EURECOM

Sophia Antipolis

Essential Mathematical Methods for Engineers

Lecture 5:

Probability and random variables

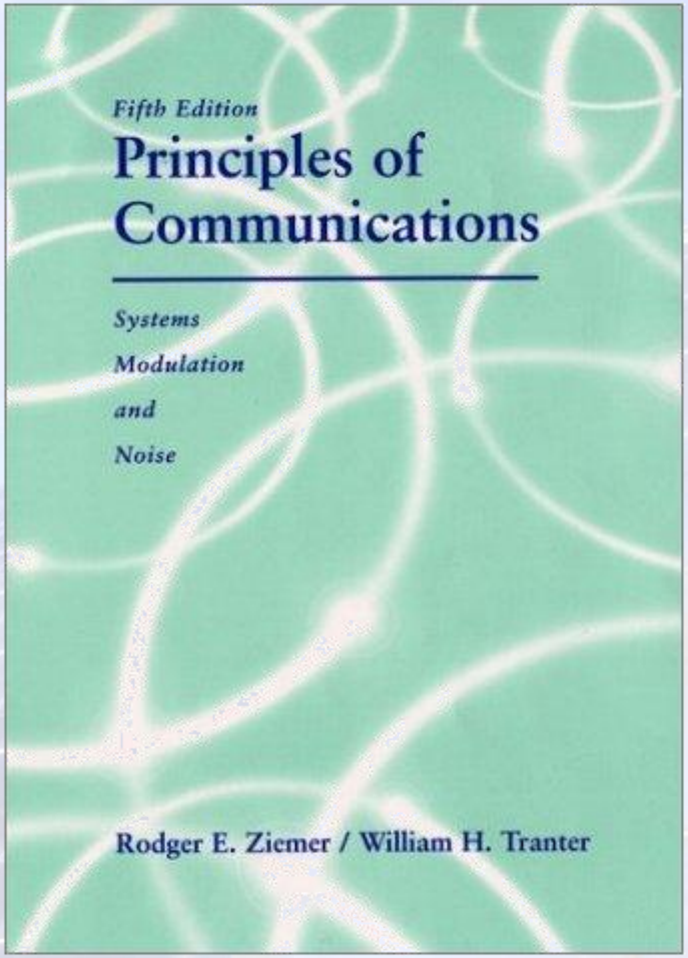
Outline

- sample spaces and the axioms of probability

- random variables and related functions
  - random variables
  - probability (cumulative) distribution functions
  - probability density functions
  - joint cdfs and pdfs
  - transformations of random variables

- statistical averages
  - average of a discrete random variable
  - average of a continuous random variable
  - average of a function of a random variable
  - average of a function of multiple random variables
  - variance of a random variable
  - average of a linear combination of *N* random variables
  - variance of a linear combination of independent random variables
  - the characteristic function
  - the PDF of the sum of two independent random variables
  - covariance and correlation coefficients

Outline

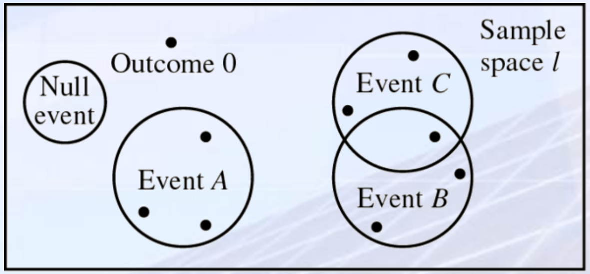


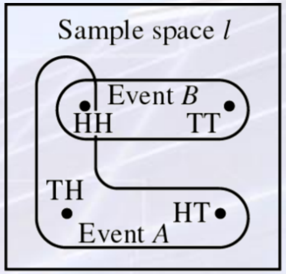
The material in this lecture is adapted from Principles of Communications, 5th ed., Ziemer & Tranter, Wiley

Sample spaces and the axioms of probability

- a chance experiment viewed geometrically
  - all possible outcomes define the sample space, S
  - an event is a collection of outcomes
  - an impossible collection of outcomes is defined as the null event

- some useful notations
  - $A \cup B$ is the event A or B or both
  - $A \cap B$ is the event A and B
  - $\overline{A}$ is the event "not A"





- a set of satisfactory axioms is the following
  - Axiom 1. $P(A) \geq 0$ for all events A in the sample space S
  - Axiom 2. The probability of all possible events occurring, $\mathrm{P}(S) = 1$.
  - Axiom 3. If the occurrence of A precludes the occurrence of B, and vice versa (i.e., A and B are mutually exclusive), then $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$

Random variables

- continuous random variables
  - e.g. noise

- discrete random variables
  - e.g. the tossing of a coin

- notation
  - capital letters denote a random variable
    - e.g. X, Θ and so on
  - lowercase letters denote the value that the random variable takes on
    - e.g. x, θ and so on

Probability (cumulative) distribution functions

- a probabilistic description of random variables
- the cdf $F_X(x)$ is defined as

$$F_X(x) = \text{probability that } X \leq x = P(X \leq x)$$

  – $F_X(x)$ is a function of x, not of the random variable X

- properties:
  – $0 \leq F_X(x) \leq 1$, with $F_X(-\infty) = 0$ and $F_X(\infty) = 1$.
  – $F_X(x)$ is continuous from the right, that is, $\lim_{X \to X_{0^+}} F_X(x) = F_X(x_0)$.
  – $F_X(x)$ is a nondecreasing function of x; that is $F_X(x_1) \leq F_X(x_2)$ if $x_1 < x_2$

