2. FUNDAMENTALS OF PROBABILITY THEORY

P. Thoft-Christensen, M. J. Baker

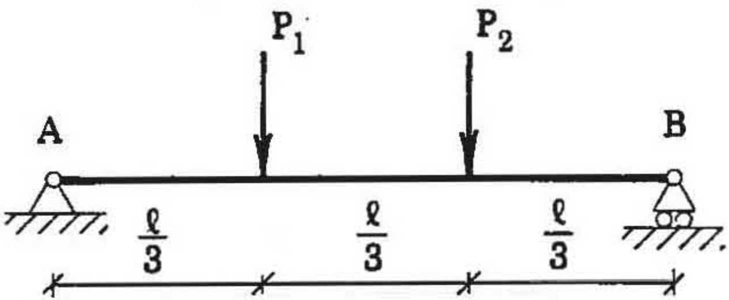
Structural reliability theory and its applications, 1982

1.1 SAMPLE SPACE

A standard way to determining e.g. the yield stress of steel is to perform a number of simple tensile specimen tests.
Each test records a steel yield stress value, we expect values varying from test to test.
Thus the yield stress is considered uncertain
- it is a *random* parameter.
The set of all possible outcomes of such tests is called *sample space*, each individual outcome is a *sample point*.
The yield stress sample space is the set of all positive real numbers, the number of sample points is infinite – a *continuous* sample space.
A sample space may also be *discrete*, when the number of sample points is finite of countable.

Example 2.1

A simply supported beam *AB* is loaded by point forces P_1 and P_2.



Let the possible values of P_1 and P_2 be 4, 5, 6 and 3, 4 [kN], respectively. The sample space for the loading will then be the set

$\Omega = \{(4, 3), (4, 4), (5, 3), (5, 4), (6, 3), (6, 4)\}$

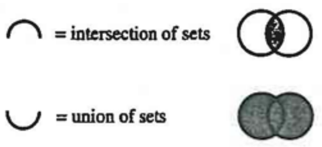
This is a discrete sample space. For its finite number of sample points we it a *finite* sample space.

A sample space with the countable infinite number of sample points is called discrete *infinite* sample space.

The sample spaces for the loads P_1 and P_2 are $\Omega_1 = \{4, 5, 6\}$ and $\Omega_2 = \{3, 4\}$, respectively. Note that $\Omega = \Omega_1 \times \Omega_2$. Show as an exercise that the sample space for the reaction R_A in point A is $\Omega_A = \{11/3, 12/3, 13/3, 14/3, 15/3, 16/3\}$.

An event (*zdarzenie*) is any sample space subset (*podzbiór*), being therefore a set of sample points.
An event of no sample points is an *impossible event (zdarzenie niemożliwe)*. An event of all sample points is a *certain event (zdarzenie pewne)*, being a sample space itself.
Example 2.2
The beam in Fig. 2.1 revisited. The sample space for the reaction R_A is $\Omega_A = \{11/3, 12/3, 13/3, 14/3, 15/3, 16/3\}$.
The subset $\{15/3, 16/3\}$ is the event that R_A is equal to 15/3 or 16/3.
Let E_1 and E_2 be two events.
The *union (suma, alternatywa)* $E_1 \cup E_2$ of E_1 and E_2 is subset of sample points that belong to E_1 and/or E_2.

The *intersection (iloczyn lub koniunkcja)* $E_1 \cap E_2$ of E_1 and E_2 is a subset of sample points belonging to both E_1 and E_2.

	A Venn diagram showing the difference between the intersection and the union of events

Two events E_1 and E_2 are *mutually exclusive (rozłączne, wyklucza-jące się)* if they are disjoint (have no common sample points).
In this case $E_1 \cap E_2 = \varnothing$,where $\varnothing$ is an *impossible event* (an empty set) (*zdarzenie niemożliwe*).
Let Ω be a sample space and E an event.
The event containing all sample points of Ω not included in E is called the *complementary event to E (zdarzenie przeciwne)*, denoted by $\bar{E}$. Obviously, $\bar{E} \cup E = \Omega$ and $\bar{E} \cap E = \varnothing$.

