POLITECNICO DI MILANO

Basic notions of probability theory

Contents

Why a Lecture on Probability?

Lecture 1, Slide 22:

Risk

RISK = POTENTIAL DAMAGE + UNCERTAINTY

Dictionary: RISK = possibility of damage or injury to people or things

1) What undesired conditions may occur? → Accident Scenario, S
2) With what probability do they occur? → Probability, p
3) What damage do they cause? → Consequence, x

RISK = $\{S_i, p_i, x_i\}$

CentraleSupélec 22

Basic Definitions

Definitions: experiment, sample space, event

- Experiment ε: process whose outcome is a priori unknown to the analyst (all possible outcomes are a priori known)
- Sample space Ω: the set of all possible outcomes of ε.
- Event E: a set of possible outcomes of the experiment ε (a subset of Ω):

  the event E occurs when the outcome of the experiment ε is one of the elements of E.

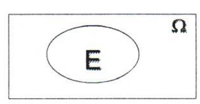
Boolean Logic

Definition: Certain events → Boolean Logic

Logic of certainty: **an event** E can either **occur** or **not occur**

Indicator variable $X_E = \begin{cases} 0, & \text{when } E \text{ does not occur} \\ 1, & \text{when } E \text{ occurs} \end{cases}$

Certain Events (Example)



(ε = die toss; Ω={1,2,3,4,5,6}; E=Odd number)

I perform the experiment and the outcome is '3'

Event E, X_E ▶ True $X_E=1$

Boolean Logic Operations

- Negation: $\overline{E}$
- Union:
  $X_{A\cup B}$
- Intersection:
  $X_{A\cap B}$

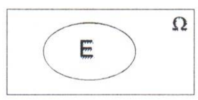
Boolean Logic Operations

- Negation: $\overline{E} \rightarrow \overline{X_E} = 1 - X_E$
- Union:
  $$X_{A\cup B} = 1-(1-X_A)(1-X_B) = 1-\prod_{j=A,B}(1-X_j) =$$
  $$= \coprod_{j=A,B} X_j = X_A + X_B - X_A X_B$$
- Intersection:
  $$X_{A\cap B} = X_A X_B$$
- Definition: A and B are mutually exclusive events if $X_{A\cap B} = 0$

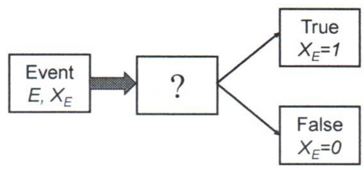
Uncertain Events

Uncertain Events

Let us consider: the experiment ε, its sample space Ω and the event *E*.

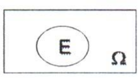


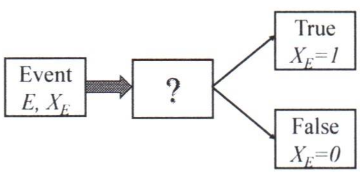
(ε = die toss; Ω={1,2,3,4,5,6}; E=Odd number)



Uncertain Events:

Let us consider: the experiment ε, its sample space Ω, event *E*.





Uncertain events can be compared → probability of E = p(E)

Probability for comparing the likelihood of events

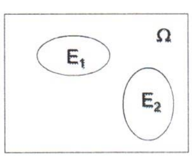
Money for comparing the value of objects

Probability theory

Probability theory: Kolmogorov Axioms

1. $0 \le p(E) \le 1$
2. $p(\Omega) = 1 \quad p(\varnothing) = 0$
3. Addition law:
   Let $E_1, \ldots, E_n$ be a finite set of mutually exclusive events: $(X_{E_i} \cap X_{E_j} = \varnothing)$.

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



Definitions of probability

Three definitions of probability

1. **Classical definition**
2. **Empirical Frequentist Definition**
3. **Subjective definition**

1. Classical Definition of Probability

- Let us consider an experiment with N possible elementary, mutually exclusive and equally probable outcomes: $A_1, A_2, \ldots, A_N$ and the event:

$$E = A_1 \,\mathrm{Y}\, A_2 \,\mathrm{Y} \ldots \mathrm{Y}\, A_M$$

$$p(E) = \frac{number\ of\ outcomes\ resulting\ in\ E}{total\ number\ of\ possible\ outcomes} = \frac{M}{N}$$

1. Classical Definition of Probability

- Let us consider an experiment with N possible elementary, mutually exclusive and equally probable outcomes: $A_1, A_2, \ldots, A_N$:

When is it applicable?
- Gambling (e.g. tossing of a die)
- If no evidence favouring one outcome over others

$$E = A_1 \,\mathrm{Y}\, A_2 \,\mathrm{Y} \ldots \mathrm{Y}\, A_M$$

$$p(E) = \frac{number\ of\ outcomes\ resulting\ in\ E}{total\ number\ of\ possible\ outcomes} = \frac{M}{N}$$

1. Classical Definition of Probability (criticisms)

- Let us consider an experiment with N possible elementary, mutually exclusive and equally probable outcomes: $A_1, A_2, \ldots, A_N$:

When is this requirement met?

In most real life situations the outcomes are not equally probable!

