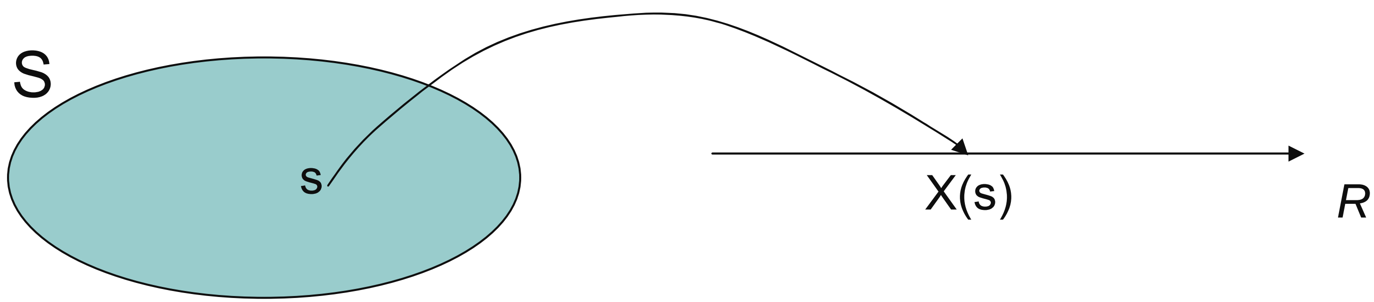
Probability theory
Random variables

In an experiment a number is often attached to each outcome.



Definition:
A **random variable** X is a function defined on S, which takes values on the real axis

$$X: \quad S \rightarrow R$$

Sample space — Real numbers

Probability theory
Random variables

Example:

Random variable	Type	
Number of eyes when rolling a dice	discrete	"counting"
The sum of eyes when rolling two dice	discrete	
Number of children in a family	discrete	
Age in years of first-time mother	discrete	
Time of running 5 km	continuous	measure
Amount of sugar in a coke	continuous	
Height of males	continuous	

Discrete: can take a finite number of values or an infinite but countable number of values.

Continuous: takes values from the set of real numbers.

Discrete random variable
Probability function

Definition:
Let $X : S \rightarrow R$ be a discrete random variable.

The function f(x) is a **probabilty function** for X, if

1. $f(x) \geq 0$ for all x
2. $\sum_x f(x) = 1$
3. $P(X = x) = f(x)$,

   where $P(X=x)$ is the probability for the outcomes $s \in S : X(s) = x$.

Discrete random variable
Probability function

Example: Flip three coins   X : # heads $X: S \rightarrow \{0,1,2,3\}$

Outcome	**Value of X**	**Probability function**
TTT	X=0	f(0) = P(X=0) = 1/8
HTT, TTH, THT	X=1	f(1) = P(X=1) = 3/8
HHT, HTH, THH	X=2	f(2) = P(X=2) = 3/8
HHH	X=3	f(3) = P(X=3) = 1/8

Notice! The definition of a probability function is fulfilled:

1. $f(x) \geq 0$
2. $\sum f(x) = 1$
3. $P(X=x) = f(x)$

Discrete random variable
Cumulative distribution function

Definition:
Let $X : S \to R$ be a discrete random variable with probability function f(x).

The cumulative distribution function for X, F(x), is defined by

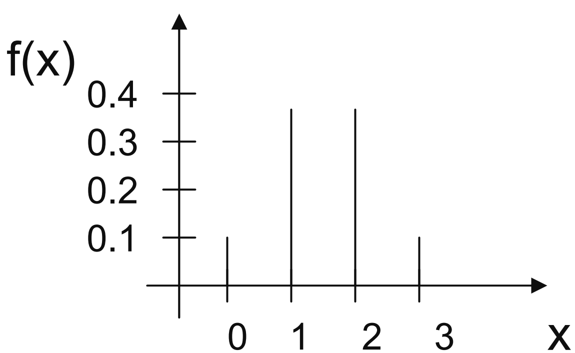
$$F(x) = P(X \le x) = \sum_{t \le x} f(t) \qquad \text{for } -\infty < x < \infty$$

