CONTINUOUS RANDOM VARIABLES

DEFINITION : A *continuous random variable* is a *function* $X(s)$ from an *uncountably infinite* sample space $\mathcal{S}$ to the real numbers $\mathbb{R}$,

$$X(\cdot) \quad : \quad \mathcal{S} \quad \rightarrow \quad \mathbb{R} \ .$$

EXAMPLE :

Rotate a *pointer* about a pivot in a plane (like a hand of a clock).

The *outcome* is the *angle* where it stops : $2\pi\theta$, where $\theta \in (0,1]$.

A good *sample space* is all values of θ , *i.e.* $\mathcal{S} = (0,1]$.

A very simple example of a *continuous random variable* is $X(\theta) = \theta$.

Suppose *any outcome*, *i.e.*, any value of θ is "equally likely".

What are the values of

$$P(0 < \theta \leq \tfrac{1}{2}) \quad , \quad P(\tfrac{1}{3} < \theta \leq \tfrac{1}{2}) \quad , \quad P(\theta = \tfrac{1}{\sqrt{2}}) \ ?$$

The (*cumulative*) *probability distribution function* is defined as

$$F_X(x) \equiv P(X \leq x) .$$

Thus

$$F_X(b) - F_X(a) \equiv P(a < X \leq b) .$$

We must have

$$F_X(-\infty) = 0 \quad \text{and} \quad F_X(\infty) = 1 ,$$

i.e.,

$$\lim_{x \to -\infty} F_X(x) = 0 ,$$

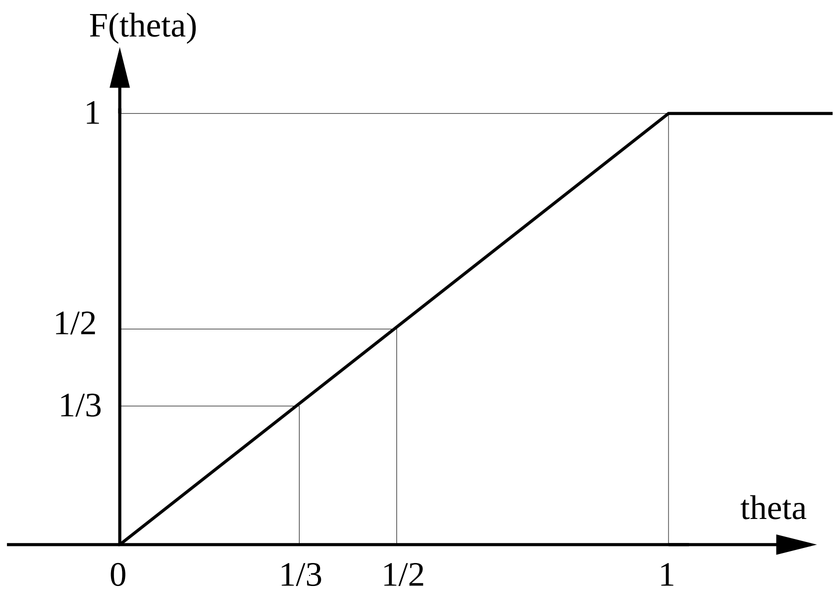
and

$$\lim_{x \to \infty} F_X(x) = 1 .$$

Also, $F_X(x)$ is a *non-decreasing* function of x . (Why ?)

NOTE : All the above is *the same* as for *discrete* random variables !

EXAMPLE : In the "*pointer example*", where $X(\theta) = \theta$, we have the *probability distribution function*



Note that

$$F(\tfrac{1}{3}) \equiv P(X \leq \tfrac{1}{3}) = \tfrac{1}{3} \quad , \quad F(\tfrac{1}{2}) \equiv P(X \leq \tfrac{1}{2}) = \tfrac{1}{2} \ ,$$

$$P(\tfrac{1}{3} < X \leq \tfrac{1}{2}) = F(\tfrac{1}{2}) - F(\tfrac{1}{3}) = \tfrac{1}{2} - \tfrac{1}{3} = \tfrac{1}{6} \ .$$

QUESTION : What is $P(\frac{1}{3} \leq X \leq \frac{1}{2})$?

The *probability density function* is the *derivative* of the probability distribution function :

$$f_X(x) \equiv F'_X(x) \equiv \frac{d}{dx} F_X(x) .$$

EXAMPLE : In the "*pointer example* "

$$F_X(x) = \begin{cases} 0 \ , & x \leq 0 \\ x \ , & 0 < x \leq 1 \\ 1 \ , & 1 < x \end{cases}$$

Thus

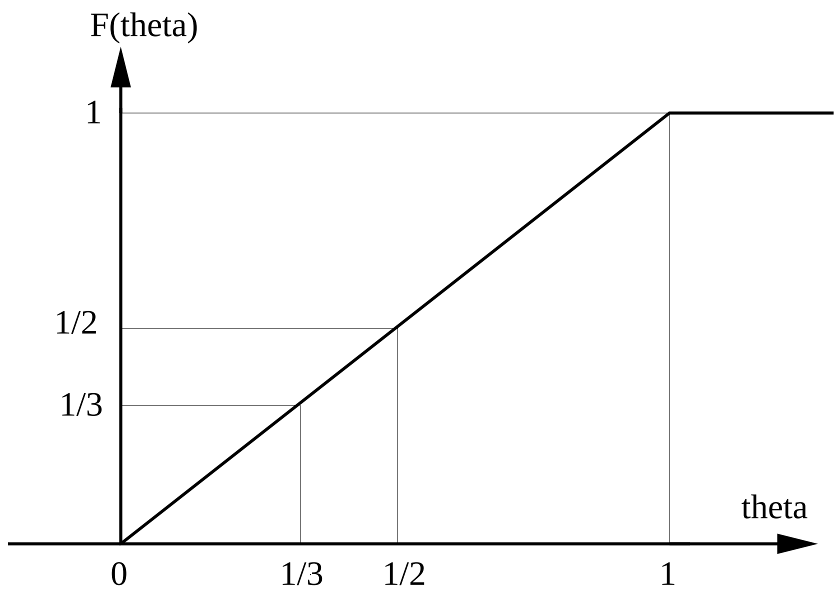
$$f_X(x) = F'_X(x) = \begin{cases} 0 \ , & x \leq 0 \\ 1 \ , & 0 < x \leq 1 \\ 0 \ , & 1 < x \end{cases}$$