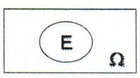
$$E = A_1 \,\mathrm{Y}\, A_2 \,\mathrm{Y} \ldots \mathrm{Y}\, A_M$$

$$p(E) = \frac{number\ of\ outcomes\ resulting\ in\ E}{total\ number\ of\ possible\ outcomes} = \frac{M}{N}$$

2. Frequentist Definition of Probability

Let us consider: the experiment ε, its sample space Ω and an event *E*.

ε = die toss; Ω={1,2,3,4,5,6}; E={Odd number}



- *n* times ε, *E* occurs *k* times

(*n* = 100 die tosses → *k*=48 odd numbers)

- *k/n* = the relative frequency of occurrence of *E*

(k/*n* = 48/100=0.48)

$$\lim_{n\to\infty} \frac{k}{n} = p$$

p is defined as the probability of *E*

2. Frequentist Definition of Probability (criticisms)

$$\lim_{n\to\infty} \frac{k}{n} = p$$

- This is not a limit from the matemathical point of view [limit of a numerical series]
- It is not possible to repeat an experiment an infinite number of times …
- We tacitly assume that the limit exists

Possible way out?

probability as a physical characteristic of the object:
- the physical characteristics of a coin (weight, center of mass, …) are such that when tossing a coin over and over again the fraction of 'head' will be *p*

2. Frequentist Definition of Probability (criticisms)

$$\lim_{n\to\infty} \frac{k}{n} = p$$

- Applicable only to those events for which we can conceive of a repeatable experiment (e.g. not to the event «your professor will be sick tomorrow»)
- The experiment conditions cannot be identical
  - let us consider the probability that a specific valve *V* of a specific Oil & Gas plant will fail during the next year
  - *what should be the population of similar valves?*
    - Large population: all the valves used in industrial plants. Considering data from past years, we will have a large number *n*, but data may include valves very different to *V*
    - Small population: valve used in Oil&Gas of the same type, made by the same manufacturer with the same technical characteristics → too small *n* for limit computation.

Similarity Vs population size dilemma

2. Frequentist Definition of Probability (criticisms)

$$\lim_{n\to\infty} \frac{k}{n} = p$$

Some events (e.g. in the nuclear industry) have very low probabilities (e.g. $p \approx 10^{-6}$) (RARE EVENTS)

Very difficult to observe

The frequentist definition is not applicable

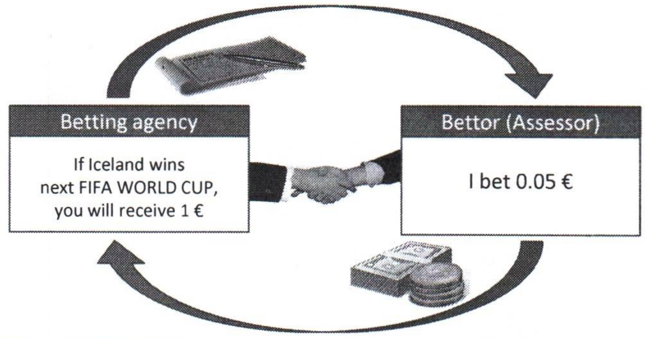
3. Subjective Definition of Probability

P(E) is the **degree of belief** that a person (assessor) has that E will occour, **given all the relevant information currently known to that person** (background knowledge)

- Probability is a numerical encoding of the state of knowledge of the assessor (De Finetti: "probability is the feeling of the analyst towards the occurrence of the event")
- P(E) is conditional on the background knowledge K of the assessor: **P(*E*)= P(*E*|*K*)**
- Background knowledge typically includes data/models/expert knowledge
- If the background knowledge changes → the probability may change
- Two interpretations of subjective probability:
  - Betting interpretation
  - Reference to a standard for uncertainty

3. Betting Interpretation

P {Iceland will win next UEFA EURO 2020|K} = 0.05

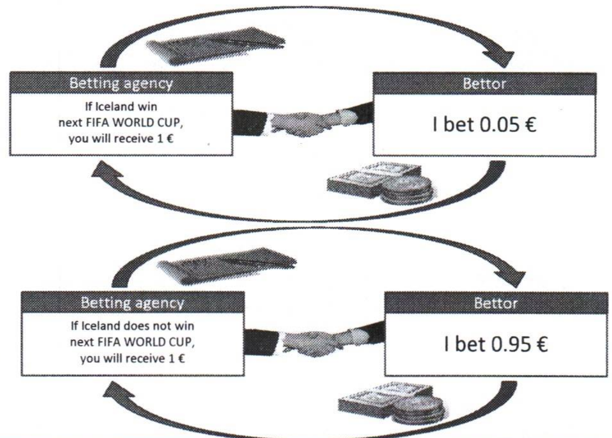


3. Betting Interpretation

P(E) is the amount of money that the person assigning the probability would be willing to bet if a single unit of payment were given in return in case event E were to occur and nothing otherwise.

The opposite must also hold: 1- P(E) is also the amount of money that the person assigning the probability would be willing to bet if a single unit of payment were given in return in case event E were not to occur and nothing otherwise.