Probability density functions

- for the purposes of computing averages the pdf, $f_X(x)$ is more useful
  - it is defined as
$$f_X(x) = \frac{dF_X(x)}{dx}$$
  - thus
$$F_X(x) = \int_{-\infty}^{x} f_X(\eta)d\eta$$

- the pdf has the following properties

$$f_X(x) = \frac{dF_X(x)}{dx} \geq 0 \qquad \int_{-\infty}^{\infty} f_X(x)dx = 1$$

$$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)dx$$

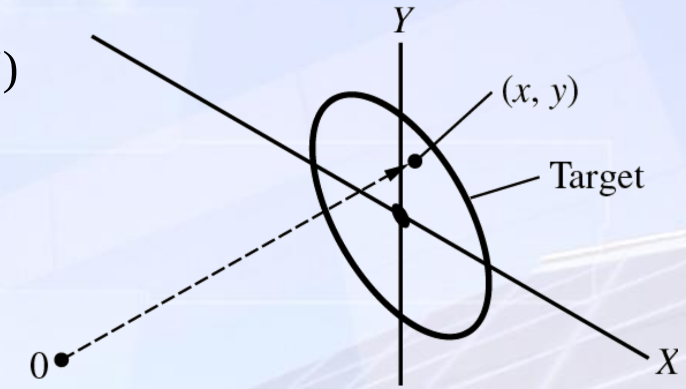
- and by setting $x_1 = x - dx$ and $x_2 = x$

$$f_X(x)dx = P(x - dx < X \leq x)$$

Joint CDFs and PDFs

- for example a dart thrown at a target – two random variables

$$F_{XY}(x,y) = P(x \le X, y \le Y)$$



- the pdf is defined as $$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

$$P(x_1 < X \le x_2, y_1 < Y \le y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{XY}(x,y)dxdy$$

$$F_{XY}(\infty,\infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$$

- letting $x_1 = x - dx$, $x_2 = x$, $y_1 = y - dy$ and $y_2 = y$ we obtain

$$f_{XY}(x,y)dxdy = P(x - dx < X \le x, y - dy < Y \le y)$$

Random variables and related functions

Joint CDFs and PDFs

- the cdf for X irrespective of the value of Y is simply

$$F_X(x) = P(X \le x, -\infty < Y \le \infty) \\ = F_{XY}(x, \infty)$$

- the cdf for Y alone is $F_Y(y) = F_{XY}(\infty, y)$

- $F_X(x)$ and $F_Y(y)$ are referred to as marginal cdf's and may be expressed as

$$F_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(x', y')dx'dy' \qquad F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{XY}(x', y')dx'dy'$$

- since $f_X(x) = \dfrac{dF_X(x)}{dx}$ and $f_Y(y) = \dfrac{dF_Y(y)}{dy}$ from

we obtain $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y')dy'$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x', y)dx'$

i.e. the marginal pdf's are obtained by integrating out the undesired variables

Random variables and related functions
Joint CDFs and PDFs

- two random variables are statistically independent if the values each takes on do not influence the values of the other

- for independent random variables, it must be true for any x and y that

$$P(x \le X, y \le Y) = P(x \le X)P(y \le Y)$$

or in terms of cdf's $$F_{XY}(x, y) = F_X(x)F_Y(y)$$

- if we differentiate first w.r.t. x and then y we obtain $f_{XY}(x, y) = f_X(x) f_Y(y)$

- if two random variables are not independent

$$\begin{aligned} f_{XY}(x, y) &= f_X(x) f_{Y|X}(y \mid x) \\ &= f_Y(y) f_{X|Y}(x \mid y) \end{aligned}$$

intuitively $f_{X|Y}(x \mid y)dx = P[x - dx < X \le x \quad \text{given} \quad Y = y]$

- for independent variables

$$f_{X|Y}(x \mid y) = f_X(x)$$

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

Example

Two random variables X and Y have the joint pdf

$$f_{XY}(x,y) = \begin{cases} Ae^{-(2x+y)}, & x, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

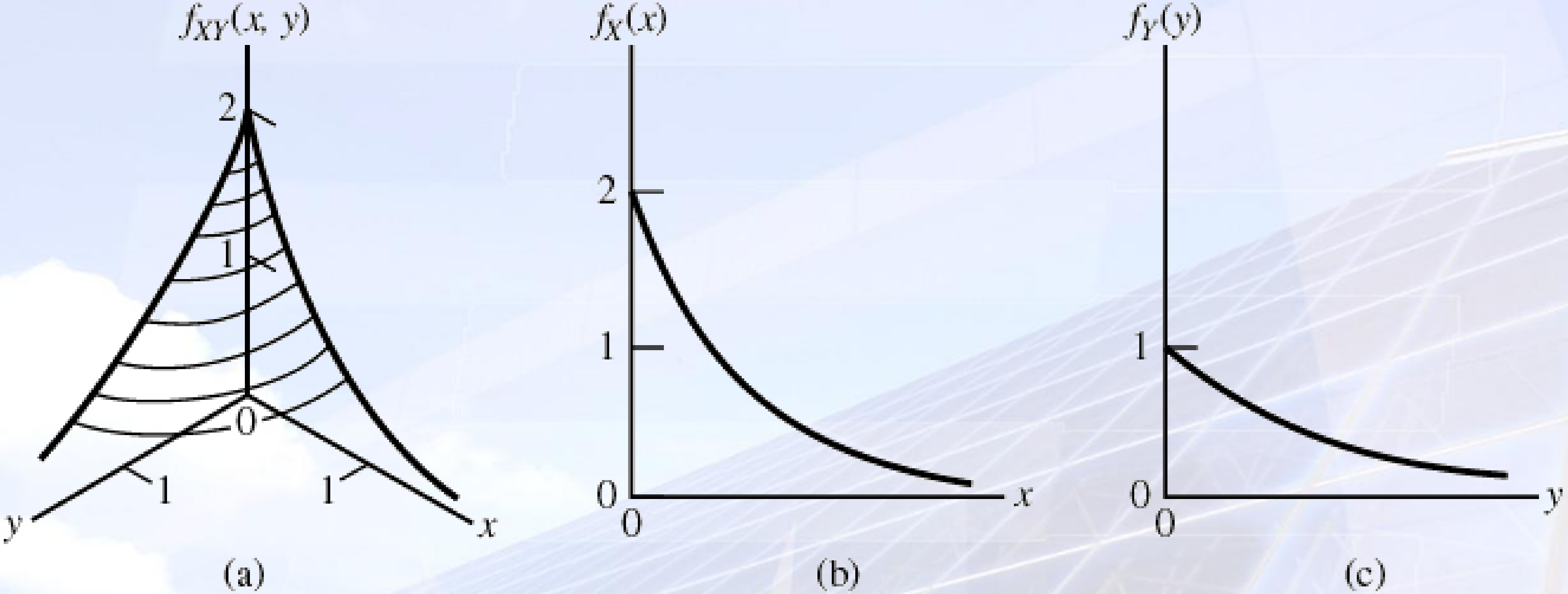
where A is a constant. Determine A and find the two marginal pdf's.