NOTATION : When it is clear what X is then we also write

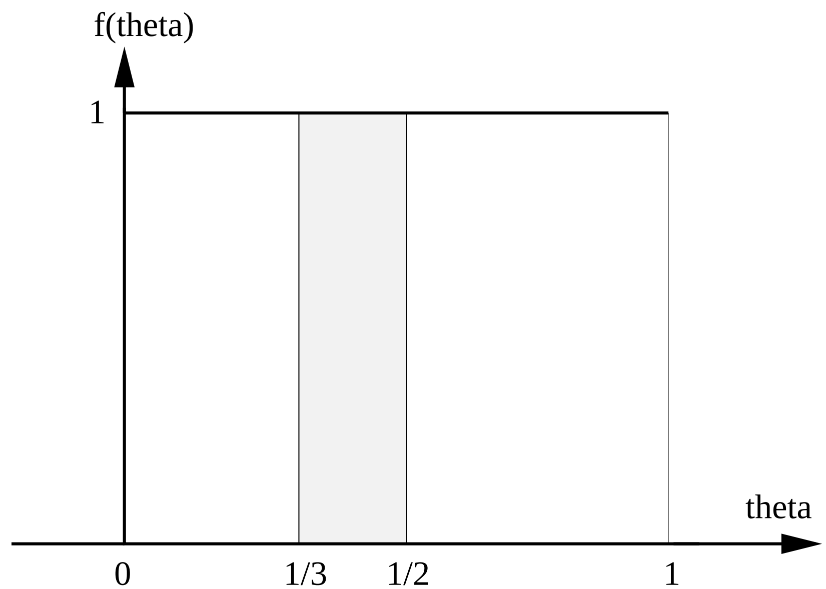
$$f(x) \quad \text{for} \quad f_X(x) \ , \qquad \text{and} \qquad F(x) \quad \text{for} \quad F_X(x) \ .$$

EXAMPLE : (continued ⋯)

$$F(x) = \begin{cases} 0\ , & x \leq 0 \\ x\ , & 0 < x \leq 1 \\ 1\ , & 1 < x \end{cases} \quad , \quad f(x) = \begin{cases} 0\ , & x \leq 0 \\ 1\ , & 0 < x \leq 1 \\ 0\ , & 1 < x \end{cases}$$



Distribution function



Density function

NOTE :

$$P(\frac{1}{3} < X \leq \frac{1}{2}) = \int_{\frac{1}{3}}^{\frac{1}{2}} f(x)\ dx = \frac{1}{6} = \text{the shaded area .}$$

In general, from

$$f(x) \equiv F'(x) ,$$

with

$$F(-\infty) = 0 \quad \text{and} \quad F(\infty) = 1 ,$$

we have from Calculus the following *basic identities* :

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} F'(x)\, dx = F(\infty) - F(-\infty) = 1 ,$$

$$\int_{-\infty}^{x} f(x)\, dx = F(x) - F(-\infty) = F(x) = P(X \leq x) ,$$

$$\int_{a}^{b} f(x)\, dx = F(b) - F(a) = P(a < X \leq b) ,$$

$$\int_{a}^{a} f(x)\, dx = F(a) - F(a) = 0 = P(X = a) .$$

EXERCISE : Draw *graphs* of the distribution and density functions

$$F(x) = \begin{cases} 0\,, & x \leq 0 \\ 1 - e^{-x}\,, & x > 0 \end{cases} \quad , \quad f(x) = \begin{cases} 0\,, & x \leq 0 \\ e^{-x}\,, & x > 0 \end{cases} \quad ,$$

and verify that

- $F(-\infty) = 0\,, \qquad F(\infty) = 1\,,$
- $f(x) = F'(x)\,,$
- $F(x) = \int_0^x f(x)\, dx\,,$ (Why is *zero* as lower limit OK ?)
- $\int_0^\infty f(x)\, dx = 1\,,$
- $P(0 < X \leq 1) = F(1) - F(0) = F(1) = 1 - e^{-1} \cong 0.63\,,$
- $P(X > 1) = 1 - F(1) = e^{-1} \cong 0.37\,,$
- $P(1 < X \leq 2) = F(2) - F(1) = e^{-1} - e^{-2} \cong 0.23\,.$

EXERCISE : For positive integer n, consider the density functions

$$f_n(x) = \begin{cases} cx^n(1-x^n) \ , & 0 \le x \le 1 \\ \\ 0 \ , & \text{otherwise} \end{cases}$$

- Determine the value of c in terms of n .
- Draw the graph of $f_n(x)$ for $n = 1, 2, 4, 8, 16$.
- Determine the distribution function $F_n(x)$.
- Draw the graph of $F_n(x)$ for $n = 1, 2, 3, 4, 8, 16$.
- Determine $P(0 \le X \le \frac{1}{2})$ in terms of n .
- What happens to $P(0 \le X \le \frac{1}{2})$ when n becomes large?
- Determine $P(\frac{9}{10} \le X \le 1)$ in terms of n .
- What happens to $P(\frac{9}{10} \le X \le 1)$ when n becomes large?

Joint distributions

A *joint probability density function* $f_{X,Y}(x,y)$ must satisfy

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy \;=\; 1 \qquad (\text{“Volume”} = 1\ ).$$

The corresponding *joint probability distribution function* is

$$F_{X,Y}(x,y) \;=\; P(X \leq x\ ,\ Y \leq y) \;=\; \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(x,y)\,dx\,dy\ .$$

By Calculus we have $$\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} \;=\; f_{X,Y}(x,y)\ .$$

Also,

$$P(a < X \leq b\ ,\ c < Y \leq d) \;=\; \int_{c}^{d}\int_{a}^{b} f_{X,Y}(x,y)\,dx\,dy\ .$$

EXAMPLE :