Intersection and union operations follow the commutative (*przemienność*), associative (*łączność*), and distributive (*rozdzielność*) laws:

$$E_1 \cap E_2 = E_2 \cap E_1 \quad \text{and} \quad E_1 \cup E_2 = E_2 \cup E_1 \quad \text{commutative}$$

$$E_1 \cap (E_2 \cap E_3) = (E_1 \cap E_2) \cap E_3 \quad \text{associative}$$

$$E_1 \cup (E_2 \cup E_3) = (E_1 \cup E_2) \cup E_3$$

$$E_1 \cap (E_2 \cup E_3) = (E_1 \cap E_2) \cup (E_1 \cap E_3) \quad \text{distributive}$$

$$E_1 \cup (E_2 \cap E_3) = (E_1 \cup E_2) \cap (E_1 \cup E_3)$$

Due to these laws intersection or union of a set of events E_1, E_2, ... , E_n may be stated, as follows:

$$\bigcap_{i=1}^{n} E_i = E_1 \cap E_2 \cap \ \dots \ \cap E_3 , \qquad \bigcup_{i=1}^{n} E_i = E_1 \cup E_2 \cup \ \dots \ \cup E_3$$

De Morgan's laws:

$$E_1 \cap E_2 = \overline{E_1 \cup E_2}$$

$$E_1 \cup E_2 = \overline{E_1 \cap E_2}$$

1.2 Axioms and theorems of probability theory

Axiom 1

For any event *E*

$$0 \le P(E) \le 1 \qquad (0.1)$$

where the function *P* is the *probability measure (miara prawdopodobieństwa)*. $P(E)$ is the probability of the event *E*.

Axiom 2

Let the sample space be Ω. Then

$$P(\Omega) = 1 \qquad (0.2)$$

Axiom 3

If $E_1, E_2, \ldots, E_n$ are mutually exclusive events (*wydarzenia wzajemnie wykluczające się)* then

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) \qquad (0.3)$$

The following theorems can be proved

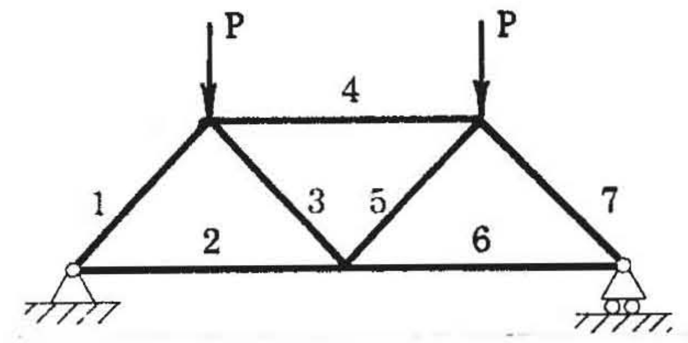
$$P(\bar{E}) = 1 - P(E)$$

$$P(\emptyset) = 0$$

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Example 2.3

Consider the statically determinate structural system (a truss) with 7 elements shown in Figure.



Let the event that element "i" fails be denoted by F_i
and let the probability of failure of element "i" be $P\left(F_i\right)$.
Further assume that failures of the individual members are statistically independent, that is $P(F_i \cap F_j) = P(F_i) \cdot P(F_j)$ for any pair of $(i,\ j)$. The failure of any member will result in system failure for a statically determinate structure. Thus

$$P\left(\text{failure of structure}\right) = P\left(F_1 \cup ... \cup F_7\right) = P\left(\bigcup_{i=1}^{7} F_i\right) =$$

$$= 1 - P\left(\overline{\bigcup_{i=1}^{7} F_i}\right) = 1 - P\left(\bigcap_{i=1}^{7} \overline{F_i}\right)$$

according to De Morgan's law.
Statistical independence assumption leads to the following

$$P(\text{failure of structure}) = 1 - P(\bar{F}_1) \cdot P(\bar{F}_2) \cdot P(\bar{F}_3) \ldots P(\bar{F}_7) =$$

$$= 1 - (1 - P(F_1))(1 - P(F_2)) \ldots (1 - P(F_7))$$

Let $P(F_1) = P(F_3) = P(F_5) = P(F_7) = 0.02$, $P(F_2) = P(F_6) = 0.01$, and $P(F_4) = 0.03$. Then

$$P(\text{failure of structure}) = 1 - 0.98^4 \cdot 0.99^2 \cdot 0.97 = 1 - 0.8769 = 0.1231$$

In practical applications the probability of occurrence of event E_1 conditional upon the occurrence of event E_2, is of great interest. This probability, called the *conditional probability*, is denoted $P(E_1|E_2)$ and defined by

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)} \qquad (0.4)$$

if $P(E_2) > 0$

Conditional probability is not defined for $P(E_2)=0$.

Event E_1 is *statistically independent* of event E_1 if

$$P(E_1|E_2)=P(E_1)$$

that is, if occurrence of E_2 does not affect the probability of E_1.