You should see that the rv's are statistically independent.

You should also be able to determine the two cdf's and prove again that the two rv's are independent.

Random variables and related functions

Joint CDFs and PDFs



Example

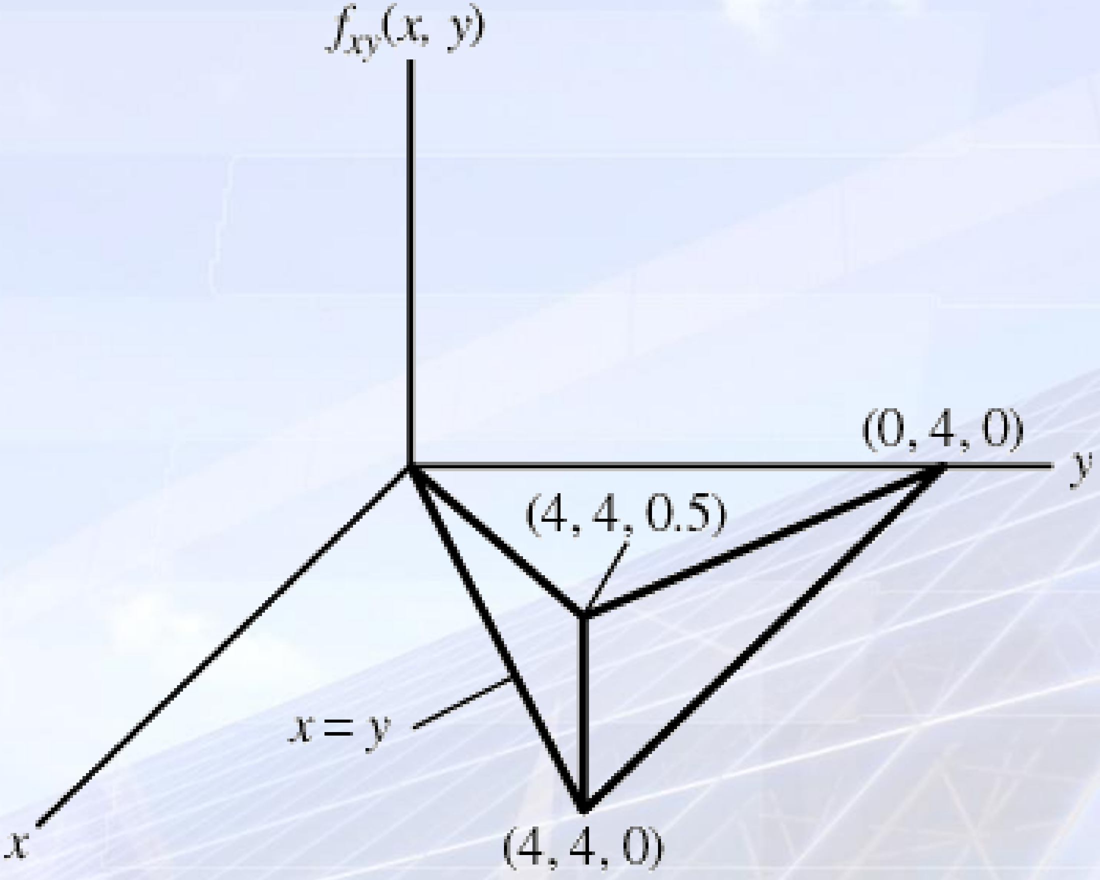
To illustrate the processes of normalisation of joint pdf's, finding marginal from joint pdf's and checking for statistical independence of the corresponding random variables, we consider the joint pdf

$$f_{XY}(x,y) = \begin{cases} \beta xy, & 0 \le x \le y, 0 \le y \le 4 \\ 0, & \text{otherwise} \end{cases}$$

Determine β and find the two marginal pdf's.

just write down the integrals – pay attention to the limits! They may not be what you first think they are.

