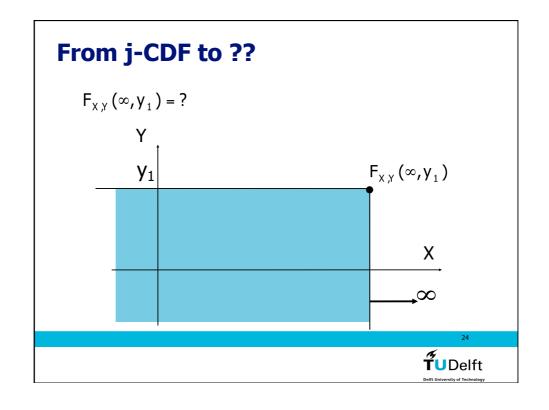
$$F_{X,Y}(x_1, y_1) = P(X \le x_1, Y \le y_1)$$

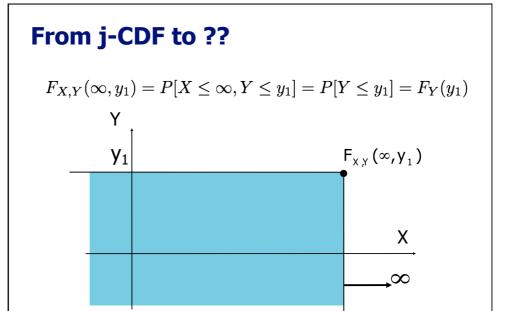












## **Warning**

- The determination of the 2-D j-CDF has to be done for all  $(x,y)\in\Re\cdot\Re\cdot$
- This is much harder than in 1-D
- For that reason, always start from the definition

$$F_{X,Y}(X_1, Y_1) = P(X \le X_1, Y \le Y_1)$$

and make sure that the CDF is defined for all (x,y) values.



**T**UDelft

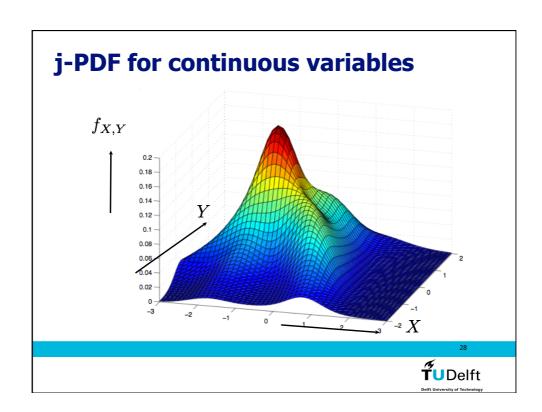
## **Probability density function**

• Analogous to the 1D case, for the 2D case:

$$f_{X,Y}(x,y) = \frac{d^2 F_{X,Y}(x,y)}{dx \, dy} = \frac{d^2 P(X \le x, Y \le y)}{dx \, dy}$$

• This is the **joint**-pdf, or j-pdf

**T**UDelft



## j-PDF ⇔ j-CDF

• 1-D Case

$$F_X(a) = \int_{-\infty}^a f_X(x) dx$$

• 2-D Case

$$F_{X,Y}(a,b) = \int_{-\infty-\infty}^{a} f_{X,Y}(x,y) dy dx$$

$$P(X \leq a, Y \leq b)$$



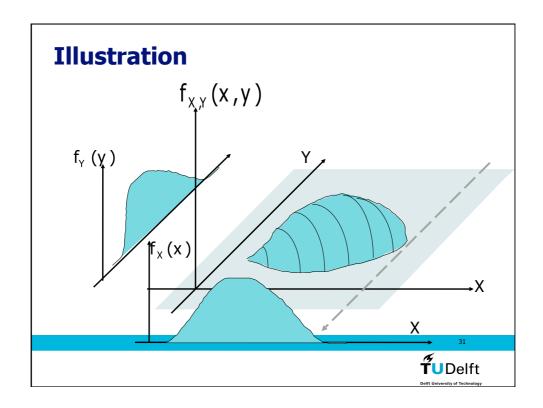
#### Marginal pdf's

• Analogous to the marginal probability mass functions, we define marginal probability density functions:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$$

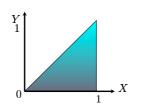
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$$





## **Example marginal pdf**

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

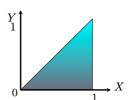


• What are the marginal pdfs?