The probability of the event $P(E_1 \cap E_2)$is expressed by

$$P(E_1 \cap E_2)=P(E_1|E_2)P(E_2)=P(E_2|E_1)P(E_1) \quad (0.5)$$

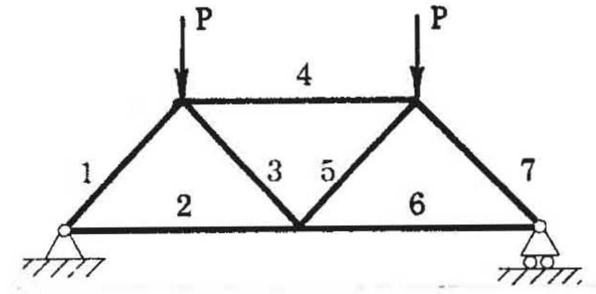
If E_1 and E_2, are statistically independent,(0.5) becomes

$$P(E_1 \cap E_2)=P(E_1)P(E_2) \quad (0.6)$$

The rule (0.6) is called *multiplication* rule and has already been used in example 2.3

Example 2.4.

Consider again the structure in Figure 2.2.



To simplify we assumeonly elements 2 and 6 failure-prone. Therefore

$$P(\text{failure of structure}) = P(F_2 \cup F_6) =$$

$$= P(F_2) + P(F_6) - P(F_2 \cap F_6) = \qquad (0.7)$$

$$= P(F_2) + P(F_6) - P(F_2 | F_6) P(F_6)$$

If F_2 and F_6 are statistically independent, as in example 2.3, and if $P(F_2) = P(F_6) = 0.01$ then

$$P(\text{failure of structure}) = 0.01 + 0.01 - 0.01 \cdot 0.01 = 0.0199$$

If F_2 and F_6 are not independent then the $P(F_2 | F_6)$ is required.

If a pair of elements is fabricated from the same steel bar it is reasonable to assume the same strength for both. Further, they are equally loaded, so in this special case, $P(F_2 | F_6)$ is close to 1.

Having $P(F_2 | F_6) = 1$ the(0.7) yields

$$P(\text{failure of structure}) = 0.01 + 0.01 - 1 \cdot 0.01 = 0.0100$$

EXAMPLE 2.11.(*A.S. Nowak, K.R. Collins, Reliability of structures*).

Consider tests of concrete beams. Two parameters are observed: cracking moment and ultimate moment. Let M_u and M_{cr} denote the ultimate bending moment and the cracking moment, respectively.

Define event E_1 by $M_u \geq 150$ kip-feet (k-ft)

and the event E_2 by $M_{cr} \geq 100$ k-ft.

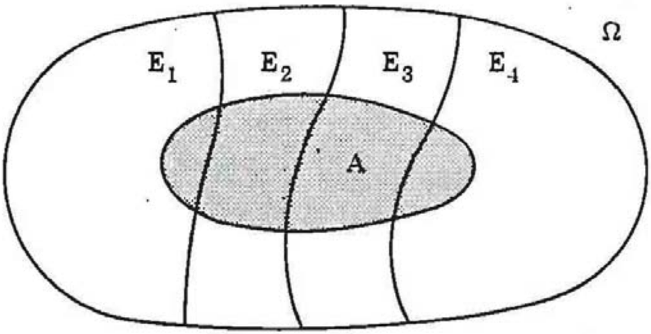
A conditional probability that the ultimate moment will be reached given the cracking moment is reached is expressed by:

$$P\left(E_1 \middle| E_2\right) = p\left(M_u \geq 150 \text{ given } M_{cr} \geq 100\right) = \frac{P\left(E_1 \cap E_2\right)}{P\left(E_2\right)} =$$

$$= \frac{P\left(M_u \geq 150 \text{ AND } M_{cr} \geq 100\right)}{M_{cr} \geq 100}$$

Bayes' theorem

Let the sample space Ω be divided into n mutually exclusive events $E_1, E_2, \ldots, E_n$ (see figure, where n = 4).



Let A be an event in the same sample space. Then

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + \ldots + P(A \cap E_n) =$$

$$= P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \ldots + P(A|E_n)P(E_n) \qquad (0.8)$$

$$= \sum_{i=1}^{n} P(A|E_i)P(E_i)$$

From the definition (0.4) $P\left(E_1 \middle| E_2\right) = \dfrac{P\left(E_1 \cap E_2\right)}{P\left(E_2\right)}$ follows

$$P\left(A \middle| E_i\right) P\left(E_i\right) = P\left(E_i \middle| A\right) P\left(A\right)$$

so that

$$P\left(E_i \middle| A\right) = \frac{P\left(A \middle| E_i\right) P\left(E_i\right)}{P\left(A\right)}$$

or using (0.8)

$$P\left(E_i \middle| A\right) = \frac{P\left(A \middle| E_i\right) P\left(E_i\right)}{\sum_{j=1}^{n} P\left(A \middle| E_j\right) P\left(E_j\right)}$$

This is the Bayes' theorem of high importance.