Random variables and related functions
Joint CDFs and PDFs



Transformations of random variables

- for situations where we know the pdf of a random variable X and desire the pdf of a second random variable Y defined as a function of X, i.e. $Y = \mathrm{g}(X)$

- consider initially monotonic functions

- the probability that X lies in the range $(x - dx, x)$ is the same as the probability that Y lies in the range $(y - dy, y)$ where $y = g(x)$, thus

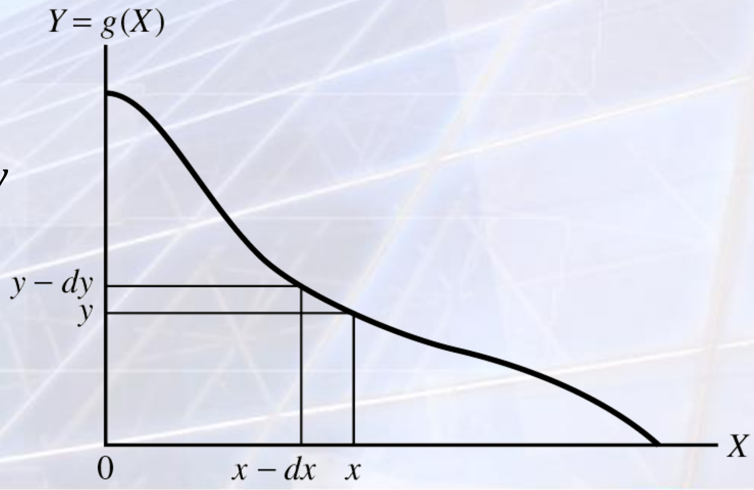
$$f_X(x)dx = f_y(y)dy$$

if $g(X)$ is monotonically increasing and

$$f_X(x)dx = -f_y(y)dy$$

if $g(Y)$ is monotonically decreasing

- for both cases $f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|_{x=g^{-1}(y)}$



Examples

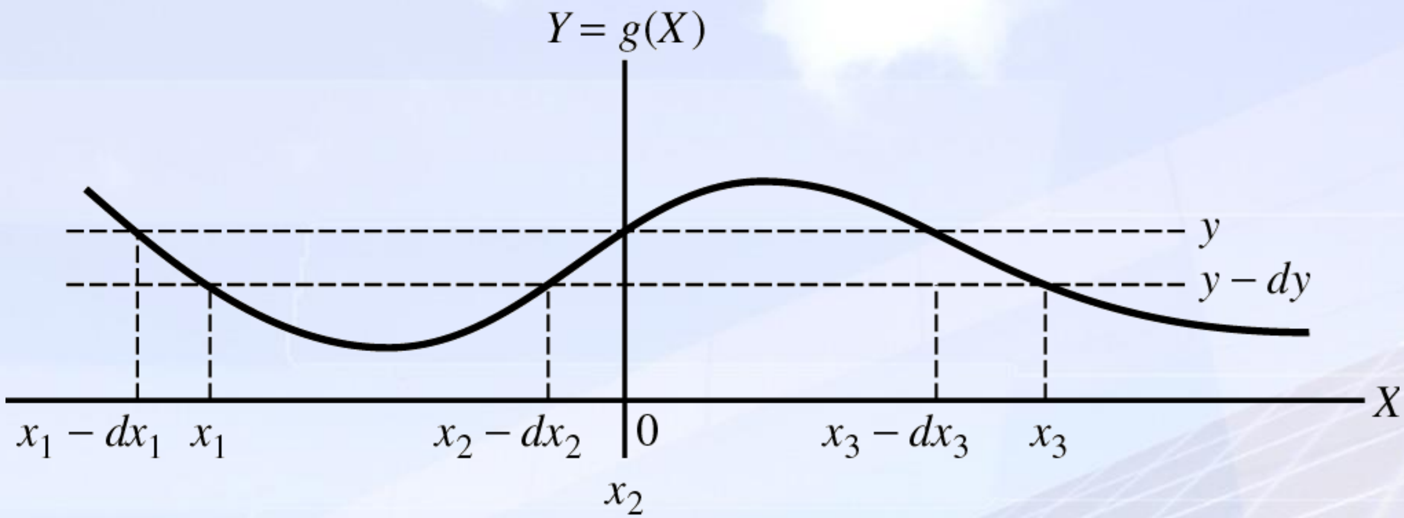
Derive the pdf of the random variable Y defined by

$$Y = -\left(\frac{1}{\pi}\right)\Theta + 1$$

where the random variable Θ has a pdf given by

$$f_{\Theta}(\theta) = \begin{cases} 1/(2\pi), & 0 \leq \theta \leq 2\pi \\ 0, & \text{otherwise} \end{cases}$$

Random variables and related functions
Transformations of random variables



- for the case where $g(x)$ is nonmonotonic as shown the infinitesimal interval $(y - dy, y)$ corresponds to three infinitesimal intervals on the x-axis: $(x_1 - dx_1, x_1)$, $(x_2 - dx_2, x_2)$ and $(x_3 - dx_3, x_3)$
- the probability that X lies in any one of these intervals is equal to the probability that Y lies in the interval $(y - dy, y)$

Random variables and related functions

Transformations of random variables

- this can be generalised to the case of N disjoint intervals where it follows that

$$P(y-dy<Y\le y)=\sum_{i=1}^{N}P(x_i-dx_i<X_i\le x_i)$$

where we have generalised to N intervals on the X axis corresponding to the interval $(y - dy, y)$ on the Y axis

- since $P(y-dy<Y\le y)=f_Y(y)dy$ and $P(x_i-dx_i<X_i\le x_i)=f_X(x_i)dx_i$

$$f_Y(y)=\sum_{i=1}^{N}f_X(x_i)\left|\frac{dx_i}{dy}\right|_{x_i=g_i^{-1}(y)}$$

$g_i^{-1}(y)$ is the i th solution to $g(x)=y$

- the absolute value signs insure that probabilities are positive

Example

Consider the transformation $y = x^2$. If $f_X(x) = 0.5 \exp(-|x|)$, find $f_Y(y)$.