Discrete random variable
Cumulative distribution function

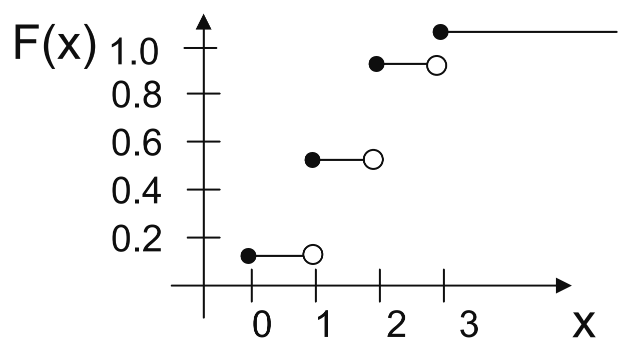
Example: Flip three coins   X : # heads $X : S \to \{0,1,2,3\}$

Outcome	**Value of X**	**Probability function**	**Cumulative dist. Func.**
TTT	X=0	f(0) = P(X=0) = 1/8	$F(0) = P(X \leq 0) = 1/8$
HTT, TTH, THT	X=1	f(1) = P(X=1) = 3/8	$F(1) = P(X \leq 1) = 4/8$
HHT, HTH, THH	X=2	f(2) = P(X=2) = 3/8	$F(2) = P(X \leq 2) = 7/8$
HHH	X=3	f(3) = P(X=3) = 1/8	$F(3) = P(X \leq 3) = 1$

Probability function:



Cumulative distribution function:



Continuous random variable

A continuous random variable X has probability 0 for all outcomes!!

Mathematically: $P(X = x) = f(x) = 0$ for all x

Hence, we cannot represent the probability function f(x) by a table or bar chart as in the case of discrete random variables.

Instead we use a continuous function – a density function.

Continuous random variable
Density function

Definition:

Let X: S $\rightarrow$ R be a continuous random variables.

A **probability density function** f(x) for X is defined by:

1. $f(x) \geq 0$ for all x
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. $P(a < X < b) = \int_a^b f(x)\,dx$

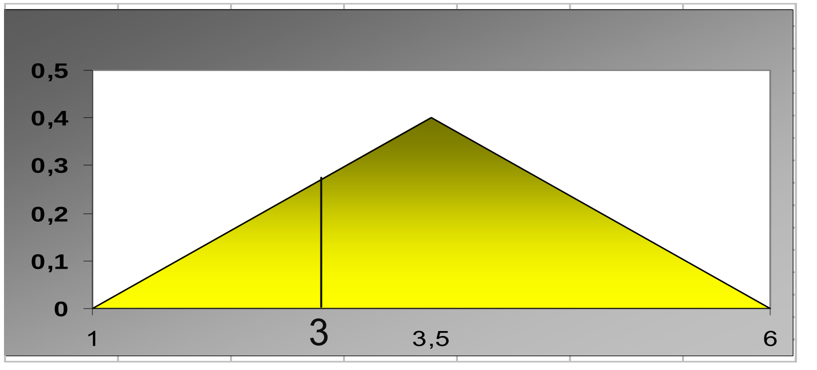
Note!! Continuity: $P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b)$

Continuous random variable
Density function

Example: X: service life of car battery in years (continuous)

Density function:

$$f(x) = \begin{cases} -0.16 + 0.16x & \text{for} \quad 1 \le x \le 3.5 \\ 0.96 - 0.16x & \text{for} \quad 3.5 < x \le 6 \\ 0 & \text{otherwise} \end{cases}$$



Probability of a service life longer than 3 years:

$$P(X > 3) = \int_3^\infty f(x)\,dx$$

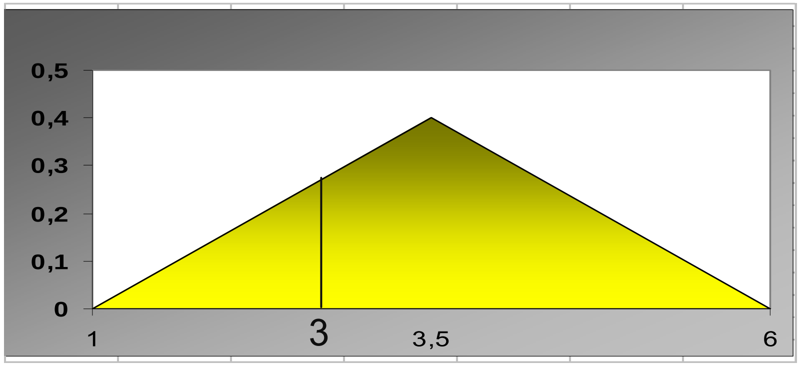
$$= \int_3^{3.5} (-0.16 + 0.16x)\,dx + \int_{3.5}^{6} (0.96 - 0.16x)\,dx$$

$$= \ldots = 0.68$$

Continuous random variable
Density function

Alternativ måde:

$$f(x) = \begin{cases} -0.16 + 0.16x & for \quad 1 \le x \le 3.5 \\ 0.96 - 0.16x & for \quad 3.5 < x \le 6 \\ 0 & othwerwise \end{cases}$$



Probability of a service life longer than 3 years:

$$P(X > 3) = 1 - P(X \le 3)$$

$$= 1 - \int_{-\infty}^{3} f(x)\,dx$$

$$= 1 - \int_{1}^{3} (-0.16 + 0.16x)\,dx$$

$$= \ldots = 1 - 0.32 = 0.68$$

Continuous random variable
Cumulative distribution function

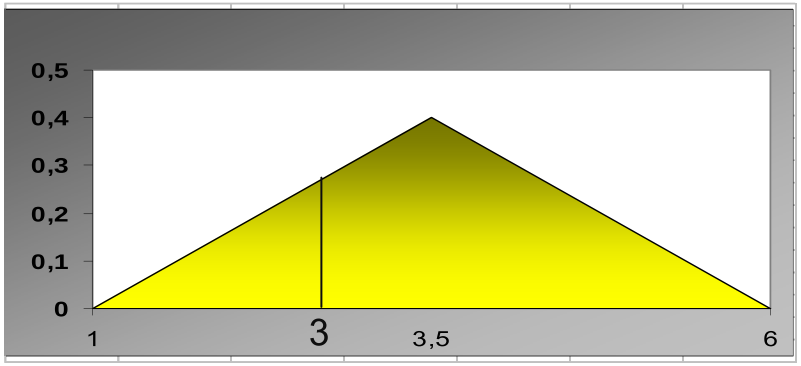
Definition:
Let $X : S \to R$ be a continuous random variable with density function f(x).

The cumulative distribution function for X, F(x), is defined by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt \quad \text{for } -\infty < x < \infty$$