If

$$f_{X,Y}(x,y) = \begin{cases} 1 & \text{for } x \in (0,1] \text{ and } y \in (0,1] \ , \\ 0 & \text{otherwise} \ , \end{cases}$$

then, for $x \in (0,1]$ and $y \in (0,1]$,

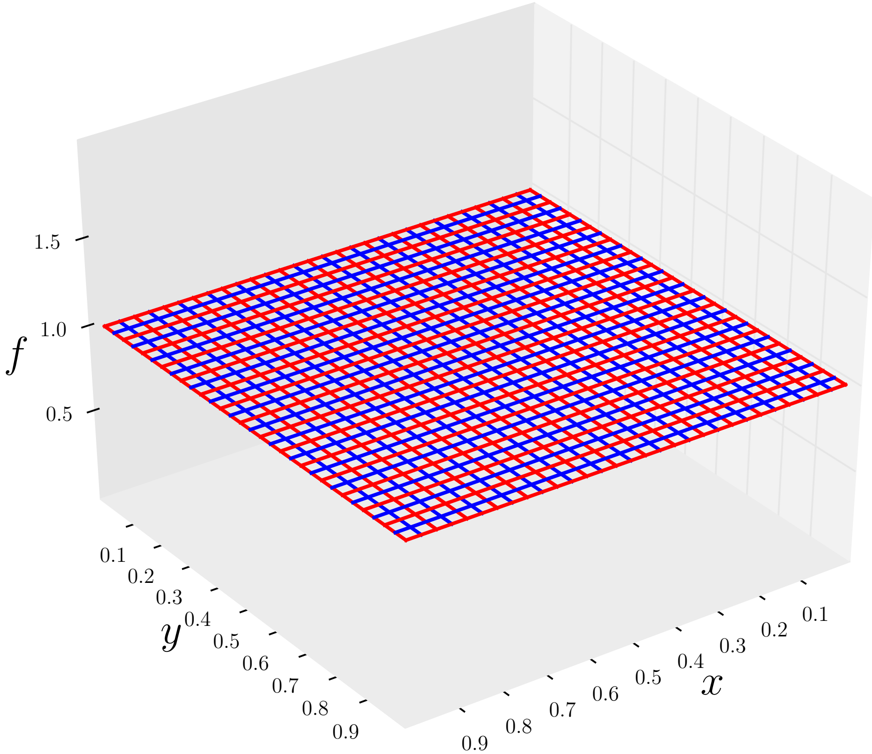
$$F_{X,Y}(x,y) = P(X \leq x \ , \ Y \leq y) = \int_0^y \int_0^x 1 \ dx \ dy = xy \ .$$

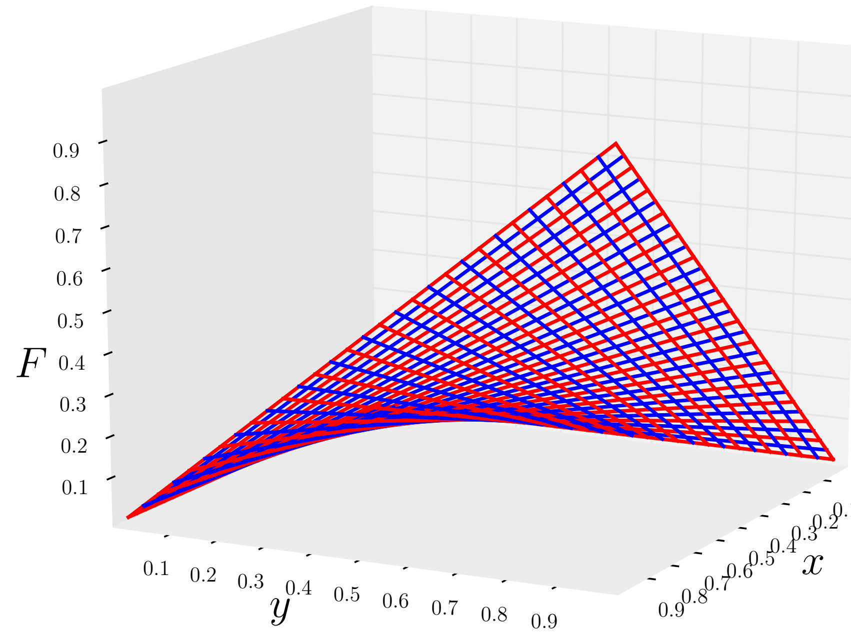
Thus

$$F_{X,Y}(x,y) = xy \ , \qquad \text{for } x \in (0,1] \text{ and } y \in (0,1] \ .$$

For example

$$P(\ X \leq \frac{1}{3} \ , \ Y \leq \frac{1}{2} \ ) = F_{X,Y}(\ \frac{1}{3} \ , \ \frac{1}{2} \ ) = \frac{1}{6} \ .$$





Also,

$$P\left(\frac{1}{3} \leq X \leq \frac{1}{2}\,,\ \frac{1}{4} \leq Y \leq \frac{3}{4}\right) \;=\; \int_{\frac{1}{4}}^{\frac{3}{4}} \int_{\frac{1}{3}}^{\frac{1}{2}} f(x,y)\,dx\,dy \;=\; \frac{1}{12}\,.$$

EXERCISE : Show that we can also compute this as follows :

$$F\left(\tfrac{1}{2},\tfrac{3}{4}\right) \;-\; F\left(\tfrac{1}{3},\tfrac{3}{4}\right) \;-\; F\left(\tfrac{1}{2},\tfrac{1}{4}\right) \;+\; F\left(\tfrac{1}{3},\tfrac{1}{4}\right) \;=\; \tfrac{1}{12}\,.$$

and explain why !