3. Betting Interpretation: two sideness of the bet



3. Betting Interpretation (Criticism)

P(E) is the amount of money that the person assigning the probability would be willing to bet if a single unit of payment were given in return in case event E were to occur and nothing otherwise.
The opposite must also hold (1- P(E)) is also the amount of money that the person assigning the probability would be willing to bet if a single unit of payment were given in return in case event E were not to occur and nothing otherwise.

probability assignment depends from the value judgment about money and event consequences (the assessor may even think that in case of LOCA in a Nuclear Power Plant he/she will die and so the payment will be useless)

3. Reference to a standard for uncertainty

P(E) is the number such that the uncertainty about the occurrence of E is considered equivalent by the person assigning the probability (assessor) to the uncertainty about drawing a red ball from an urn containing P(E)*100% red balls

E={Germany will win next FIFA WORLD CUP}

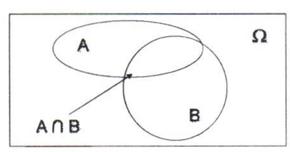
$P(E)$=0.33

urn

Probability laws

Probability laws (1)

- Union of two non-mutually exclusive events



$$P_{A\cup B} = P_A + P_B - P_{A\cap B}$$

It can be demonstrated by using the three Kolmogorov axioms*

$$P_{A\cup B} \leq P_A + P_B$$

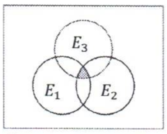
- Rare event approximation: A and B events are considered as mutually exclusive ($A \cap B = \emptyset$) → $P(A \cap B) = 0$ →

$$P_{A\cup B} = P_A + P_B$$

* http://www.ucs.louisiana.edu/~jcb0773/Berry_probbook/425chpt2.pdf

Probability laws (2)

- Union of non-mutually exclusive events: $E_{\cup} = \bigcup_{i=1,\dots,n} E_i$



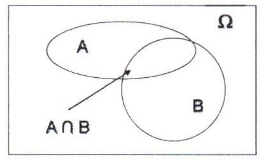
$$P(E_{\cup}) = \sum_{i=1}^{n} P(E_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(E_i \cap E_j) + \dots + (-1)^{n+1} P(E_1 \cap E_2 \cap \dots \cap E_n)$$

- Upper bound $P(E_U) \le \sum_{j=1}^{n} P(E_j)$
- Lower bound $P(E_U) \ge \sum_{j=1}^{n} P(E_j) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(E_i \cap E_j)$
- Rare event approximation: events are considered as mutually exclusive $(E_i \cap E_j = \emptyset, \forall i,j, i \neq j) \rightarrow P(E_{\cup}) = \sum_{i=1}^{n} P(E_i)$

Probability laws (3)

- Conditional Probability of A given B

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$



- Event A is said to be statistically independent from event B if:

$$P(A \mid B) = P(A)$$

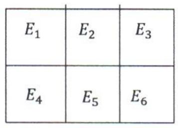
- If A and B are statistically independent then:

$$P(A \cap B) = P(A)P(B)$$

Theorem of Total Probability

- Let us consider a partition of the sample space Ω into n mutually exclusive and exhaustive events. In terms of Boolean events:

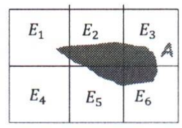
$$E_i \cap E_j = 0 \quad \forall i \neq j \qquad \bigcup_{j=1}^{n} E_j = \Omega$$



Theorem of Total Probability

- Let us consider a partition of the sample space Ω into n mutually exclusive and exhaustive events. In terms of Boolean events:

$$E_i \cap E_j = 0 \quad \forall i \neq j \qquad \bigcup_{j=1}^{n} E_j = \Omega$$



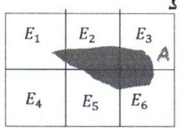
- Given any event A in Ω, its probability can be computed in terms of the partitioning events and the conditional probabilities of A on these events: $A = \cup_j (A \cap E_j) \rightarrow P(A) = \sum_j P(A \cap E_j)$

$$P(A) = P(A \mid E_1)P(E_1) + P(A \mid E_2)P(E_2) + \dots + P(A \mid E_n)P(E_n)$$

Bayes Theorem

- Let us consider a partition of the sample space Ω into n mutually exclusive and exhaustive events E_j. We know
- Event A has occurred

Can I use this information to update the probability of $P(E_j)$?



$$P(E_i \mid A) = \frac{P(E_i A)}{P(A)} = \frac{P(A \mid E_i)P(E_i)}{\sum_{j=1}^{n} P(A \mid E_j)P(E_j)}$$

theorem of total probability

The Bayesian Subjective Probability Framework

P(E|K) is the **degree of belief** of the **assigner** with regard to the occurrence of E (numerical encoding of the **state of knowledge** – K - of the assessor)

Bayes Theorem to update the probability assignment in light of new information

Updated

Old

$$P(E_i|A,K) = \frac{P(A|E_i,K) \cdot P(E_i|K)}{\sum_{j=1}^{n} P(A|E_j,K) \cdot P(E_j|K)}$$

new information

POLITECNICO DI MILANO

Basic notions of probability theory

(Part 2)