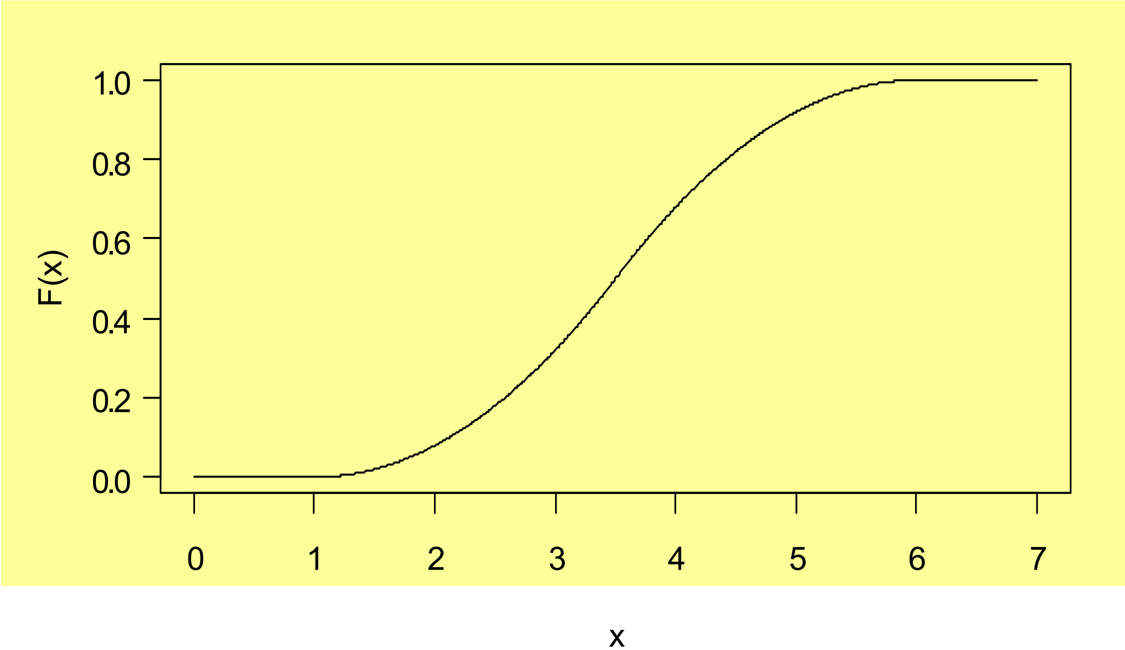
Note: $F'(x) = f(x)$

$$F(3) = P(X \leq 3)$$
$$= \int_{-\infty}^{3} f(x)\,dx$$
$$= \int_{1}^{3} -0.16 + 0.16x\,dx$$
$$= \ldots = 0.32$$



Continuous random variable
Cumulative distribution function

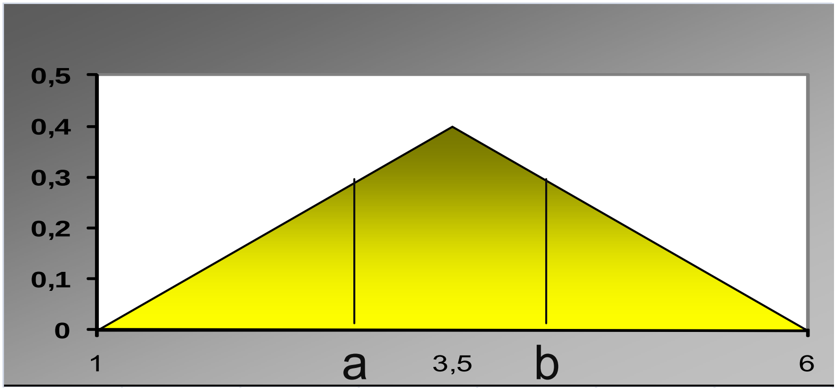
$$F(x)=\begin{cases} 0 & \text{for} \quad x<1 \\ 0.08-0.16x+0.08x^2 & \text{for} \quad 1\le x\le 3.5 \\ -1.88+0.96x-0.08x^2 & \text{for} \quad 3.5<x\le 6 \\ 1 & \text{for} \quad x>6 \end{cases}$$



Continuous random variable
Cumulative distribution function

From the definition of the cumulative distribution funct. we get:

$$P(a < X < b) = P(a \le X \le b) = P(a \le X < b) = P(a < X \le b)$$
$$= P(X \le b) - P(X \le a)$$
$$= F(b) - F(a)$$



Random variables

Discrete ~ Continous

Probability function ~ Density function

Discrete random variable

- Sample space is finite or has countable many outcomes
- Probability function f(x)
  Is often given by table
- Calculation of probabilities
  $$P(a < X < b) = \sum_{a<t<b} f(t)$$

Continuous random variable

- The sample space contains infinitely many outcomes
- Density function f(x) is a
  continuous function
- Calculation of probabilities
  $$P(a < X < b) = \int_a^b f(t)\, dt$$

Joint distribution
Joint probability function

Definition:
Let X and Y be two **discrete** random variables.
The **joint probability function** f(x,y) for X and Y
Is defined by

1. $f(x,y) \geq 0$ for all x og y
2. $\sum_{x} \sum_{y} f(x,y) = 1$
3. $P(X = x, Y = y) = f(x,y)$ (the probability that both X = x and Y=y)