Marginal density functions

The *marginal density functions* are

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy \quad , \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx \; .$$

with corresponding *marginal distribution functions*

$$F_X(x) \equiv P(X \leq x) = \int_{-\infty}^{x} f_X(x)\, dx = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy\, dx \; ,$$

$$F_Y(y) \equiv P(Y \leq y) = \int_{-\infty}^{y} f_Y(y)\, dy = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy \; .$$

By Calculus we have

$$\frac{dF_X(x)}{dx} = f_X(x) \quad , \quad \frac{dF_Y(y)}{dy} = f_Y(y) \; .$$

EXAMPLE : If

$$f_{X,Y}(x,y) = \begin{cases} 1 & \text{for } x \in (0,1] \text{ and } y \in (0,1] \ , \\ 0 & \text{otherwise} \ , \end{cases}$$

then, for $x \in (0,1]$ and $y \in (0,1]$,

$$f_X(x) = \int_0^1 f_{X,Y}(x,y)\, dy = \int_0^1 1\, dy = 1 \ ,$$

$$f_Y(y) = \int_0^1 f_{X,Y}(x,y)\, dx = \int_0^1 1\, dx = 1 \ ,$$

$$F_X(x) = P(X \le x) = \int_0^x f_X(x)\, dx = x \ ,$$

$$F_Y(y) = P(Y \le y) = \int_0^y f_Y(y)\, dy = y \ .$$

For example

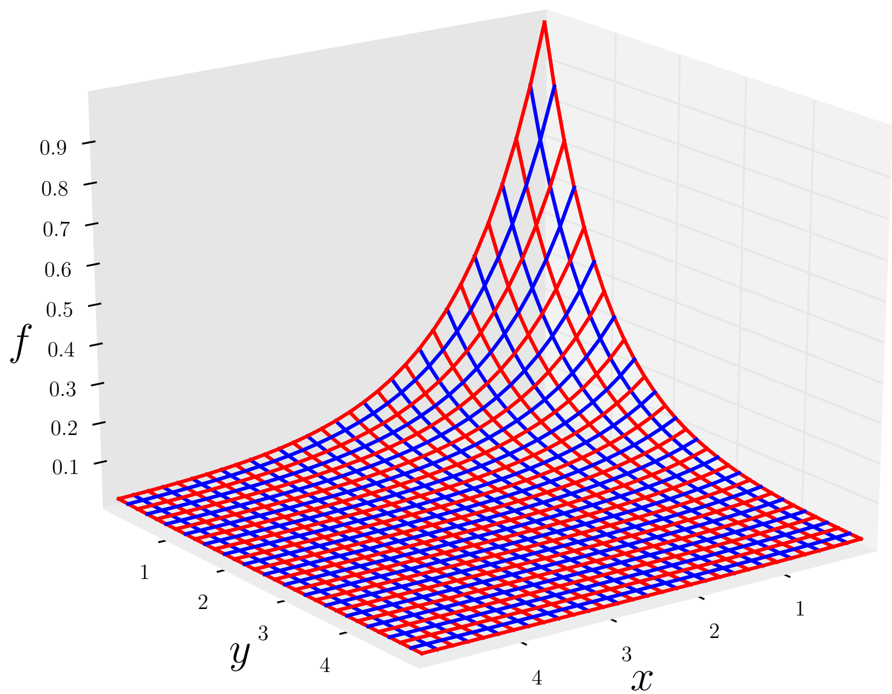
$$P(\ X \le \tfrac{1}{3}\ ) = F_X(\tfrac{1}{3}) = \tfrac{1}{3} \quad , \quad P(\ Y \le \tfrac{1}{2}\ ) = F_Y(\tfrac{1}{2}) = \tfrac{1}{2} \ .$$

EXERCISE :

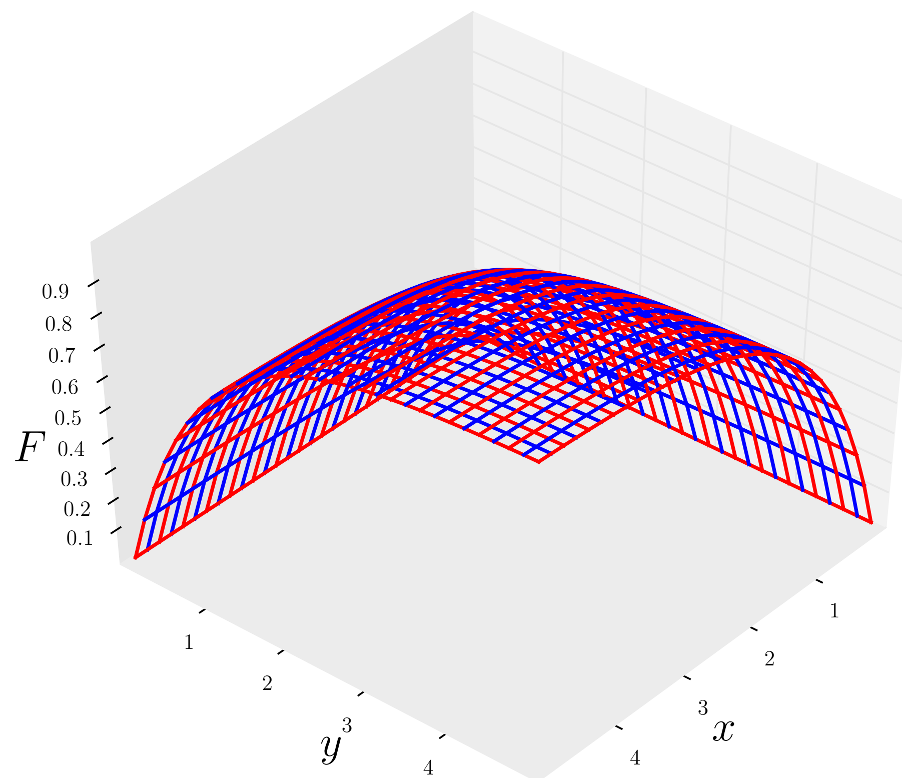
Let $$F_{X,Y}(x,y) = \begin{cases} (1-e^{-x})(1-e^{-y}) & \text{for } x \geq 0 \text{ and } y \geq 0 , \\ 0 & \text{otherwise} . \end{cases}$$

- Verify that

$$f_{X,Y}(x,y) = \frac{\partial^2 F}{\partial x \partial y} = \begin{cases} e^{-x-y} & \text{for } x \geq 0 \text{ and } y \geq 0 , \\ 0 & \text{otherwise} . \end{cases}$$



Density function $f_{X,Y}(x,y)$



Distribution function $F_{X,Y}(x,y)$

EXERCISE : (continued $\cdots$)

$$F_{X,Y}(x,y) = (1-e^{-x})(1-e^{-y}) \quad , \quad f_{X,Y}(x,y) = e^{-x-y} \; , \quad \text{for } x, y \geq 0 \, .$$

Also verify the following :

- $F(0,0) = 0 \quad , \quad F(\infty,\infty) = 1$,
- $\int_0^\infty \int_0^\infty f_{X,Y}(x,y)\,dx\,dy = 1$, (Why *zero* lower limits ?)
- $f_X(x) = \int_0^\infty e^{-x-y}\,dy = e^{-x}$,
- $f_Y(y) = \int_0^\infty e^{-x-y}\,dx = e^{-y}$.
- $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$. (So ?)