Contents

Random variables

Random variables

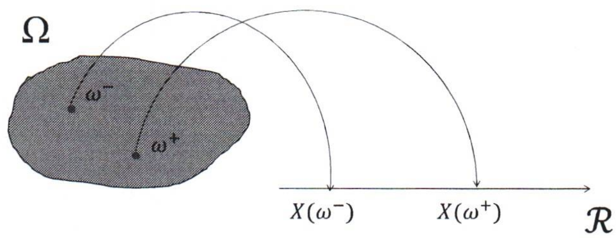
Experiment: ε

Sample space: Ω

Generic outcome: ω

$\Rightarrow$ $X(\omega)$ random variable in $\mathcal{R}$

Univocal mapping



Random variable - Example

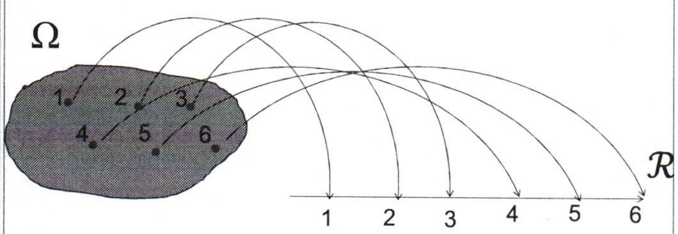
Experiment: ε = {tossing a dice}

Sample space: Ω={1,2,3,4,5,6}

Outcome: ω

$\Rightarrow$ $X(\omega)$ in $\Re$

Univocal mapping



Random variable - Event

Experiment: ε = {tossing a dice}

Sample space: Ω={1,2,3,4,5,6}

Event: E_1= {1,2,3,4}

$E_2 = \emptyset$

$E_3 = \Omega$

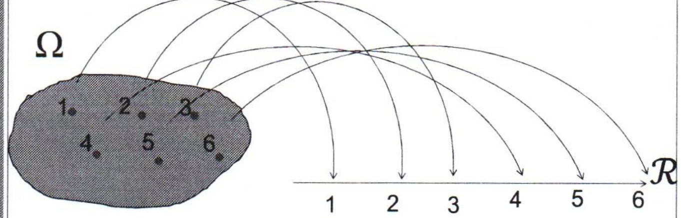
$\Rightarrow$ X(ω)

E_1={X<4.236}

E_2={X<0}

E_3={X<+∞}

Univocal mapping



Random variables

Experiment: ε

Sample space: Ω

Generic outcome: ω

$\Rightarrow$ $X(\omega)$ random variable in $\mathcal{R}$

▼

General mathematical models of random behaviours

▼

They apply to different physical phenomena which behave similarly

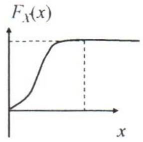
Probability distributions for reliability, safety and risk analysis

Probability functions (I)

- **Cumulative Distribution Function (cdf)**
  - $F_X(x)$ gives the probability of the event $\{X \le x\}$ for any numerical value x.
  - Properties:

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

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

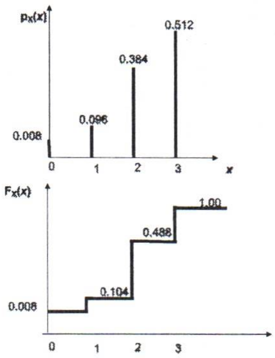


  - $F_X(x)$ is a non-decreasing function of x
  - The probability that X takes on a value in the interval $[a,b]$ is:

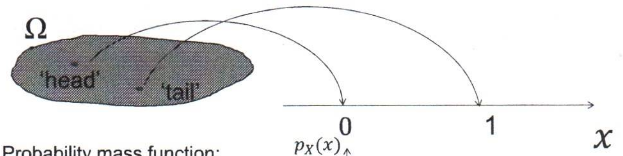
$$P\{a < X \le b\} = F_X(b) - F_X(a)$$

Probability functions (II, discrete random variables)

- **Probability Mass Function (pmf)**
  - X – random variable takes discrete values x_i, $i = 1,2,\dots n$:
    - $p_i = P\{X = x_i\}$
    - $F_X(x) = \sum_{x_i \le X} p_i$

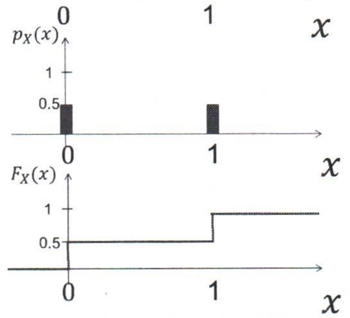


Probabilty Mass Function - Example



Probability mass function:
$p_0 = P\{X = 0\} = 0.5$
$p_1 = P\{X = 1\} = 0.5$

Cumulative distribution



Probability functions (II, continuous random variables)

- X – random variable takes continuous values in R
- Definition:

probability density function

$$P\{x \le X < x + dx\} = F_X(x+dx) - F_X(x) = f_X(x)dx$$

- $dx \to 0$:

$$f_X(x) = \lim_{dx\to0} \frac{F_X(x+dx) - F_X(x)}{dx} = \frac{dF_X}{dx}$$

- $f_X(x)$ is not a probability but a probability per unit of x (probability density)