Example 2.5.

Assume that a steel girder is tested before application. Experience shows that 95% of all girders pass the test, but the test is only 90% reliable, so there is margin of 0.1 for this test of being erroneous. How probable is that a perfect girder will pass the test? Let E denote that the girder is perfect and let A be the event that it passes the test.

$P(E|A)=0.90$ and $P(E|\overline{A})+P(E|A)=1$

so that $P(E|\overline{A})=1-0.90=0.10$

Experience gives $P(A)=0.95$. The problem is to find $P(A|E)$. The events A and $\overline{A}$ are mutually exclusive, so according to (0.8)

$$P(E)=P(E|A)P(A)+P(E|\overline{A})P(\overline{A})=0.90\cdot 0.95+0.10\cdot 0.05=0.860$$

Finally, $P(A|E)=\dfrac{P(E|A)\cdot P(A)}{P(E)}=\dfrac{0.90\cdot 0.95}{0.86}=0.994$

Example 2.6.

Consider a number of tensile specimens supporting a load of2 kN. Estimate the probability that a specimen supports aload of 2.5 kN. Based on previous experiments the probability of 0.80 is estimated that a specimen can carry 2.5 kN.

Further we know that 50% ofspecimens not able to support 2.5 kN fail at loads less than 2.3 kN.

The probability of 0.80 mentioned above can now be updated if the following test issuccessful.

A single specimen is loaded to 2.3 kN.

Let E be the event that the specimen can support 2.5 kN and A the event that the test is successful (the specimen can support 2.3 kN).

Then $P\left(\bar{E}\middle|\bar{A}\right)=0.50$, and $P\left(E\right)=0.80$.Further $P\left(A\middle|E\right)=1.0$ so that Bayes' theorem gives

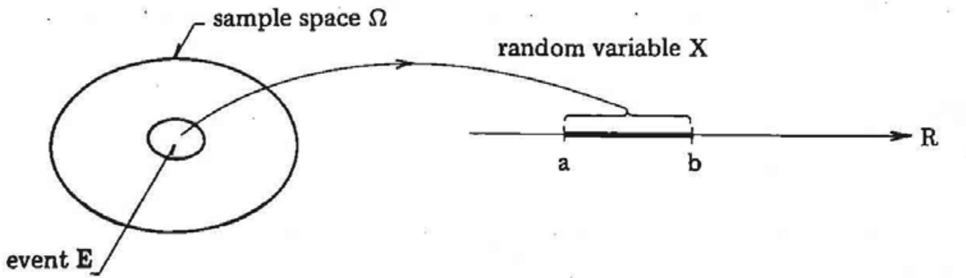
$$P\left(E\middle|A\right)=\frac{P\left(A\middle|E\right)\cdot P\left(E\right)}{P\left(A\middle|E\right)\cdot P\left(E\right)+P\left(A\middle|\bar{E}\right)\cdot P\left(\bar{E}\right)}=\frac{1.0\cdot 0.80}{1.0\cdot 0.80+0.5\cdot 0.20}=0.89$$

The previous value of 0.80 of the probability that a specimen can carry 2.5 kN is updated to 0.89, by means of Bayes' theorem.

1.3 Random variables

The outcomes of experiments are numerical values in most cases. Of course, there are exceptions. Checking a structure to carry a given load the outcome may be **yes or no**.
In such a case it is possible to present the outcome, e.g. **1** forthe event of carrying the load and **0** to the event of collapse.
The numbers 0 and 1 are virtually assigned, to be easily replacedby other values. Generally, it is possible to identify possible outcomes of a random phenomenon by numerical values.
In most cases the values will simply be outcomes of a phenomenon but it may be necessary to state the numerical values artificially.
An outcome or event is denoted by a value of a function called a *random variable (zmienna losowa)*.
A random variable is a function which maps events of the sample space Ω into the real line R.

A random variable is denoted by a capital letter, e.g.X. Pointing out the domain of X the random variable we write $X : \Omega \rightarrow R$.
A concept of continuous random variable is presented in Figure



The function X is a mapping(*odwzorowanie*) of a sample space into the interval $[a, b] \subset R$

If the sample space is discrete we enter a discrete random variable. In the previous section the probability of an event E is denoted by a probability measure P.

In this section it is shown how a numerical value is associated with any event by the random variable. This leads a convenient analytical and graphical description of events and associated probabilities. The argument ω in $X(\omega)$ is usually dropped. Similarly, the abbreviation $P(X \le x)$ is used for $P(\{\omega : X(\omega)\} \le x)$.