EXERCISE : (continued $\cdots$)

$$F_{X,Y}(x,y) = (1-e^{-x})(1-e^{-y}) \quad , \quad f_{X,Y}(x,y) = e^{-x-y} , \quad \text{for } x, y \geq 0 .$$

Also verify the following :

- $F_X(x) = \int_0^x f_X(x)\, dx = \int_0^x e^{-x}\, dx = 1 - e^{-x}$,
- $F_Y(y) = \int_0^y f_Y(y)\, dy = \int_0^y e^{-y}\, dy = 1 - e^{-y}$,
- $F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)$. (So ?)
- $P(1 < x < \infty) = F_X(\infty) - F_X(1) = 1-(1-e^{-1}) = e^{-1} \cong 0.37$,
- $P(\, 1 < x \leq 2 \, , \, 0 < y \leq 1 \,) = \int_0^1 \int_1^2 e^{-x-y}\, dx\, dy$

$$= (e^{-1} - e^{-2})(1 - e^{-1}) \cong 0.15 ,$$

Independent continuous random variables

Recall that two events E and F are *independent* if

$$P(EF) \; = \; P(E)\, P(F) \; .$$

Continuous random variables $X(s)$ and $Y(s)$ are *independent* if

$$P(X \in I_X \; , \; Y \in I_Y) \; = \; P(X \in I_X) \cdot P(Y \in I_Y) \; ,$$

for *all* allowable sets I_X and I_Y (typically *intervals*) of *real numbers*.

Equivalently, $X(s)$ and $Y(s)$ are independent if for all such sets I_X and I_Y the *events*

$$X^{-1}(I_X) \quad \text{and} \quad Y^{-1}(I_Y) \; ,$$

are independent *in the sample space* $\mathcal{S}$.

NOTE :

$$X^{-1}(I_X) \equiv \{s \in \mathcal{S} \; : \; X(s) \in I_X\} \; ,$$
$$Y^{-1}(I_Y) \equiv \{s \in \mathcal{S} \; : \; Y(s) \in I_Y\} \; .$$

FACT : $X(s)$ and $Y(s)$ are *independent* if for all x and y

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

EXAMPLE : The random variables with density function

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & \text{for } x \geq 0 \text{ and } y \geq 0 , \\ 0 & \text{otherwise} , \end{cases}$$

are *independent* because (by the preceding exercise)

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

NOTE :

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-x})(1-e^{-y}) & \text{for } x \geq 0 \text{ and } y \geq 0 , \\ 0 & \text{otherwise} , \end{cases}$$

also satisfies (by the preceding exercise)

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y) .$$

PROPERTY :

For *independent* continuous random variables X and Y we have

$$F_{X,Y}(x,y) \quad = \quad F_X(x) \cdot F_Y(y) \ , \qquad \text{for all } x, y \ .$$

PROOF :

$$\begin{aligned}
F_{X,Y}(x,y) \quad &= \quad P(X \leq x \ , \ Y \leq y) \\
&= \quad \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y) \ dy \ dx \\
&= \quad \int_{-\infty}^{x} \int_{-\infty}^{y} f_X(x) \cdot f_Y(y) \ dy \ dx \qquad \text{(by independence)} \\
&= \quad \int_{-\infty}^{x} [\ f_X(x) \cdot \int_{-\infty}^{y} f_Y(y) \ dy \ ] \ dx \\
&= \quad [\ \int_{-\infty}^{x} f_X(x) \ dx \ ] \ \cdot \ [\ \int_{-\infty}^{y} f_Y(y) \ dy \ ] \\
&= \quad F_X(x) \cdot F_Y(y) \ .
\end{aligned}$$

REMARK : Note how the proof parallels that for the discrete case !

Conditional distributions

Let X and Y be continuous random variables.

For given allowable sets I_X and I_Y (typically *intervals*), let

$$E_x = X^{-1}(I_X) \quad \text{and} \quad E_y = Y^{-1}(I_Y) ,$$

be their corresponding *events* in the sample space $\mathcal{S}$.

We have $$P(E_x|E_y) \equiv \frac{P(E_xE_y)}{P(E_y)} .$$

The *conditional probability density function* is defined as

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

When X and Y are *independent* then

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

(assuming $f_Y(y) \neq 0$).

EXAMPLE : The random variables with density function

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & \text{for } x \geq 0 \text{ and } y \geq 0 \,, \\ 0 & \text{otherwise} \,, \end{cases}$$

have (by previous exercise) the marginal density functions

$$f_X(x) = e^{-x} \quad , \quad f_Y(y) = e^{-y} \,,$$

for $x \geq 0$ and $y \geq 0$, and zero otherwise.

Thus for such x, y we have

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

$i.e.$, information about Y does not alter the density function of X .

Indeed, we have already seen that X and Y are *independent* .

Expectation

The *expected value* of a continuous random variable X is

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

which represents the *average value* of X over many trials.

The expected value of a *function of a random variable* is

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

The expected value of a function of *two* random variables is

$$E[g(X,Y)] \quad = \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \; f_{X,Y}(x,y) \; dy \; dx \; .$$

EXAMPLE :

For the *pointer* experiment

$$f_X(x) = \begin{cases} 0\,, & x \leq 0 \\ 1\,, & 0 < x \leq 1 \\ 0\,, & 1 < x \end{cases}$$

we have

$$E[X] = \int_{-\infty}^{\infty} x\ f_X(x)\ dx = \int_0^1 x\ dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}\,,$$

and

$$E[X^2] = \int_{-\infty}^{\infty} x^2\ f_X(x)\ dx = \int_0^1 x^2\ dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}\,.$$

EXAMPLE : For the joint density function

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & \text{for } x > 0 \text{ and } y > 0 \, , \\ 0 & \text{otherwise} \, . \end{cases}$$

we have (by previous exercise) the marginal density functions

$$f_X(x) = \begin{cases} e^{-x} & \text{for } x > 0 \, , \\ 0 & \text{otherwise} \, , \end{cases} \quad \text{and} \quad f_Y(y) = \begin{cases} e^{-y} & \text{for } y > 0 \, , \\ 0 & \text{otherwise} \, . \end{cases}$$

Thus $E[X] = \int_0^\infty x \, e^{-x} \, dx = -[(x+1)e^{-x}]\Big|_0^\infty = 1 \, .$ (Check !)