Random variables and related functions
Transformations of random variables

- suppose two new random variables are defined in terms of two old random variables X and Y by the relations

$$U = g_1(X, Y) \quad \text{and} \quad V = g_2(X, Y)$$

- the new pdf $f_{UV}(u,v)$ is obtained from the old pdf $f_{XY}(x,y)$ by writing

$$P(u - du < U \le u, v - dv < V \le v) = P(x - dx < X \le x, y - dy < Y \le y)$$

or

$$f_{UV}(u,v)dA_{UV} = f_{XY}(x,y)dA_{XY}$$

where dA_{UV} is the infinitesimal area in the uv plane corresponding to the infinitesimal area dA_{XY} in the xy plane through the transformation

- the ratio of elementary area dA_{XY} to dA_{UV} is given by the Jacobian $\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$

so that $\quad f_{UV}(u,v) = f_{XY}(x,y) \left| \frac{\partial(x,y)}{\partial(u,v)} \right|_{\substack{x = g_1^{-1}(u,v) \\ y = g_2^{-1}(u,v)}}$

where the inverse functions $g_1^{-1}(u,v)$ and $g_2^{-1}(u,v)$ exist if we assume a one-to-one transformation

Random variables and related functions
Transformations of random variables

Example

Consider the dart-throwing game discussed in connection with joint cdf's and and pdf's. We assume that the joint pdf in terms of rectangular coordinates for the impact point is:

$$f_{XY}(x,y) = \frac{\exp[-(x^2+y^2)/2\sigma^2]}{2\pi\sigma^2}$$

where σ^2 is a constant. This is a special case of the joint Gaussian pdf. Instead of rectangular coordinates, we wish to use polar coordinates R and Θ, defined by:

$$R = \sqrt{X^2+Y^2} \qquad \Theta = \tan^{-1}\left(\frac{Y}{X}\right)$$

Determine the pdf of R and Θ.

EURECOM

Statistical averages

- for a discrete random variable

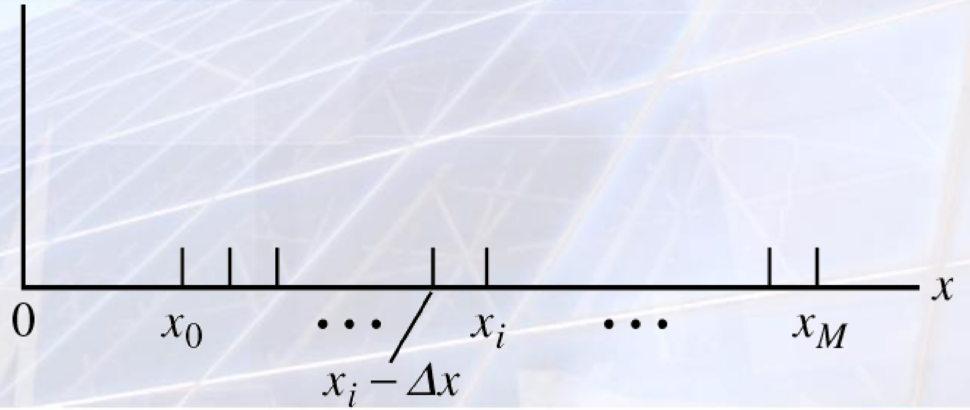
$$\overline{X} = E[X] = \sum_{j=1}^{M} x_j P_j$$

- for a continuous random variable we break up the range of values that X takes into a large number of small sub intervals with length Δx

$$P(x_i - \Delta x < X \le x_i) \cong f_X(x)\Delta x, \quad i = 1,2,\ldots, M$$

  - X is approximated by a discrete random variable that takes on values $x_0, x_1, \ldots x_M$ with probabilities $f_X(x_0), f_X(x_1), \ldots, f_X(x_M)$
  - as Δx approaches dx the expectation is

$$E[X] \cong \sum_{i=0}^{M} x_i f_X(x_i)\Delta x = \int_{-\infty}^{\infty} x\, f_X(x)dx$$



EURECOM

Average of a function of a random variable

- for a function $y = g(x)$

$$E[Y] = \int_{-\infty}^{\infty} y \; f_Y(y) dy$$

where $f_Y(y)$ is the pdf of Y which can be found from $f_X(x)$ from the transformation of a random variable

- sometimes it is more convenient to find the expectation of $g(X)$

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

- the next two examples illustrate this

Example

Suppose the random variable Θ has the pdf

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & |\theta| < \pi \\ 0, & \text{otherwise} \end{cases}$$

Determine $E[\Theta^n]$, referred to as the n^{th} moment of Θ.

The first moment or mean of Θ, $E[\Theta]$, is a measure of the location of $f_\Theta(\theta)$ (i.e. the "centre of mass"). Since $f_\Theta(\theta)$ is symmetrically located about $\theta = 0$, it is not surprising that $E[\Theta] = 0$.