Discrete random variable X

Discrete random variableisa function whose set of values is finite or countable infinite.
A discrete random variable is described by a *probability mass function (PMF) (funkcja rozkładu prawdopodobieństwa)* p_X :

$$p_X(x) = P(X = x) \qquad (0.9)$$

X is the random variable, $x = x_1, x_2, ..., x_n$,
n can be finite or countable infinite.

$p_X(x)$ - probability that a discrete random variable X is equal to a specific value x,where x is a real number.
Different symbols are used for the random variable and its values, namely X and x, respectively.
It is a direct consequence of the axioms (0.1) - (0.3) that

$$0 \le p_X(x) \le 1$$

$$\sum_{i=1}^{n} p_X(x) = 1$$

$$P(a \le X \le b) = \sum_{x_i \le b} p_X(x_i) - \sum_{x_1 \le = a} p_X(x_i)$$

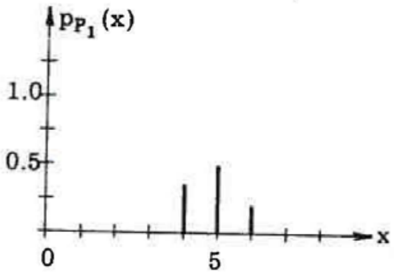
The probability distribution function or cumulative distribution function CDF*(dystrybuanta)* $P_X : R \to R$ is related to p_X by

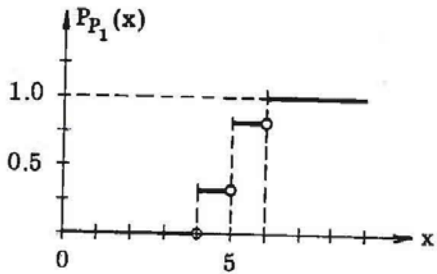
$$P_X(x) = P(X \le x) = \sum_{x_i \le x} p_X(x_i) \qquad (0.10)$$

The definition (0.10)expresses $P_X(x)$ the probability of the event that the random variable X takes on values equal to or less than x.

Example 2.7.

Consider again example 2.1 and let $P(P_1 = 4) = 0.3, P(P_1 = 5) = 0.5$ and $P(P_1 = 6) = 0.2$. Both probability mass functions p_{P_1} and the probability distribution function P_{P_1} for the random variable P_1 are shown in Fig.





Note that the circled points are not included in $P_{P_1}(x)$.

Continuous random variable X

A continuous random variableis a function of any possible value within one or several intervals, the sample space is infinite.
The probability of a specific value of a continuous random variable is zero. Thus the probability mass function defined in (0.9) for discrete random variables is no longer valid.
However, the probability distribution function, called cumulative distribution function (CDF) *(dystrybuanta)* $F_X : R \to R$ can still be defined by

$$F_X(x) = P(X \le x), \quad x \in R$$

The derivative probability function is used for continuous random variables. This function is called *probability density function(PDF)(funkcja gęstości prawdopodobieństwa)* $f_X : R \to R$, defined by

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{0.11}$$

assuming that the derivative exists.

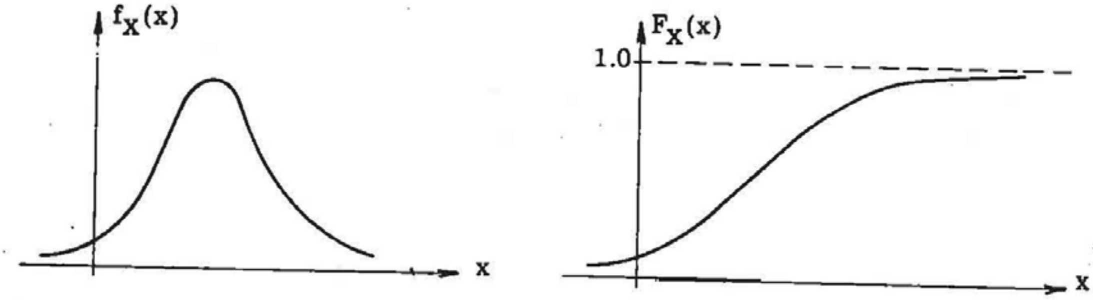
Note that the symbol $p_X(x)$ is used for the probability mass function (discrete variables only) and the symbol $f_X(x)$ for the probability density function (continuous variables only).

Inversion of the equation (0.11) gives

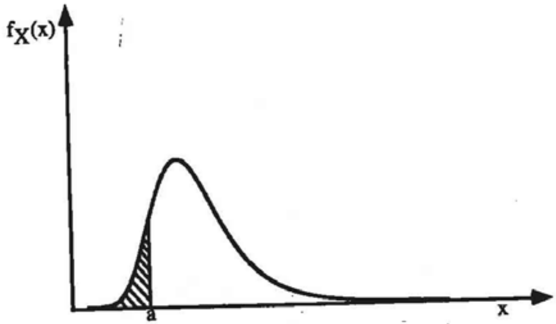
$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

for continuous random variables.