Similarly $E[Y] = \int_0^\infty y \, e^{-y} \, dy = 1 \, ,$

and

$$E[XY] = \int_0^\infty \int_0^\infty xy \, e^{-x-y} \, dy \, dx = 1 \, . \qquad (\text{ Check ! })$$

EXERCISE :

Prove the following for *continuous* random variables :

- $E[aX] \quad = \quad a\; E[X]$,

- $E[aX + b] \quad = \quad a\; E[X] \quad + \quad b$,

- $E[X + Y] \quad = \quad E[X] \quad + \quad E[Y]$,

and *compare* the proofs to those for *discrete* random variables.

EXERCISE :

A stick of length 1 is split at a randomly selected point X.
(Thus X is uniformly distributed in the interval $[0, 1]$.)
Determine the expected length of the piece containing the point $1/3$.

PROPERTY : If X and Y are *independent* then

$$E[XY] \quad = \quad E[X] \cdot E[Y] \, .$$

PROOF :

$$\begin{aligned}
E[XY] &= \int_{\mathbb{R}} \int_{\mathbb{R}} \; x \, y \, f_{X,Y}(x,y) \; dy \, dx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \; x \, y \, f_X(x) \, f_Y(y) \; dy \, dx \qquad \text{(by independence)} \\
&= \int_{\mathbb{R}} [\, x \, f_X(x) \; \int_{\mathbb{R}} \; y \, f_Y(y) \, dy \,] \, dx \\
&= [\int_{\mathbb{R}} \; x \, f_X(x) \, dx \,] \cdot [\int_{\mathbb{R}} \; y \, f_Y(y) \, dy \,] \\
&= E[X] \cdot E[Y] \, .
\end{aligned}$$

REMARK : Note how the proof parallels that for the discrete case !

EXAMPLE : For

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & \text{for } x > 0 \text{ and } y > 0 , \\ 0 & \text{otherwise} , \end{cases}$$

we already found

$$f_X(x) = e^{-x} \quad , \quad f_Y(y) = e^{-y} ,$$

so that

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

i.e., X and Y are *independent* .

Indeed, we also already found that

$$E[X] = E[Y] = E[XY] = 1 ,$$

so that

$$E[XY] = E[X] \cdot E[Y] .$$

Variance

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

Then the *variance* of the continuous random variable X is

$$Var(X) \equiv E[\,(X-\mu)^2] \equiv \int_{-\infty}^{\infty} (x-\mu)^2\, f_X(x)\, dx\ ,$$

which is the average weighted *square distance* from the mean.

As in the discrete case, we have

$$\begin{aligned} Var(X) &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - \mu^2\ . \end{aligned}$$

The *standard deviation* of X is

$$\sigma(X) \equiv \sqrt{Var(X)} = \sqrt{E[X^2] - \mu^2}\ .$$

which is the average weighted *distance* from the mean.

EXAMPLE : For $f(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \leq 0, \end{cases}$

we have

$$E[X] = \mu = \int_0^\infty x\, e^{-x}\, dx = 1 \quad (\text{ already done ! }),$$

$$E[X^2] = \int_0^\infty x^2\, e^{-x}\, dx = -\left[(x^2 + 2x + 2)e^{-x}\right]\Big|_0^\infty = 2,$$

$$Var(X) = E[X^2] - \mu^2 = 2 - 1^2 = 1,$$

$$\sigma(X) = \sqrt{Var(X)} = 1.$$

NOTE : The two integrals can be done by "*integration by parts*".

EXERCISE :

Also use the *Method of Moments* to compute $E[X]$ and $E[X^2]$.

EXERCISE : For the random variable X with density function

$$f(x) = \begin{cases} 0\ , & x \leq -1 \\ \text{c}\ , & -1 < x \leq 1 \\ 0\ , & x > 1 \end{cases}$$

- Determine the value of c
- Draw the graph of $f(x)$
- Determine the distribution function $F(x)$
- Draw the graph of $F(x)$
- Determine $E[X]$
- Compute $Var(X)$ and $\sigma(X)$
- Determine $P(X \leq -\frac{1}{2})$
- Determine $P(\mid X \mid \geq \frac{1}{2})$

EXERCISE : For the random variable X with density function

$$f(x) = \begin{cases} x+1\ , & -1 < x \leq 0 \\ 1-x\ , & 0 < x \leq 1 \\ 0\ , & \text{otherwise} \end{cases}$$

- Draw the graph of $f(x)$
- Verify that $\int_{-\infty}^{\infty} f(x)\ dx = 1$
- Determine the distribution function $F(x)$
- Draw the graph of $F(x)$
- Determine $E[X]$
- Compute $Var(X)$ and $\sigma(X)$
- Determine $P(X \geq \frac{1}{3})$
- Determine $P(\mid X \mid \leq \frac{1}{3})$

EXERCISE : For the random variable X with density function

$$f(x) = \begin{cases} \frac{3}{4}\left(1 - x^2\right), & -1 < x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

- Draw the graph of $f(x)$
- Verify that $\int_{-\infty}^{\infty} f(x)\, dx = 1$
- Determine the distribution function $F(x)$
- Draw the graph of $F(x)$
- Determine $E[X]$
- Compute $Var(X)$ and $\sigma(X)$
- Determine $P(X \leq 0)$
- Compute $P(X \geq \frac{2}{3})$
- Compute $P(\mid X \mid \geq \frac{2}{3})$

EXERCISE : Recall the density function

$$f_n(x) = \begin{cases} cx^n(1-x^n) \ , & 0 \le x \le 1 \\ \\ 0 \ , & \text{otherwise} \end{cases}$$

considered earlier, where n is a positive integer, and where

$$c = \frac{(n+1)(2n+1)}{n} \ .$$

- Determine $E[X]$.
- What happens to $E[X]$ for *large* n ?
- Determine $E[X^2]$
- What happens to $E[X^2]$ for *large* n ?
- What happens to $Var(X)$ for *large* n ?