Example

Consider a random variable X that is defined in terms of the uniform random variable Θ considered in the last example by

$$X = \cos\Theta$$

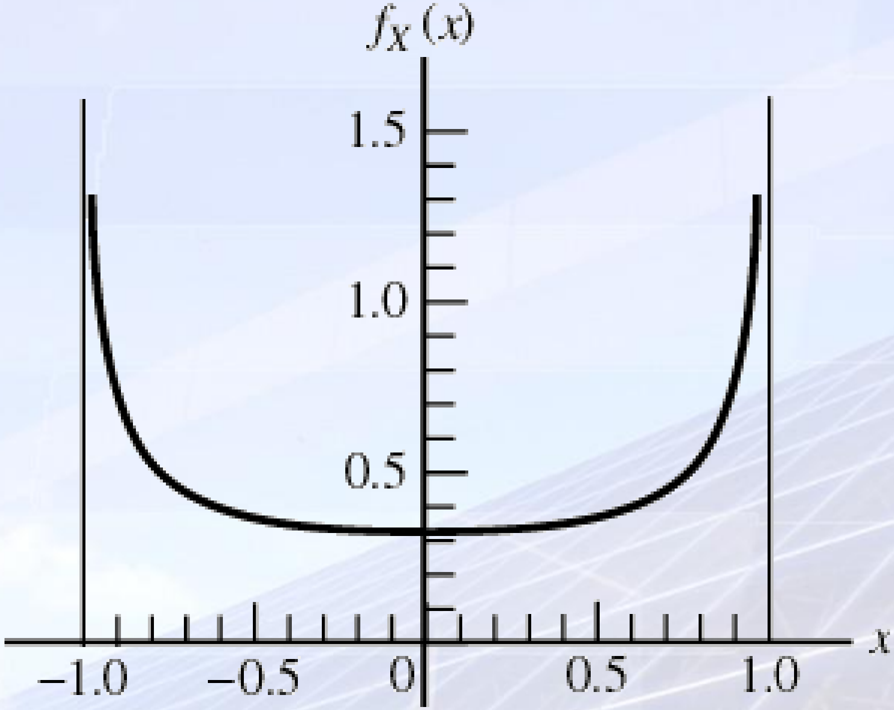
Determine the density function of X, $f_X(x)$ and the first and second moments.

You will need to use: $$\frac{d}{dx}\cos^{-1}x = \frac{1}{\sqrt{1-x^2}}$$

and: $$\int_{-1}^{1}\frac{x^2}{\pi\sqrt{1-x^2}}dx = \frac{1}{2}$$ … unless … ?

Statistical averages

Average of a function of a random variable



Average of a function of multiple random variables

- if $f_{XY}(x,y)$ is the joint pdf of X and Y the expectation of $g(X,Y)$ is

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

  and the generalisation to more than two random variables is obvious

- note that if $g(X,Y)$ is replaced by a function of X alone, say $h(X)$, we obtain

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

another marginal

$$= \int_{-\infty}^{\infty} h(x)\, f_X(x)dx$$

- the concept of conditional expectation may be easier, e.g. for a function $g(X,Y)$ of two random variables X and Y with the joint pdf $f_{XY}(x,y)$

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

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)dx\right] f_Y(y)dy$$

$$= E\{E[g(X,Y)|Y]\}$$

Example

Consider the joint pdf

$$f_{XY}(x,y) = \begin{cases} Ae^{-(2x+y)}, & x, y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Noting that X and Y are statistically independent, determine the expectation of $g(X, Y) = X\,Y$.

★ it's the same example as earlier!

Example

As a specific example of conditional expectation, consider the firing of projectiles at a target. Projectiles are fired until the target is hit for the first time, after which firing ceases. Assume that the probability of a projectile's hitting the target is p and that the firings are independent of one another. Find the average number of projectiles fired at the target.

Variance of a random variable

- the variance, $\mathrm{var}\{X\}$ or σ_X^2 is given by

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

- the standard deviation σ_X measures the concentration of the pdf of X, or $f_X(x)$, about the mean

- a useful relation for obtaining σ_X^2 is

$$\sigma_X^2 = E[X^2] - E^2[X]$$

which is the second moment minus the mean squared

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x)dx = \int_{-\infty}^{\infty} (x^2 - 2xm_X + m_X^2) f_X(x)dx$$
$$= E[X^2] - 2m_X^2 + m_X^2 = E[X^2] - E^2[X]$$

which follows since $\int_{-\infty}^{\infty} x\, f_X(x)dx = m_X$

Statistical averages
Variance of a random variable

Example

Determine the variance of the uniform pdf

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

Statistical averages

Average of a linear combination of N random variables

- the expected value of a linear combination of random variables is the same as the linear combination of their respective means

$$E\left[\sum_{i=1}^{N} a_i X_i\right] = \sum_{i=1}^{N} a_i E[X_i]$$

where $X_1, X_2, \ldots, X_N$ are random variables and $a_1, a_2, \ldots, a_N$ are arbitrary constants

- demonstrating for the case where $N = 2$ and $f_{X1X2}(x_1,x_2)$, the joint pdf of X_1 and X_2

$$E[a_1 X_1 + a_2 X_2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (a_1 x_1 + a_2 x_2) f_{X_1 X_2}(x_1, x_2) dx_1 dx_2$$

$$= a_1 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f_{X_1 X_2}(x_1, x_2) dx_1 dx_2$$

$$+ a_2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f_{X_1 X_2}(x_1, x_2) dx_1 dx_2$$

EURECOM

- considering the first double integral we find that

another marginal

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f_{X_1X_2}(x_1,x_2)dx_1dx_2 = \int_{-\infty}^{\infty} x_1\left\{\int_{-\infty}^{\infty} f_{X_1X_2}(x_1,x_2)dx_2\right\}dx_1$$

$$= \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)dx_1 = E[X_1]$$

- similarly it can be shown that the second double integral equals $E[X_2]$
- the result holds for any N
- does the result hold for dependent and independent variables?