Summary measures: *percentiles, median, mean, variance*

- Distribution Percentiles (x_α):

$$F_X(x_\alpha) = \frac{\alpha}{100}$$

- Median of the distribution (x_{50}):

$$F_X(x_{50}) = 0.5$$

- Mean Value (Expected Value):

$$\mu_X = E[X] = \langle X \rangle = \sum_{i=1}^{n} x_i p_i \quad \text{(discrete random variables)}$$

$$= \int_{-\infty}^{\infty} x f_X(x) dx \quad \text{(continuous random variables)}$$

- Variance (var[X]):

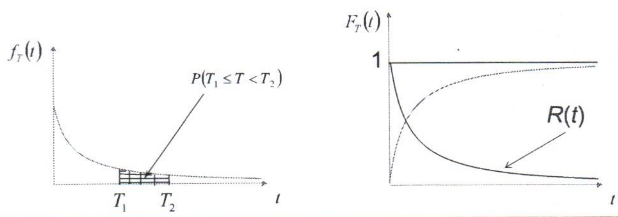
$$\sigma_X^2 = \sum_i (x_i - \mu_X)^2 p_i \quad \text{(discrete random variables)}$$

$$= \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \quad \text{(continuous random variables)}$$

Reliability

- T = Time to failure of a component (random variable)
  - Probability density function (pdf) at time t: $f_T(t)$
  - Cumulative distribution function (cdf) at time t = probability of having a failure before t : $F_T(t) = P(T<t)$
  - Reliability at time t = Probability that the component does not fail up to t:

$R(t) = 1 - F_T(t)$



The Hazard Function – *failure rate* (I)

$$\boxed{h_T(t)dt = P(t < T \le t+dt \mid T > t) = \frac{P(t < T \le t+dt)}{P(T>t)} = \frac{f_T(t)dt}{R(t)}}$$

Hazard Function and Reliability

$$\boxed{h_T(t)dt = P(t < T \le t+dt \mid T > t) = \frac{P(t < T \le t+dt)}{P(T>t)} = \frac{f_T(t)dt}{R(t)}}$$

$$h_T(t)dt = -\frac{dR(t)}{R(t)}$$

$$\int_0^t h_T(t')dt' = -\ln R(t)$$

$$R(t) = e^{-\int_0^t h_T(t')dt'}$$

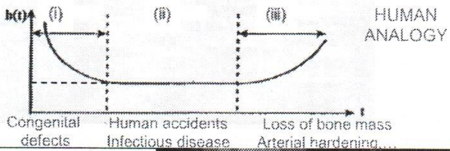
$$f(t) = h(t)R(t) = h(t)e^{-\int_0^t h_T(t')dt'}$$

Hazard Function: the Bath-Tub Curve

- Usually, the hazard function shows three distinct phases:
  i. Decreasing - *infant mortality* or *burn in period:*
     - Failures due to defective pieces of equipment not manufactured or constructed properly (missing parts, substandard material batches, damage in shipping, ...)

  ii. Constant - *useful life*
     - Random failures due to unavoidable loads coming from without (earthquakes, power surges, vibration, temperature fluctuations,...)

  iii. Increasing – *ageing*
     - Aging failures due to cumulative effects such as corrosion, embrittlement, fatigue, cracking, ...



Univariate probability distributions

Univariate Discrete Distributions: Binomial Distribution (I)

Y = discrete random variable with only two possible outcomes:

- Y=1 (success) with $P\{Y=1\}=p$
- Y=1 (failure) with $P\{Y=0\}=1-p$

Bernoulli process

We perform n different trials of the experiment, $Y_1, \dots, Y_n$

X= discrete random variable describing the number of success out of the n trials, independently from the sequence with which success appear:

$X = \sum_{i=1}^{n} Y_i \qquad \Omega = \{1,2,\dots,n\}$

The probability mass function:

$$\boxed{b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}} \qquad k=1,2,\dots n$$

Univariate Discrete Distributions: Binomial Distribution (II)

$$\boxed{b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}} \qquad k=1,2,\dots n$$

$$\boxed{\begin{array}{l} E[X] = np \\ Var[X] = np(1-p) \end{array}}$$

Univariate Discrete Distributions, Geometric Distribution

T= trial of the first success (or number of trials between two successive occurrences of success)

The probability mass function:

$$\boxed{g(t;p) = (1-p)^{t-1} p} \qquad t=1,2,\dots$$

Expected value (return period):

$$\boxed{E[T] = \sum_{t=1}^{\infty} t(1-p)^{t-1} p = p[1 + 2(1-p) + 3(1-p)^2 + \dots] = \frac{p}{[1-(1-p)]^2} = \frac{1}{p}}$$

Univariate Discrete Distributions, Poisson Distribution

Stochastic events that occur in a continuum period:

- Independent events
- Rate of occurrence, λ, is constant

- Discrete Random Variable, K = number of events in the period of observation $(0,t)$

- Probability mass function: $\boxed{p(k;(0,t),\lambda) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}}$ $\quad k=1,2,\dots$

$$\boxed{\begin{array}{l} E[K] = \lambda t \\ Var[K] = \lambda t \end{array}}$$

Continuous Distributions: Exponential Distribution

- $h_T(t)=\lambda$ constant
- T=failure time

0 —— t ——→

$P\{T>t\}$=P{no failure in $(0,t)$}=

=Poisson(k=0,(0,t),λ) $= \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}$

$$F_T(t) = 1 - P\{T > t\} = 1 - e^{-\lambda t}$$
$$f_T(t) = \lambda e^{-\lambda t}$$

- It is the only distribution characterized by a constant failure rate

Exponential Distribution and bath tub curve

Poisson vs. Exponential

1.