#### **Example marginal pdf**

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



· What are the marginal pdfs?

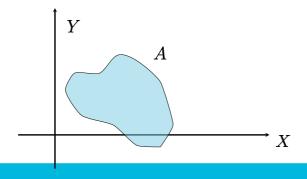
$$f_X(x) = \int_0^x 6y dy = 3x^2, \quad 0 \le x \le 1$$
  
 $f_Y(y) = \int_y^1 6y dx = 6y(1-y), \quad 0 \le y \le 1$ 

33



#### **Probability of events**

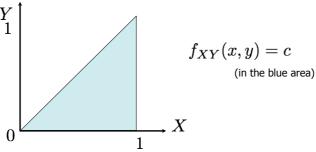
- Event A:  $P(A) = \iint_A f_{X,Y}(x,y) dx dy$
- Can be very hard due to boundaries of the event A



34



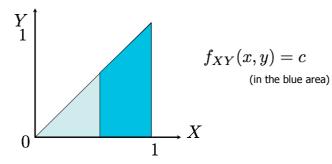
## **Example**



• Value of c?  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y) dy \, dx = \int_{0}^{1} \int_{0}^{x} c dy \, dx = 1$   $\to c = 2$ 

**T**UDelft

## **Example**



- Value of c?  $\rightarrow c=2$
- P[X > 1/2] ?  $\int_{0.5}^{1} \int_{0}^{x} c dy \, dx = \dots$   $\to \dots = 3/4$



#### **Pairs of random variables**

- Independence
- Conditional PMF
- Expectation, variance and COvariance
- Correlation



#### **Independent Random variables**

• X and Y are (stochastically) independent if for any events A and B:

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

• Consequence:

$$X, Y$$
 independent  $\Leftrightarrow P_{X,Y}(x,y) = P_X(x)P_Y(y)$   
 $X, Y$  independent  $\Leftrightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y)$ 



#### **Test procedure (PMFs)**

- Assume we are given  $P_{X,Y}(x,y)$
- Compute the marginal PMFs

$$P_X(x) = \sum_{\text{all } y} P_{X,Y}(x,y) \qquad P_Y(y) = \sum_{\text{all } x} P_{X,Y}(x,y)$$

• Test if

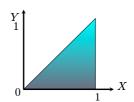
$$P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y)$$

for all x,y



#### **Example independence**

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



$$f_X(x) = \int_0^x 6y dy = 3x^2, \quad 0 \le x \le 1$$
$$f_Y(y) = \int_y^1 6y dx = 6y(1-y), \quad 0 \le y \le 1$$

• Are X and Y independent?



#### **Conditional PMF/PDF**

- $P_{X,Y}(x,y)$  contains all information about the behavior of
  - X
  - Y
  - X and Y jointly
  - dependency
- What about the influence of X on the behavior of Y, and vice versa?

$$P_{X|Y}(x|y)$$
 and  $P_{Y|X}(y|x)$   
 $f_{X|Y}(x|y)$  and  $f_{Y|X}(y|x)$ 



#### **Conditional PMF/PDF**

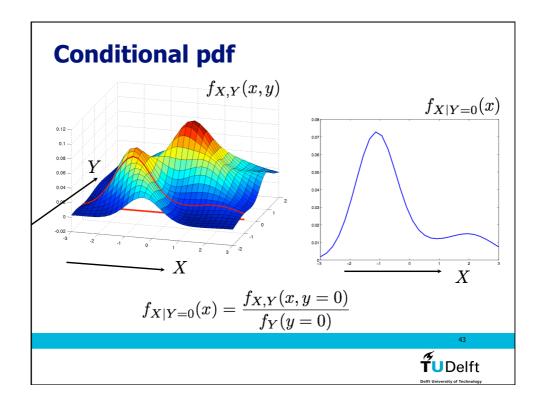
- Behavior of a random variable, given (the outcome of) another variable
- Discrete

$$P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

 Continuous (is a little tricky, because denominator cannot be a probability of an outcome)