Covariance

Let X and Y be continuous random variables with *mean*

$$E[X] = \mu_X \quad , \quad E[Y] = \mu_Y .$$

Then the *covariance* of X and Y is

$$Cov(X,Y) \equiv E[\,(X - \mu_X)\,(Y - \mu_Y)\,]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_X)\,(y - \mu_Y)\, f_{X,Y}(x,y)\, dy\, dx .$$

As in the discrete case, we have

$$\begin{aligned} Cov(X,Y) &= E[\,(X - \mu_X)\,(Y - \mu_Y)\,] \\ &= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \\ &= E[XY] - E[X]\,E[Y] . \end{aligned}$$

As in the discrete case, we also have

PROPERTY 1 :

- $Var(X+Y) \;=\; Var(X) + Var(Y) + 2\,Cov(X,Y)$,

and

PROPERTY 2 : If X and Y are *independent* then

- $Cov(X,Y) \;=\; 0$,

- $Var(X+Y) \;=\; Var(X) + Var(Y)$.

NOTE :

- The proofs are identical to those for the discrete case !

- As in the discrete case, if $Cov(X,Y) \;=\; 0$ then X and Y are not necessarily independent!

EXAMPLE : For

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & \text{for } x > 0 \text{ and } y > 0 , \\ 0 & \text{otherwise} , \end{cases}$$

we already found

$$f_X(x) = e^{-x} \quad , \quad f_Y(y) = e^{-y} ,$$

so that

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

$i.e.,$ X and Y are *independent* .

Indeed, we also already found

$$E[X] = E[Y] = E[XY] = 1 ,$$

so that

$$Cov(X,Y) = E[XY] - E[X]\,E[Y] = 0 .$$

EXERCISE :

Verify the following properties :

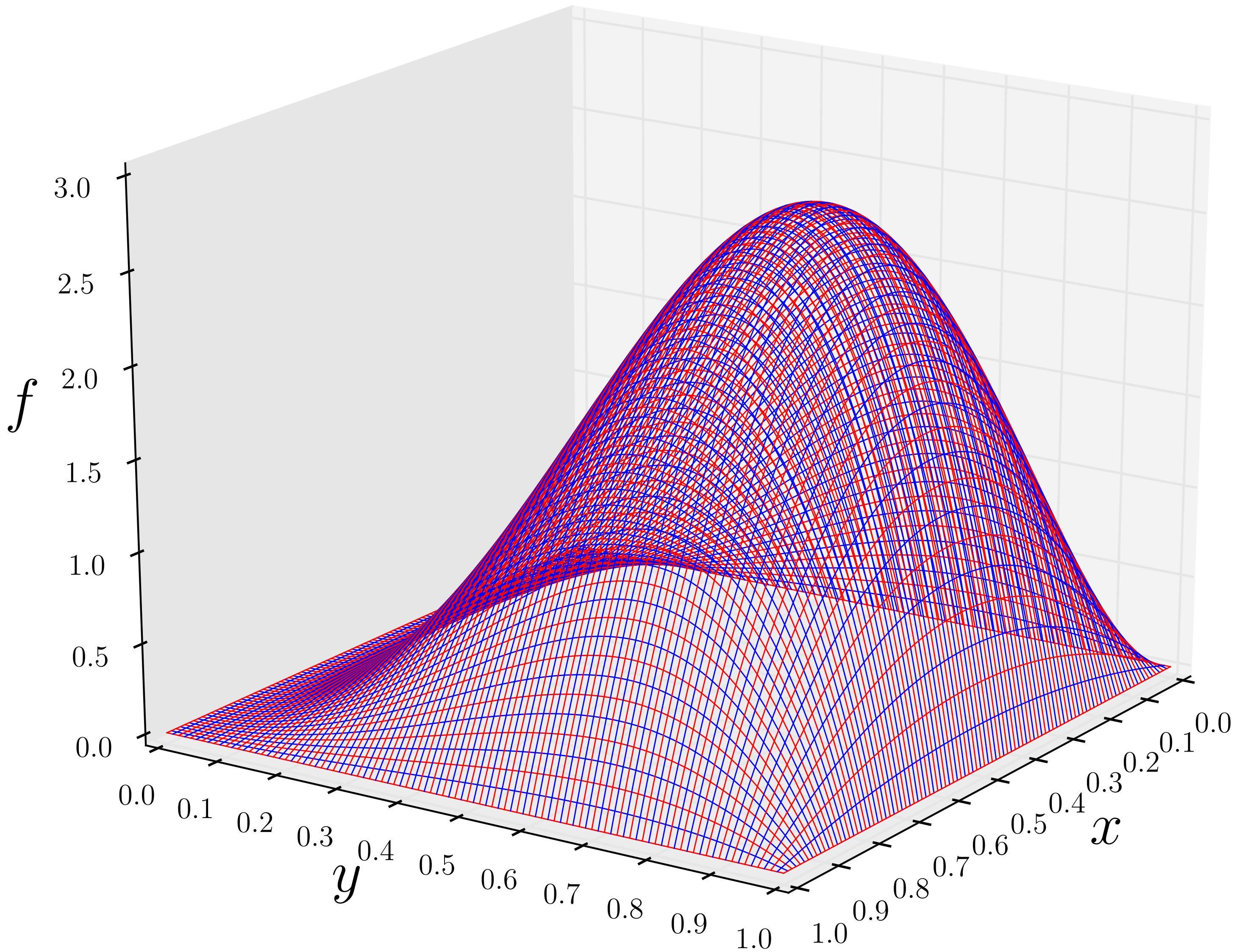
- $Var(cX + d) \; = \; c^2 \, Var(X)$,
- $Cov(X, Y) \; = \; Cov(Y, X)$,
- $Cov(cX, Y) \; = \; c \, Cov(X, Y)$,
- $Cov(X, cY) \; = \; c \, Cov(X, Y)$,
- $Cov(X + Y, Z) \; = \; Cov(X, Z) \; + \; Cov(Y, Z)$,
- $Var(X + Y) \; = \; Var(X) \; + \; Var(Y) \; + \; 2 \, Cov(X, Y)$.