Suppose to wait at a bus stop. A bus usually arrives at the stop in every 10 mins. The rate of arrival of a bus per minute is λ = 1/10. What is the probability that no bus arrives in the next minute?

Poisson	λ = 1/10	Probability of 0 arrivals in the next minute: P(X = 0) = 0.9048
Exponential	λ = 1/10	Probability of waiting more than 1 minute: P(X>1) = 0.9048

Note: Look at the expected values of both the distributions.
– Poisson: average number of buses arriving per minute E(X) = λ = 0.10 buses
– Exponential: average waiting time for a bus to arrive E(X) = (1/λ) = 10 min

2.

N(t) ~ P(λ) = number of arrivals during a tme period t
X(t) ~ E(λ) = the time it takes for one additional arrival to arrive assuming that someone arrived at time t

P(X(t) > x) = P(N(t) = N(t+x))

The probability that no one has arrived in the time interval [t, t+x] is equal to the probability that the count of the number of arrivals at time t+x is identical to the count at time t.

P(N(t+x) – N(t) = 0) = e^(–λ*x)
P(X(t) > x) = 1 – P(X(t) ≤ x) = e^(–λ*x)

Exponential Distribution moments

$$E[T] = \int_0^{+\infty} t f(t) dt = \int_0^{+\infty} t\lambda e^{-\lambda t} dt = \frac{1}{\lambda}$$

$$\mathrm{Var}[T] = \frac{1}{\lambda^2}$$

Exponential distribution: memorylessness

- A component with constant failure rate, λ, is found still operational at a given time t_1.

$$P\{T \le t_2 | T > t_1\} = \frac{P(t_1 < T \le t_2)}{P(T > t_1)} = \frac{F(t_2) - F(t_1)}{R(t_1)} =$$

$$\frac{(1-e^{-\lambda t_2}) - (1-e^{-\lambda t_1})}{e^{-\lambda t_1}} = \frac{(e^{-\lambda t_1} - e^{-\lambda t_2})}{e^{-\lambda t_1}} = 1 - e^{-\lambda(t_2 - t_1)}$$

- Still exponential with failure rate λ!
- The probability that it will fail in some period of time of lengths $\tau = t_2 - t_1$ does not depend from the component age t_1

Continuous Distributions : the Weibull Distribution

- In practice, the age of a component influences its failure process so that the hazard rate does not remain constant throughout the lifetime

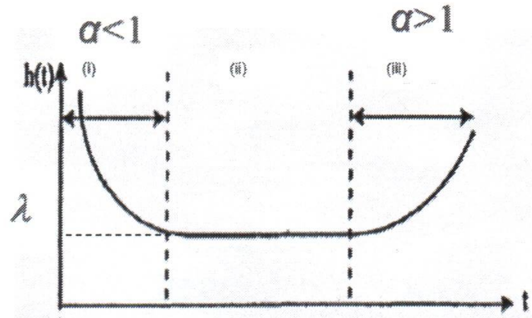
$$F_T(t) = P(T \le t) = 1 - e^{-\lambda t^{\alpha}} \qquad \begin{cases} f_T(t) = \lambda \alpha t^{\alpha-1} e^{-\lambda t^{\alpha}} & t \ge 0 \\ \quad = 0 & t < 0 \end{cases}$$

$$E[T] = \frac{1}{\lambda}\Gamma\left(\frac{1}{\alpha}+1\right) \quad ; \quad Var[T] = \frac{1}{\lambda^2}\left(\Gamma\left(\frac{2}{\alpha}+1\right) - \Gamma\left(\frac{1}{\alpha}+1\right)\right)^2$$

$$\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x} dx \qquad k > 0$$

Weibull Distribution and bath tub curve

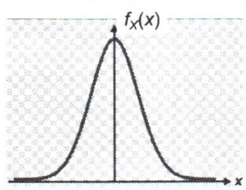
$$h(t) = \frac{f(t)}{1 - F(t)} = \lambda \alpha t^{\alpha - 1}$$



Continuous Distributions: Normal or Gaussian Distribution

Probability density function: $\boxed{X \sim N(\mu_X, \sigma_X)}$

$$\boxed{f_X(x;\mu_X,\sigma_X) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} \qquad -\infty < x, \mu_X < \infty; \sigma_X > 0}$$



It is the only distribution with a symmetric bell shape!