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$





#### **Expectation**

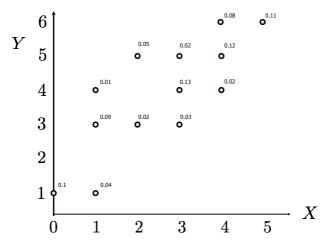
- Expectation of a random vector is a vector
- Compute element-wise the expectation:

$$E[\mathbf{X}] = E\left[\begin{pmatrix} X \\ Y \end{pmatrix}\right] = \begin{bmatrix} E[X] \\ E[Y] \end{bmatrix}$$
$$= \begin{bmatrix} \int_{-\infty}^{\infty} x f_X(x) dx \\ \int_{-\infty}^{\infty} y f_Y(y) dy \end{bmatrix}$$

• With variance it is slightly more involved...



#### **COvariance**



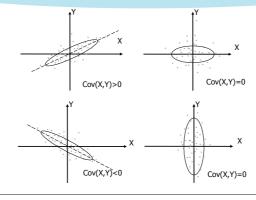
• Characterize the joint behavior of two random variables



#### **Covariance of X and Y**

- Describes the joint dispersion (global dependency) of X and Y
- Product of the fluctuations of X and Y

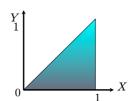
$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$





## **Example covariance**

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



• What is the covariance between X and Y?

$$Cov(X,Y) = E[XY] - E[X]E[Y]$$
  
 $E[XY] = ?$   
 $E[X] = ?$ 

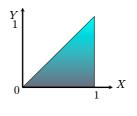


#### **Example covariance**

E[Y] = ?

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

$$E[XY] = \int_0^1 \int_y^1 xy \, 6y \, dx \, dy = 0.4$$





#### **Example covariance**

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

$$E[XY] = \int_0^1 \int_y^1 xy \, 6y \, dx \, dy = 0.4$$

$$f_X(x) = \int_0^x 6y \, dy = 3x^2, \quad 0 \le x \le 1$$

$$f_Y(y) = \int_y^1 6y \, dx = 6y(1-y), \quad 0 \le y \le 1$$

$$E[X] = \int_0^1 x \, 3x^2 \, dx = 0.75 \qquad E[Y] = \int_0^1 y \, 6y(1-y) \, dy = 0.5$$



#### **Example covariance**

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

$$E[XY] = \int_0^1 \int_y^1 xy \, 6y \, dx \, dy = 0.4$$

$$f_X(x) = \int_0^x 6y \, dy = 3x^2, \quad 0 \le x \le 1$$

$$f_Y(y) = \int_y^1 6y \, dx = 6y(1-y), \quad 0 \le y \le 1$$

$$E[X] = \int_0^1 x \, 3x^2 \, dx = 0.75 \qquad E[Y] = \int_0^1 y \, 6y(1-y) \, dy = 0.5$$

$$Cov(X,Y) = 0.4 - 0.75 \cdot 0.5 = 0.025$$



#### **Covariance and correlation**

• Two related concepts:

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])]$$

$$= E[XY] - E[X]E[Y]$$
correlation

- Cov(X,Y)=0  $\Leftrightarrow$  X and Y are uncorrelated E[XY]=0  $\Leftrightarrow$  X and Y are orthogonal
- Correlation **coefficient**:  $ho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$



## Independence in uncorrelated

• Property if X and Y are independent

$$E[h(X)g(Y)] = \iint_{-\infty}^{\infty} h(X)g(Y)f_{XY}(x,y)dxdy$$

$$= \iint_{-\infty}^{\infty} h(X)g(Y)f_{X}(x)f_{Y}(y)dxdy$$

$$= \int_{-\infty}^{\infty} h(X)f_{X}(x)dx \int_{-\infty}^{\infty} g(Y)f_{Y}(y)dy$$

$$= E[h(X)]E[g(Y)]$$

- Correlation E[XY] = E[X]E[Y]
- $\bullet \ \, \text{Covariance} \ \, Cov(X,Y) = E[XY] E[X]E[Y] = 0 \\$