Statistical averages

Variance of a linear combination of independent random variables

- if $X_1, X_2, \ldots, X_N$ are statistically independent random variables then

$$\operatorname{var}\left\{\sum_{i=1}^{N} a_i X_i\right\} = \sum_{i=1}^{N} a_i^2 \operatorname{var}\{X_i\}$$

where $a_1, a_2, \ldots, a_N$ are arbitrary constants and $\operatorname{var}\{X_i\} = E[(X_i - \overline{X_i})^2]$

- again demonstrating the case for $N = 2$
  - let $Z = a_1X_1 + a_2X_2$ and $f_{Xi}(x_i)$ be the marginal pdf of X_i
  - the joint pdf of X_1 and X_2 is $f_{X1}(x_1)\, fX_2(x_2)$ due to independence
  - in addition $\overline{Z} = a_1\overline{X_1} + a_2\overline{X_2}$ and $\operatorname{var}\{Z\} = E[(Z - \overline{Z})^2]$
  - since $Z = a_1X_1 + a_2X_2$ we may write $\operatorname{var}\{Z\}$ as

$$\begin{aligned}
\operatorname{var}\{Z\} &= E\left\{[(a_1X_1 + a_2X_2) - (a_1\overline{X_1} + a_2\overline{X_2})]^2\right\} \\
&= E\left\{[a_1(X_1 - \overline{X_1}) + a_2(X_2 - \overline{X_2})]^2\right\} \\
&= a_1^2 E[(X_1 - \overline{X_1})^2] + 2a_1a_2E[(X_1 - \overline{X_1})(X_2 - \overline{X_2})] + a_2^2 E[(X_2 - \overline{X_2})^2]
\end{aligned}$$

Statistical averages

Variance of a linear combination of independent random variables

$$\begin{aligned}\mathrm{var}\{Z\} &= E\left\{[(a_1X_1 + a_2X_2) - (a_1\overline{X_1} + a_2\overline{X_2})]^2\right\} \\ &= E\left\{[a_1(X_1 - \overline{X_1}) + a_2(X_2 - \overline{X_2})]^2\right\} \\ &= a_1^2 E[(X_1 - \overline{X_1})^2] + 2a_1a_2E[(X_1 - \overline{X_1})(X_2 - \overline{X_2})] + a_2^2 E[(X_2 - \overline{X_2})^2]\end{aligned}$$

- the first and last terms are $\mathrm{var}\{X_1\}$ and $\mathrm{var}\{X_2\}$
- the middle term is zero since

$$\begin{aligned}E[(X_1 - \overline{X_1})(X_2 - \overline{X_2})] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(X_1 - \overline{X_1})(X_2 - \overline{X_2})\, f_{X_1}(x_1) f_{X_2}(x_2)\, dx_1 dx_2 \\ &= \int_{-\infty}^{\infty}(X_1 - \overline{X_1})\, f_{X_1}(x_1)\, dx_1 \int_{-\infty}^{\infty}(X_2 - \overline{X_2})\, f_{X_2}(x_2)\, dx_2 \\ &= (\overline{X_1} - \overline{X_1})(\overline{X_2} - \overline{X_2}) = 0\end{aligned}$$

The characteristic function

- letting $g(X) = \exp(jvX)$ we obtain the characteristic function of X

$$M_X(jv) = E[e^{jvX}] = \int_{-\infty}^{\infty} f_X(x) e^{jvx} dx$$

- with jv in the exponent replaced by $-j\omega$, $M_X(jv)$ would be the Fourier transform of $f_X(x)$

- $f_X(x)$ is obtained from $M_X(jv)$ according to the inverse transform

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M_X(jv) e^{-jvx} dv$$

which is useful when the pdf of a random variable is sought, but the characteristic function is more easily obtained

Statistical averages
The characteristic function

- the characteristic function may be used to obtain the moments of a random variable

- if we differentiate $M_X(jv)$ w.r.t. v

$$\frac{\partial M_X(jv)}{\partial v} = j\int_{-\infty}^{\infty} x f_X(x) e^{jvx} dx$$

- setting $v = 0$ and dividing by j, we obtain

$$E[X] = (-j)\frac{\partial M_X(jv)}{\partial v}\bigg|_{v=0}$$

and by repeated differentiation

$$E[X^n] = (-j)^n \frac{\partial^n M_X(jv)}{\partial v^n}\bigg|_{v=0}$$

Statistical averages
The characteristic function

Example

Use a table of Fourier transforms to obtain the characteristic function of the one-sided exponential pdf

$$f_X(x) = \exp(-x)u(x)$$

and determine an expression for its n^{th} moment.

$$\exp(-at)u(t) \leftrightarrow \frac{1}{a + j2\pi f}$$

The PDF of the sum of two independent random variables

- we can use the characteristic function to determine the pdf of a sum of two independent random variables X & Y, i.e. $Z = X + Y$

$$M_Z(jv) = E[e^{jvZ}] = E[e^{jv(X+Y)}]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{jv(x+y)} f_X(x) f_Y(y) dx dy$$

since the joint pdf of X and Y is $f_X(x)f_Y(y)$ due to independence

- we can write this expression as the product of two integrals since $\exp(jv[x+y]) = \exp(jvx)\exp(jvy)$

$$M_Z(jv) = \int_{-\infty}^{\infty} f_X(x)e^{jvx}dx \int_{-\infty}^{\infty} f_Y(y)e^{jvy}dy$$
$$= E[e^{jvX}]E[e^{jvY}]$$