EXERCISE :

For the random variables X , Y with *joint density function*

$$f(x,y) = \begin{cases} 45xy^2(1-x)(1-y^2) \ , & 0 \leq x \leq 1 \ , \ 0 \leq y \leq 1 \\ 0 \ , & \text{otherwise} \end{cases}$$

- Verify that $\int_0^1 \int_0^1 f(x,y)\ dy\ dx = 1$.
- Determine the *marginal density functions* $f_X(x)$ and $f_Y(y)$.
- Are X and Y *independent* ?
- What is the value of $Cov(X,Y)$?



The joint probability density function $f_{XY}(x, y)$.

Markov's inequality.

For a continuous *nonnegative* random variable X , and $c > 0$, we have

$$P(X \geq c) \quad \leq \quad \frac{E[X]}{c} \; .$$

PROOF :

$$E[X] \;=\; \int_0^\infty x f(x)\, dx \;=\; \int_0^c x f(x)\, dx \;+\; \int_c^\infty x f(x)\, dx$$

$$\geq \; \int_c^\infty x f(x)\, dx$$

$$\geq \; c \int_c^\infty f(x)\, dx \qquad \text{(Why ?)}$$

$$= \; c\, P(X \geq c) \; .$$

EXERCISE :

Show Markov's inequality also holds for *discrete* random variables.

Markov's inequality : For continuous *nonnegative* X , $c > 0$:

$$P(X \geq c) \quad \leq \quad \frac{E[X]}{c} \; .$$

EXAMPLE : For

$$f(x) \; = \; \begin{cases} e^{-x} & \text{for } x > 0 \; , \\ 0 & \text{otherwise} \; , \end{cases}$$

we have

$$E[X] \; = \; \int_0^\infty x \, e^{-x} \, dx \; = \; 1 \qquad (\text{ already done ! })$$

Markov's inequality gives

$$c = \mathbf{1} \; : \qquad P(X \geq \mathbf{1}) \quad \leq \quad \frac{E[X]}{\mathbf{1}} \; = \; \frac{1}{\mathbf{1}} \; = \; 1 \;\; (\; ! \;)$$

$$c = \mathbf{10} \; : \qquad P(X \geq \mathbf{10}) \quad \leq \quad \frac{E[X]}{\mathbf{10}} \; = \; \frac{1}{10} \; = \; 0.1$$

QUESTION : Are these estimates "*sharp*" ?

QUESTION : Are these estimates "*sharp*" ?

Markov's inequality gives

$$c = 1 \ : \qquad P(X \geq 1) \ \leq \ \frac{E[X]}{1} \ = \ \frac{1}{1} \ = \ 1 \ \ (\ !\ )$$

$$c = 10 \ : \qquad P(X \geq 10) \ \leq \ \frac{E[X]}{10} \ = \ \frac{1}{10} \ = \ 0.1$$

The actual values are

$$P(X \geq 1) \ = \ \int_1^\infty e^{-x} \, dx \ = \ e^{-1} \ \cong \ 0.37$$

$$P(X \geq 10) \ = \ \int_{10}^\infty e^{-x} \, dx \ = \ e^{-10} \ \cong \ 0.000045$$

EXERCISE : Suppose the score of students taking an examination is a random variable with mean 65 .

Give an upper bound on the probability that a student's score is greater than 75 .

Chebyshev's inequality: For (practically) any random variable X:

$$P(\ |X-\mu|\ \geq\ k\ \sigma\ )\ \leq\ \frac{1}{k^2}\ ,$$

where $\mu = E[X]$ is the *mean*, $\sigma = \sqrt{Var(X)}$ the *standard deviation*.

PROOF : Let $Y \equiv (X-\mu)^2$, which is nonnegative.

By Markov's inequality

$$P(Y \geq c)\ \leq\ \frac{E[Y]}{c}\ .$$

Taking $c = k^2\sigma^2$ we have

$$P(\ |X-\mu|\ \geq\ k\sigma\ )\ =\ P(\,(X-\mu)^2\ \geq\ k^2\sigma^2\,)\ =\ P(\,Y\ \geq\ k^2\sigma^2\,)$$

$$\leq\ \frac{E[\ Y\ ]}{k^2\sigma^2}\ =\ \frac{Var(X)}{k^2\sigma^2}\ =\ \frac{\sigma^2}{k^2\sigma^2}\ =\ \frac{1}{k^2}\ .\quad \textbf{QED}\ !$$

NOTE : This inequality also holds for discrete random variables.

EXAMPLE : Suppose the value of the Canadian dollar in terms of the US dollar over a certain period is a random variable X with

$$\textit{mean} \quad \mu = 0.98 \quad \text{and} \quad \textit{standard deviation} \quad \sigma = 0.05 \; .$$

What can be said of the probability that the Canadian dollar is valued

$$\text{between} \quad \$0.88\text{US} \quad \text{and} \quad \$1.08\text{US} \; ,$$

that is,

$$\text{between} \quad \mu - 2\sigma \quad \text{and} \quad \mu + 2\sigma \; ?$$

SOLUTION : By Chebyshev's inequality we have

$$P(\; | X - \mu | \; \geq \; 2 \, \sigma \;) \; \leq \; \frac{1}{2^2} \; = \; 0.25 \; .$$

Thus

$$P(\; | X - \mu | \; < \; 2 \, \sigma \;) \; > \; 1 \; - \; 0.25 \; = \; 0.75 \; ,$$

that is,

$$P(\; \$0.88\text{US} \; < \; \text{Can}\$ \; < \; \$1.08\text{US} \;) \; > \; 75 \; \% \; .$$

EXERCISE :

The score of students taking an examination is a random variable with mean $\mu = 65$ and standard deviation $\sigma = 5$.

- What is the probability a student scores between 55 and 75 ?
- How many students would have to take the examination so that the probability that their average grade is between 60 and 70 is at least 80% ?

HINT : Defining

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) , \quad (\text{ the average grade })$$

we have

$$\mu_{\bar{X}} = E[\bar{X}] = \frac{1}{n} n \mu = \mu = 65 ,$$

and, assuming independence,

$$Var(\bar{X}) = n \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n} = \frac{25}{n} , \quad \text{and} \quad \sigma_{\bar{X}} = \frac{5}{\sqrt{n}} .$$