Figure shows the probability density function $f_X(x)$ and the probability distribution function F_X for a continuous random variable *X*.



Example of a PDF and example of a CDF



Graphical relation between PDF and CDF (Fig.):
for a given x PDF value is an ordinate, CDF is the shaded area.

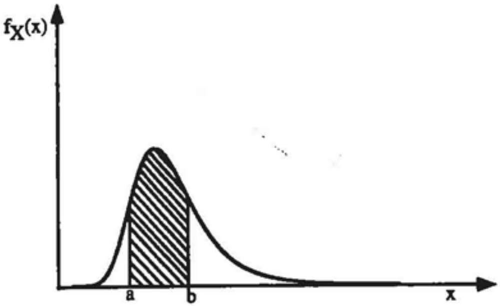
Properties of probability functions (CDF, PDF and PMF)

Here are several important properties of the CDF, sny function satisfying these six conditions may be considered a CDF.

1. The CDF definition is the same for both discrete and continuous randomvariables,
2. The CDF is a positive, nondecreasing function whose value lies between 0 and 1, including boundaries, $0 \le F_X(x) \le 1$,
3. F_X is non-decreasing, thus if $x_1 < x_2$ then $F(x_1) \le F(x_2)$
4. the limit $F_X(-\infty) = 0$
5. the limit $F_X(\infty) = 1$, or $\int_{-\infty}^{\infty} f_X(t)\,dt = F_X(\infty) = 1$
6. For a continuous random variable it holds

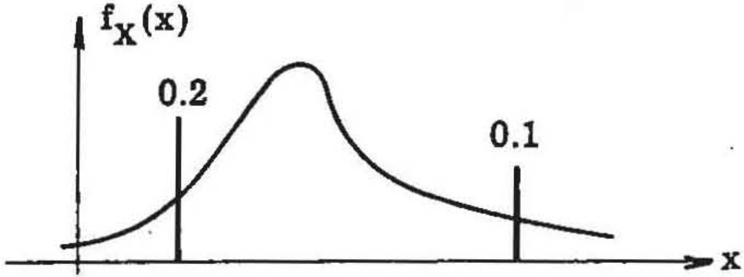
$$P(a \le x \le b) = F_X(b) - F_X(a) = \int_a^b f_X(t)\,dt \qquad (0.12)$$

The Figure below provides a graphical interpretation of Eq. (0.12)



Graphical representation of $F_X(b) - F_X(a)$ in Eq. (0.12)

Sometimes a *mixed continuous-discrete random variable*is used, i.e. a continuous random variable added a countable number of non-zero probability as shown in Figure below.



In this case the area under the curve is equal to 1 – 0.2 – 0.1 = 0.7.

1.4 Parameters of random variables - moments

Let X be a continuous random variable of a CDF function F_X .
However, in many cases the analytical form of F_X is not known.
An **approximate description of a random variable** is derived,capturing its most important features. Having F_X or f_X known,
it is also convenient to have a simplified random description.
The so-called random variable*moments*are introduced.

Assuming X a random variable $Y = X^k$ is a random variable too (k is a positive integer) for $P\left(\left\{\omega : X^k \le y\right\}\right)$ exists for every y.

The expected value *(wartość oczekiwana)* is defined for a continuous random variable X – integral form

$$E\left[X\right] = \int_{-\infty}^{\infty} x f_X\left(x\right) dx \qquad (0.13)$$

and for a discrete random variable X – summation form

$$E\left[X\right] = \sum_{x_i} x_i \, p_X\left(x_i\right) \qquad (0.14)$$

The expected value is also called the *mean value (wartość średnia, ensemble average)* or the *first moment* of X, denoted by μ_X :

$$E\left[X\right] = \mu_X$$

The n-th *moment* of X is called $E\left[X^n\right]$ and is defined below

$$E\left[X^n\right] = \int_{-\infty}^{\infty} x^n f_X\left(x\right) dx \qquad (0.15)$$

$$E\left[X^n\right] = \sum_{x_i} x_i^n p_X(x_i)$$ for both types of random variables

The *variance (wariancja)* of *X*, denoted σ_X^2, is the expected value of $(X - \mu_X)$, equal to

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\,dx \qquad (0.16)$$

$$\sigma_X^2 = \sum_{x_i} (x_i - \mu_X)^2 p_X(x_i)$$ for both types of random variables

Note that the first moment of *X* in (0.13) corresponds to the location of the centroid of a unit mass. Likewise, the second moment can be compared with the unit mass moment of inertia.