Statistical averages

The PDF of the sum of two independent random variables

- from the definition of the characteristic function we see that

$$M_Z(jv) = M_X(jv)M_Y(jv)$$

- remembering the similarity to the Fourier transform and that a product in the frequency domain corresponds to convolution in the time domain

$$f_Z(z) = f_X(x) * f_Y(y) = \int_{-\infty}^{\infty} f_X(z-u) f_Y(u) du$$

Example

First sketch and then determine the pdf of Z, the sum of four identically distributed, independent random variables,

$$Z = X_1 + X_2 + X_3 + X_4$$

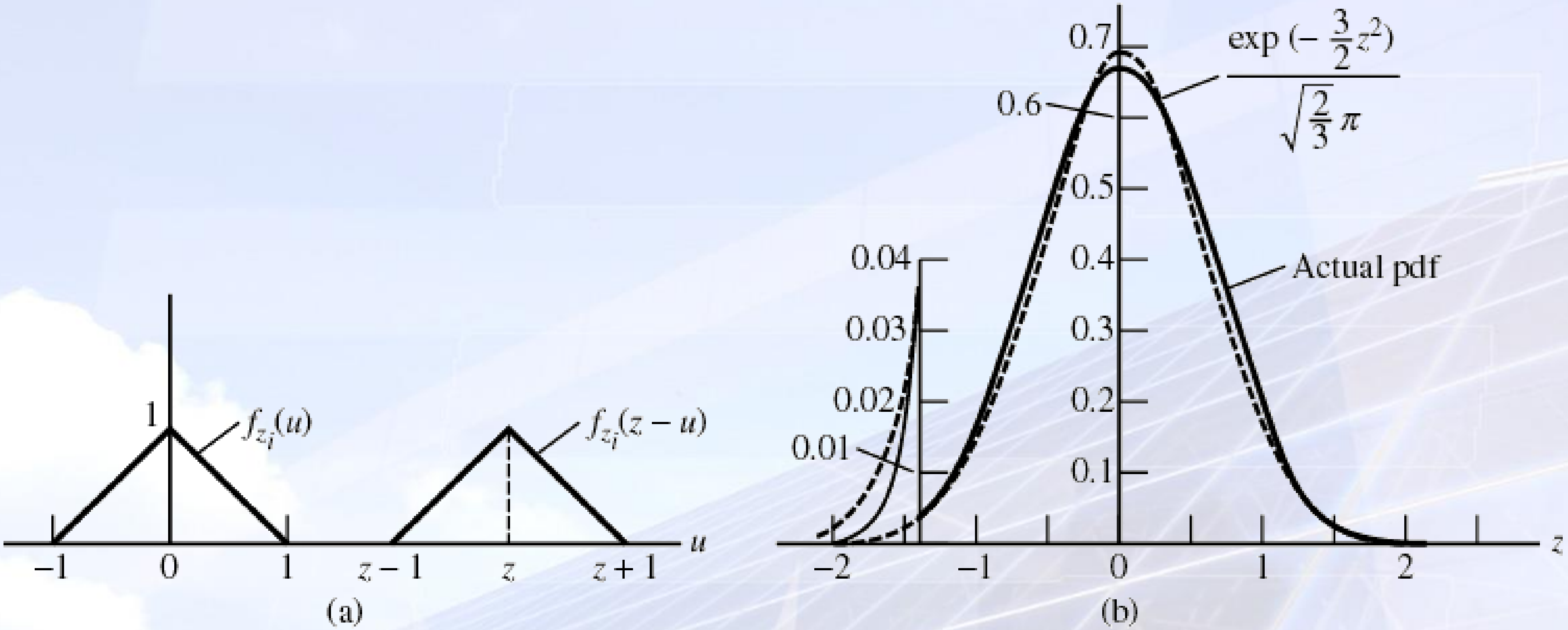
where the pdf of each X_i is given by

$$f_{X_i}(x_i) = \prod(x_i) = \begin{cases} 1, & |x_i| \leq \frac{1}{2} \\ 0, & \text{otherwise}, \; i = 1,2,3,4 \end{cases}$$

and where $\prod(x_i)$ is the unit rectangular pulse function.

Statistical averages

The PDF of the sum of two independent random variables



Covariance and correlation coefficients

- the covariance of two random variables X and Y is defined by

$$\mu_{XY} = E[(X - \overline{X})(Y - \overline{Y})] = E[XY] - E[X]E[Y]$$

- the correlation coefficient is defined by

$$\rho_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}$$

- thus we have the relationship

$$E[XY] = \sigma_X \sigma_Y \rho_{XY} + E[X]E[Y]$$

- both ρ_{XY} and μ_{XY} are measures of the interdependence of X and Y

- the normalisation of the correlation coefficient is such that $-1 \leq \rho_{XY} \leq 1$

Statistical averages
Covariance and correlation coefficients

- if X and Y are independent, their joint pdf $f_{XY}(x,y)$ is the product of the two respective marginal pdfs, that is $f_{XY}(x,y) = f_X(x)f_Y(y)$, thus

$$\mu_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\overline{X})(y-\overline{Y})f_X(x)f_Y(y)dxdy$$
$$= \int_{-\infty}^{\infty}(x-\overline{X})f_X(x)dx\int_{-\infty}^{\infty}(y-\overline{Y})f_Y(y)dy$$
$$= (\overline{X}-\overline{X})(\overline{Y}-\overline{Y}) = 0$$

- considering the case where $X = \pm\alpha Y$, where α is a positive constant

$$\mu_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\pm\alpha y \mp \alpha\overline{Y})(y-\overline{Y})f_{XY}(x,y)dxdy$$
$$= \pm\alpha\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y-\overline{Y})^2 f_{XY}(x,y)dxdy$$
$$= \pm\alpha\sigma_Y^2$$

- we can write the variance of X as $\sigma_X^2 = \alpha^2\sigma_Y^2$, thus

$$\rho_{XY} = +1, \text{ for } X = +\alpha Y \text{ and } \rho_{XY} = -1, \text{ for } X = -\alpha Y, \ \alpha > 0$$