For a set A in the xy plane: $$P((X,Y) \in A) = \sum \sum_{(x,y) \in A} f(x,y)$$

Joint distribution
Marginal probability function

Definition:
Let X and Y be two **discrete** random variables with joint probability function $f(x,y)$.
The **marginal probability function** for X is given by

$$g(x) = \sum_{y} f(x,y) \quad \text{for all } x$$

The **marginal probability function** for Y is given by

$$h(y) = \sum_{x} f(x,y) \quad \text{for all } y$$

Joint distribution
Marginal probability function

Example 3.14 (modified):
The joint probability function f(x,y) for X and Y is given by

y \ x	0	1	2
0	3/28	9/28	3/28
1	3/14	3/14	0
2	1/28	0	0

- $h(1) = P(Y = 1) = 3/14 + 3/14 + 0 = 3/7$
- $g(2) = P(X = 2) = 3/28 + 0 + 0 = 3/28$
- $P(X+Y < 2) = 3/28 + 9/28 + 3/14 = 18/28 = 9/14$

Joint distribution
Joint density function

Definition:
Let X og Y be two **continuous** random variables.
The **joint density function** f(x,y) for X and Y is defined by

1. $f(x,y) \geq 0$ for all x
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$
3. $P(a < X < b, c < Y < d) = \int_{c}^{d}\int_{a}^{b} f(x,y)\,dx\,dy$

For a region A in the xy-plane: $P[(X,Y)\in A] = \int\int_A f(x,y)\,dxdy$

Joint distribution
Marginal density function

Definition:
Let X and Y be two **continuous** random variables with joint density function f(x,y).
The **marginal density function** for X is given by

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \quad \text{for all } x$$

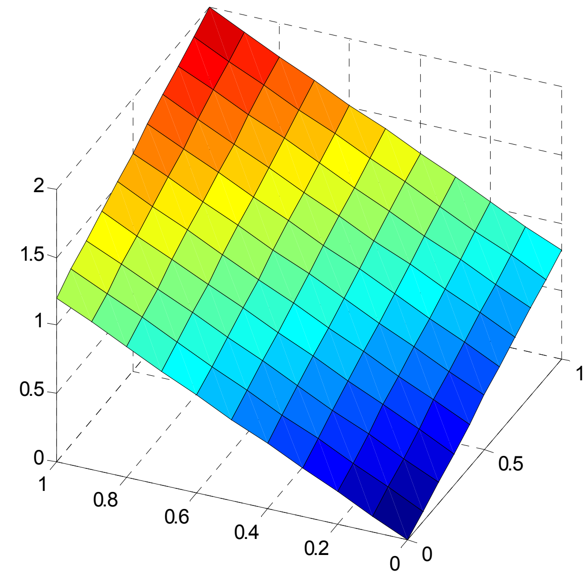
The **marginal density function** for Y is given by

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \quad \text{for all } y$$

Joint distribution
Marginal density function

Example 3.15 + 3.17 (modified):
Joint density f(x,y) for X and Y:

$$f(x,y)=\begin{cases}\frac{2}{5}(2x+3y) & 0\le x\le 1, 0\le y\le 1\\ 0 & \text{otherwise}\end{cases}$$



Marginal density function for X:

$$g(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$$

$$=\int_0^1 \frac{2}{5}(2x+3y)\,dy$$

$$=\left[\frac{2}{5}2xy+\frac{1}{5}3y^2\right]_0^1=\frac{4}{5}x+\frac{3}{5}$$

Joint distribution
Conditional density and probability functions

Definition:
Let X and Y be random variables (continuous or discrete) with joint density/probability function $f(x,y)$. Then the Conditional density/probability function for Y given X=x is

$$f(y|x) = f(x,y) / g(x) \quad g(x) \neq 0$$

where $g(x)$ is the marginal density/probability function for X, and the conditional density/probability function for X given Y=y is

$$f(x|y) = f(x,y) / h(y) \quad h(y) \neq 0$$

where $h(y)$ is the marginal density/probability function for Y.