An important relationship exists among the mean, variance, and second moment of a random variable *X*:

$$\sigma_X^2 = E\left(X^2\right) - \mu_X^2$$

The *standard deviation* of *X (odchylenie standardowe)* is a positive square root of the variance:

$$\sigma_X = \sqrt{\sigma_X^2}$$

The standard deviation σ_X is a measure of scatter of the variable *X*values around the expected value $E\left[X\right]$. It is difficult to call the dispersion high or low on the basis of σ_X only.

A non-dimensional *coefficient of variation*, V_X *(współczynnik zmienności)* of *X* is its standard deviation in the units of its mean:

$$V_X = \frac{\sigma_X}{\mu_X}$$

This parameter (often denoted c.o.v.) is always taken positive even though the mean may be negative.

1.5 Sample parameters

Parameters defined in the previous section are theoretical properties of random variables, based on theirknown probability distributions.

Many practical applications do not state these distributions, so the need arises to estimate parameters using test data.

If a set of n observations $\{x_1, x_2, ..., x_n\}$ represents a random variable X, then its mean μ_X may be estimated by a *sample mean* $\overline{x}$ standard deviation μ_X - by the *sample standard deviation* s_X.

The sample mean

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \left(\mu_X = \sum_{x_i} x_i \, p_X\left(x_i\right) \; p_X\left(x_i\right) = \frac{1}{n} \right)$$

The sample standard deviation

$$s_X = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i^2\right) - n\left(\overline{x}\right)^2}{n-1}}$$

Sample variance is defined too

$$s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \qquad (\sigma_X^2 = \sum_{x_i}\left(x_i - \mu_X\right)^2 p_X\left(x_i\right), \quad p_X\left(x_i\right) = \frac{1}{n})$$

$$E\left(\overline{X}_n\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n} nE\left(X\right) = E\left(X\right)$$

n estimator (estimator)

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2\overline{X}_n X_i + \overline{X}_n^2\right) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2\right) - \overline{X}_n^2$$

$$E\left(S_n^2\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^2\right) - E\left(\overline{X}_n^2\right) = E\left(X^2\right) - E\left(\overline{X}_n^2\right)$$

$$E\left(\bar{X}_n^2\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \frac{1}{n^2} E\left(\sum_{i=1}^{n} X_i^2 + \sum_{i \neq j}^{n} X_i X_j\right) =$$

$$= \frac{1}{n} E\left(X^2\right) + \frac{n-1}{n}\left(E\left(X\right)\right)^2$$

$$E\left(S_n^2\right) = E\left(X^2\right) - \frac{1}{n} E\left(X^2\right) - \frac{n-1}{n}\left(E\left(X\right)\right)^2 =$$

$$= \frac{n-1}{n}\left(E\left(X^2\right) - \left(E\left(X\right)\right)^2\right) = \frac{n-1}{n}\sigma^2$$

Standard for of a random variable

The standard form of a random variable *X*, denoted *Z*, is defined by

$Z = \dfrac{X - \mu_X}{\sigma_X}$. Let us compute the mean value and variance of*Z*.

The mathematical expectation (mean value) of an arbitrary function, *g*(*X*), of the random variable *X* is defined as

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

Using this definition with Z = g(X) we derive

$$\mu_Z = E\left[\frac{X-\mu_X}{\sigma_X}\right] = \frac{1}{\sigma_X}\left[E\left(X\right) - E\left(\mu_X\right)\right] = \frac{1}{\sigma_X}\left(\mu_X - \mu_X\right) = 0$$

and

$$\sigma_Z^2 = E\left(Z^2\right) - \mu_Z^2 = E\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2\right] - 0 = \frac{1}{\sigma_X^2}\left[E\left(X-\mu_X\right)^2\right] = \frac{\sigma_X^2}{\sigma_X^2} = 1$$

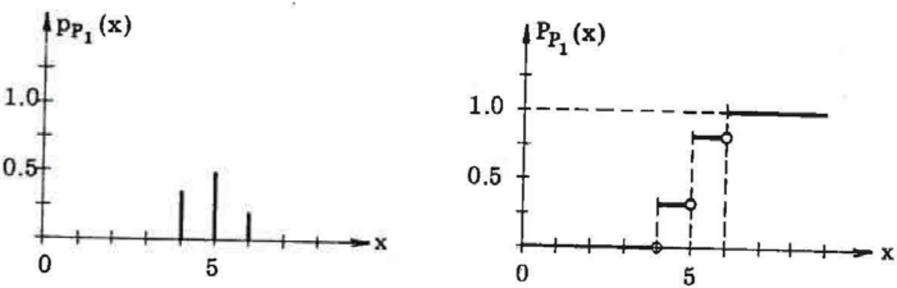
Thus the mean of the standard form of a random variable is 0, its variance is 1.