## **Independent vs. uncorrelated**

X and Y are independent random variables

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





X and Y are uncorrelated random variables

$$Cov(X, Y) = 0$$

53



#### Independent vs. uncorrelated

- Make sure you understand:
  - the difference between
    - covariance, correlation, correlation coefficient
    - uncorrelated and orthogonal
  - independence implies uncorrelated
  - uncorrelated does NOT automatically imply independence
  - correlation implies dependence
  - dependence does not automatically imply correlated

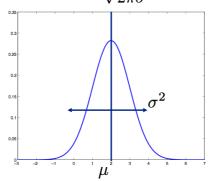
54



#### **Gaussian random variable**

When we have a single Gaussian random variable:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$





#### **Multivariate Gaussian**

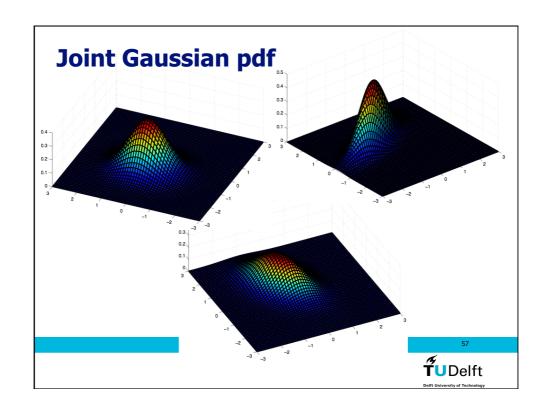
• When we have a **pair** of Gaussian random variables, and we write the pair  $\mathbf{X} = \left[ egin{array}{c} X_1 \ X_2 \end{array} 
ight]$ 

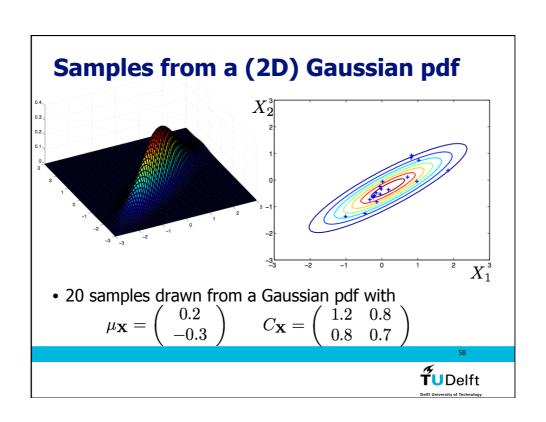
$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi \det(C_{\mathbf{X}})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_{\mathbf{X}})'C_{\mathbf{X}}^{-1}(\mathbf{x} - \mu_{\mathbf{X}})\right)$$

- $m{\cdot}$  Here, the mean vector  $\mu_{\mathbf{X}} = \left[ egin{array}{c} E[X_1] \\ E[X_2] \end{array} 
  ight]$
- and the covariance matrix:

$$C_{\mathbf{X}} = \begin{bmatrix} Var[X_1] & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var[X_2] \end{bmatrix}$$



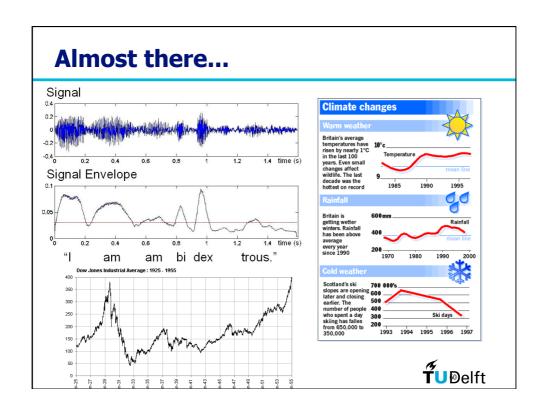




#### **Covered Today**

- Chapter 4
- Key terms
  - Random vectors
  - Joint and marginal PMF and PDF
  - Probability of events
  - Expectation, variance and covariance
  - Independence and uncorrelated
  - Multivariate Gaussian
- Next week:
  - Stochastic processes





#### **Before Next Time**

Exercises corresponding to lecture 2:

- 2<sup>nd</sup> ed: 4.1.1, 4.2.1, 4.2.2, 4.3.2, 4.4.1, 4.5.2, 4.5.6, 4.7.8
- 3<sup>rd</sup> 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