Joint distribution
Conditional probability function

Examples 3.16 + 3.18 (modified):

Joint probability function
f(x,y) for X and Y is given by:

y \ x	0	1	2
0	3/28	9/28	3/28
1	3/14	3/14	0
2	1/28	0	0

- marginal pf.

$$g(x) = \begin{cases} \frac{10}{28} & \text{for} \quad x = 0 \\ \frac{15}{28} & \text{for} \quad x = 1 \\ \frac{3}{28} & \text{for} \quad x = 2 \end{cases}$$

- P(Y=1 | X=1) = f(1|1)
  = f(1,1) / g(1)
  = (3/14) / (15/28)
  = 6/15

Joint distribution
Independence

Definition:
Two random variables X and Y (continuous or discrete) with joint density/probability functions $f(x,y)$ and marginal density/probability functions $g(x)$ and $h(y)$, respectively, are said to be independent if and only if

$$f(x,y) = g(x)\,h(y) \qquad \text{for all } x,y$$

or if $f(x|y) = g(x)$ (x indep. of y) or $f(y|x)=h(y)$ (y indep. af x)

Mean / Expected value
Definition

Definition:
Let X be a random variable with probability / Density function f(x). The mean or expected value of X is give by

$$\mu = E(X) = \sum_{x} x\, f(x)$$

if X is discrete, and

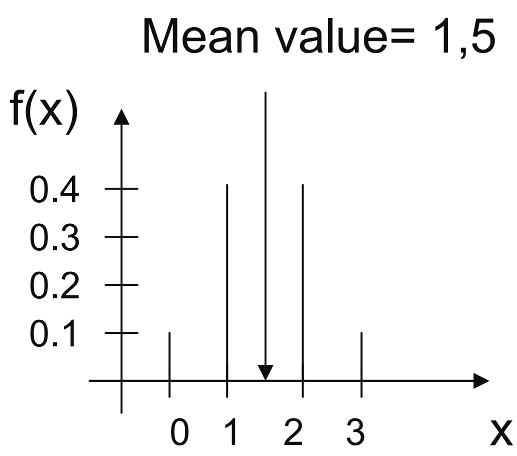
$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x) dx$$

if X is continuous.

Mean / Expected value
Interpretation

Interpretation:
The total contribution of a value multiplied by the probability of the value – a weighted average.

Example:



Mean / Expected value
Example

Problem:

- A private pilot wishes to insure his plane valued at 1 mill kr.
- The insurance company expects a loss with the following probabilities:
    - Total loss with probability 0.001
    - 50% loss with probability 0.01
    - 25% loss with probability 0.1

1. What is the expected loss in kroner ?
2. What premium should the insurance company ask if they want an expected profit of 3000 kr

Mean / Expected value
Function of a random variable

Theorem:
Let X be a random variable with probability / density function f(x). The expected value of g(X) is

$$\mu_{g(X)} = E[g(X)] = \sum_{x} g(x) f(x)$$

if X is discrete, and

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

if X is continuous.

Expected value
Linear combination

Theorem: Linear combination
Let X be a random variable (discrete or continuous), and let a and b be constants. For the random variable $aX + b$ we have

$$E(aX+b) = aE(X)+b$$

Mean / Expected value
Example

Problem:

- The pilot from before buys a new plane valued at 2 mill kr.
- The insurance company’s expected losses are unchanged:
  - Total loss with probability 0.001
  - 50% loss with probability 0.01
  - 25% loss with probability 0.1

1. What is the expected loss for the new plane?

Mean / Expected value
Function of a random variables

Definition:

Let X and Y be random variables with joint probability / density function f(x,y). The expected value of g(X,Y) is

$$\mu_{g(X,Y)} = E[g(X,Y)] = \sum_x \sum_y g(x,y) f(x,y)$$

if X and Y are discrete, and

$$\mu_{g(X,Y)} = E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) \, dx \, dy$$

if X and Y are continuous.

Mean / Expected value
Function of two random variables

Problem:

Burger King sells both via “drive-in” and “walk-in”.

Let X and Y be the fractions of the opening hours that “drive-in” and “walk-in” are busy.

Assume that the joint density for X and Y is given by

$$f(x,y) = \begin{cases} 4xy & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

The turn over g(X,Y) on a single day is given by

$$g(X,Y) = 6000\,X + 9000Y$$

What is the expected turn over on a single day?

Mean / Expected value
Sums and products

Theorem: Sum/Product

Let X and Y be random variables then

$$E[X+Y] = E[X] + E[Y]$$

If X and Y are independent then

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