Example 2.9.

Consider the discrete random variable *X* defined in example 2.7. The discrete version (0.14) gives

$$E[X] = 4 \cdot 0.3 + 5 \cdot 0.5 + 6 \cdot 0.2 = 4.9$$

The most probable value is called the *mode*, in this case equal to 5.0 (see Figure).



Further

$$E\left[X^2\right] = 16 \cdot 0.3 + 25 \cdot 0.5 + 36 \cdot 0.2 = 24.5$$

A new random variable $Y = X^k$ was considered above.
It is a function of a random variable whose distribution is known.
Let $Y = f(X)$, f is a function with only finite discontinuities.

We derive that Y is a random variable too.
If the f is monotonic the distribution function F_Y is given by

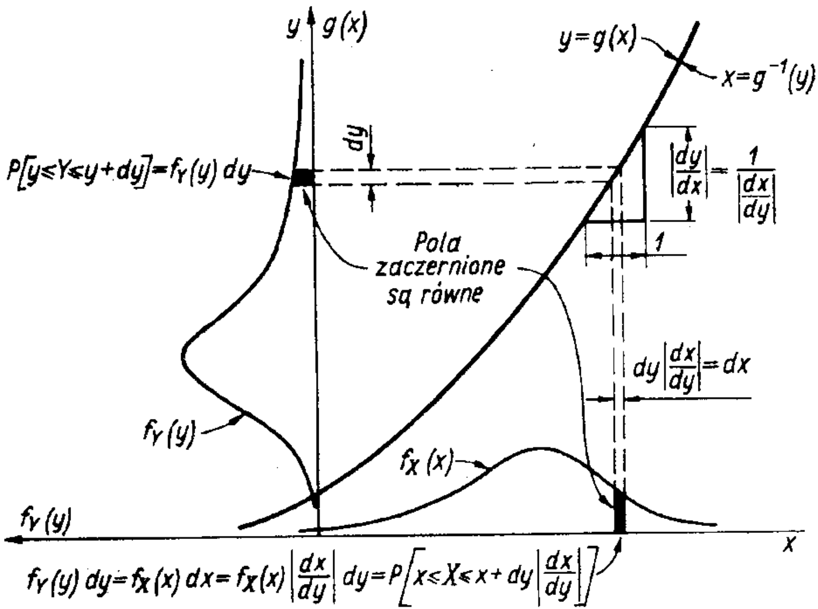
$$F_Y(y) = P(Y \le y) = P\left(X \le f^{-1}(y)\right) = F_X\left(f^{-1}(y)\right)$$

and its density function f_Y by

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(f^{-1}(y)\right) = f_X\left(f^{-1}(y)\right)\left|\frac{df^{-1}(y)}{dy}\right|$$

or simply

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$



Rys. 2.3.4. Interpretacja graficzna równości (2.3.5) i (2.3.12): $f_Y(y) = |dx/dy| f_X(x)$

Example 2.10.

Let $Y = aX + b$. It yields, $X = \dfrac{Y-b}{a}$, so $f_Y(y) = f_X\left(\dfrac{y-b}{a}\right)\left|\dfrac{1}{a}\right|$

Expected value of $Y = f(X)$ is computed without the f_Y function.

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy = \int_{-\infty}^{\infty} f(x)\, f_X(x)\,dx = E\left[f(X)\right]$$

It holds

$$E\left[\sum_{i=1}^{n} f_i(X)\right] = \sum_{i=1}^{n} E\left[f_i(X)\right]$$

for the expectation and summation operators commute.

The n^{th} *central moment* of X is $E\left[(X-\mu_X)^n\right]$, where $\mu_X = E[X]$.

Example 2.11.

Consider the same discrete random variable X as in example 2.9, where $E\left[X\right]=4.9$ and $E\left[X^2\right]=24.5$.

The variance is

$$\sigma_X^2 = 24.5 - 4.9^2$$

and the standard deviation

$$\sigma_X = \sqrt{0.49} = 0.7$$

Thus the coefficient of variation

$$V_X = \frac{\sigma_X}{\mu_X} = \frac{0.7}{4.9} = 0.14$$

The third central moment is a measure of the *asymmetry (asymetria)* or *skewness (skośność)* of the distribution of a random variable. For a continuous random variable it is defined by

$$E\left[\left(X-\mu_X\right)^3\right] = \int_{-\infty}^{\infty} \left(x-\mu_X\right)^3 f_X\left(x\right) dx$$