Expected value and variance: $\boxed{\begin{array}{l} E[X] = \mu_X \\ Var[X] = \sigma_X^2 \end{array}}$

Central limit theorem

- For any sequence of n independent random variable X_i, their sum $X = \sum_{i=1}^{n} X_i$ is a random variable which, for large n, tends to be distributed as a normal distribution

If X_i are independent, identically distributed random variables with mean μ and finite variance given by σ^2

$$S_n = \frac{\sum_{i=0}^{n} X_i}{n} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$$

Standard Normal Variable

$$\boxed{P(a < X < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx}$$

$$\boxed{S = \frac{X - \mu}{\sigma}} \qquad S \sim N(0,1)$$

$$\boxed{P(a < X < b) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}s^2} \sigma \, ds =}$$

$$\boxed{\Rightarrow P(a < X < b) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}s^2} ds = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

Table of Standard Normal Probability

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\xi^2} d\xi$$

x	Φ(x)	x	Φ(x)	x	Φ(x)
0.00	0.500000	0.50	0.691462		
0.01	0.503989	0.51	0.694974		
0.02	0.507978	0.52	0.698468		
0.03	0.511966	0.53	0.701944		
0.04	0.515954	0.54	0.705401	1.04	0.850830
0.05	0.519939	0.55	0.708840	1.05	0.853141
0.06	0.523922	0.56	0.712260	1.06	0.855428
0.07	0.527904	0.57	0.715661	1.07	0.857690
0.08	0.531862	0.58	0.719043	1.08	0.859929
0.09	0.535857	0.59	0.722405	1.09	0.862143
0.10	0.539828	0.60	0.725747	1.10	0.864334
0.11	0.543796	0.61	0.729069	1.11	0.866500
0.12	0.547759	0.62	0.732371	1.12	0.868643
0.13	0.551717	0.63	0.735653	1.13	0.870762
0.14	0.555671	0.64	0.738914	1.14	0.872857
0.15	0.559618	0.65	0.742154	1.15	0.874928
0.16	0.563560	0.66	0.745374	1.16	0.876976
0.17	0.567494	0.67	0.748572	1.17	0.878999
0.18	0.571423	0.68	0.751748	1.18	0.881000
0.19	0.575345	0.69	0.754903	1.19	0.882977
0.20	0.579260	0.70	0.758036	1.20	0.884930
0.21	0.583166	0.71	0.761148	1.21	0.886860
0.22	0.587064	0.72	0.764238	1.22	0.888767
0.23	0.590954	0.73	0.767305	1.23	0.890651
0.24	0.594835	0.74	0.770350	1.24	0.892512
0.25	0.598706	0.75	0.773373	1.25	0.894350
0.26	0.602568	0.76	0.776373	1.26	0.896165
0.27	0.606420	0.77	0.779350	1.27	0.897958
0.28	0.610262	0.78	0.782305	1.28	0.899727
0.29	0.614092	0.79	0.785236	1.29	0.901475
0.30	0.617912	0.80	0.788145	1.30	0.903199
0.31	0.621720	0.81	0.791030	1.31	0.904902
0.32	0.625517	0.82	0.793892	1.32	0.906582
0.33	0.629301	0.83	0.796731	1.33	0.908241
0.34	0.633072	0.84	0.799546	1.34	0.909877
0.35	0.636831	0.85	0.802337	1.35	0.911492
0.36	0.640576	0.86	0.805105	1.36	0.913085
0.37	0.644309	0.87	0.807850	1.37	0.914656
0.38	0.648027	0.88	0.810570	1.38	0.916207
0.39	0.651732	0.89	0.813267	1.39	0.917736
0.40	0.655422	0.90	0.815940	1.40	0.919243
0.41	0.659097	0.91	0.818589	1.41	0.920730
0.42	0.662757	0.92	0.821214	1.42	0.922196
0.43	0.666402	0.93	0.823815	1.43	0.923641
0.44	0.670032	0.94	0.826391	1.44	0.925066
0.45	0.673645	0.95	0.828944	1.45	0.926471
0.46	0.677242	0.96	0.831472	1.46	0.927855
0.47	0.680823	0.97	0.833977	1.47	0.929219
0.48	0.684387	0.98	0.836457	1.48	0.930563
0.49	0.687933	0.99	0.838913	1.49	0.931888

Table of Standard Normal Probability

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Table of Standard Normal Probability

x	$\Phi(x)$	x	$\Phi(x)$	x	$1-\Phi(x)$
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Probability distributions:

Univariate Continuous Distributions Log-normal Distribution

Probability density function:

$$\boxed{g_X(x;\mu_Z,\sigma_Z)=\frac{1}{\sqrt{2\pi}\sigma_Z}\frac{1}{x}e^{-\frac{1}{2}\left(\frac{\ln x-\mu_Z}{\sigma_Z}\right)^2} \qquad x,\sigma_Z>0}$$

Note: if

$$X\sim Log-normal(\mu_Z,\sigma_Z)\Rightarrow Z=\ln X\sim N(\mu_Z,\sigma_Z)$$

Expected value and variance:

$$\boxed{\begin{aligned}&E[X]=e^{\mu_Z+\frac{\sigma_Z^2}{2}}\\&Var[X]=e^{2\mu_Z+\sigma_Z^2}(e^{\sigma_Z^2}-1)\end{aligned}}$$

Bivariate probability distributions

Probability distributions:

Joint Probability Distributions, Bivariate Distribution

Joint cumulative distribution function:

X, Y - random variables

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

constraints:

a) $F_{XY}(-\infty,-\infty)=0;$ $\qquad F_{XY}(\infty,\infty)=1$

b) $F_{XY}(-\infty,y)=0;$ $\qquad \left.\begin{aligned}F_{XY}(\infty,y)=F_Y(y)\\F_{XY}(x,\infty)=F_X(x)\end{aligned}\right\}$ marginal CDFs

$F_{XY}(x,-\infty)=0;$

c) $F_{XY}(x,y)$ is nonegative, and nondecreasing function of x and y

Joint probability density function:

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

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

Probability distributions:

Joint Probability Distributions, Bivariate Distribution

Conditional probability density functions:

$$\boxed{P[X\le x|Y\le y]=F_{X|Y}(x|y)=\frac{F_{XY}(x,y)}{F_Y(y)},\text{ if } F_Y(y)\ne 0}$$

$$\boxed{P[Y\le y|X\le x]=F_{Y|X}(y|x)=\frac{F_{XY}(x,y)}{F_X(x)},\text{ if } F_X(x)\ne 0}$$

The marginal PDFs can be obtained from the joint PDF by applying the theorem of total probability:

$$\boxed{f_X(x)=\int_{-\infty}^{\infty}f_{XY}(x,y)dy} \qquad \boxed{f_Y(y)=\int_{-\infty}^{\infty}f_{XY}(x,y)dx}$$

X, Y - statistically independent

Joint probability density function:

$$\boxed{f_{XY}(x,y)=f_X(x)\cdot f_Y(y)}$$

Probability distributions:

Joint Probability Distributions, Bivariate Distribution

Covariance:

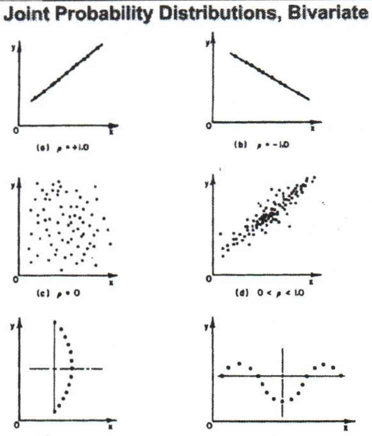
$$\boxed{\mathrm{cov}[X,Y]=E[(X-\mu_X)\cdot(Y-\mu_Y)]=E[XY]-E[X]\cdot E[Y]}$$

Correlation coefficient (normalized covariance):

$$\boxed{\rho_{XY}=\frac{\mathrm{cov}[X,Y]}{\sigma_X\cdot\sigma_Y}, \qquad -1\le\rho_{XY}\le 1}$$

Probability distributions:

Joint Probability Distributions, Bivariate Distribution



- ρ just gives a 'descriptive' relationship: $\rho\ne 0$ does not mean that there is a casual relationship; the variables remain random
- ρ is a measure of linear relationship; you might have random variables that are completely dependent but have $\rho=0$
- $\rho=0$ $\not\Rightarrow$ X and Y are independent

if X and Y are independent $\Rightarrow \rho=0$

Functions of random variables

Functions of Random Variables: Derived Probability Distributions

$X, Y = g(X)$

Cumulative distribution function:

$$\boxed{\begin{aligned} F_Y(y) &= P[Y \le y] = P[g(X) \le y] = \\ &= P[X \le g^{-1}(y)] = F_X[g^{-1}(y)] \end{aligned}}$$

assume that $g(x)$ is a monotonically increasing function

Probability density function:

$$f_Y(y)dy = f_X(x)dx \qquad , Y = g(X)$$

$$x_i = g^{-1}(y) \qquad x_i \text{ is the } i\text{-th root of } g^{-1}(y)$$

$$g'(x) = \frac{dy}{dx}$$

- derived density function:

$$\Rightarrow f_Y(y) = \sum_{i=1}^{k} \frac{f_X(x_i)}{|g'(x_i)|}$$

Functions of Random Variables:

Mean and Variance of a General Function

$X, Y = g(X)$

Mean:

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

Variance:

$$Var[Y] = \int_{-\infty}^{\infty} \left[g(x) - E[Y] \right]^2 f_X(x) dx$$

Expanding $g(X)$ in a Taylor series about the mean value $\bar{X} = \mu_x$

$$Y = g(\bar{X}) + (X - \bar{X}) \frac{dg}{dX}\bigg|_{\bar{X}} + \frac{1}{2}(X - \bar{X})^2 \frac{d^2 g}{dX^2}\bigg|_{\bar{X}} + \ldots$$

Functions of Random Variables:

Approximate Mean and Variance of a General Function

First-order approximate mean:

$$E[Y] \approx g(\bar{X})$$

First-order approximate variance:

$$Var[Y] \approx Var[X] \left(\frac{dg}{dX}\bigg|_{\bar{X}} \right)^2$$

Second-order approximate mean:

$$E[Y] \approx g(\bar{X}) + \frac{1}{2} Var[X] \frac{d^2 g}{dX^2}\bigg|_{\bar{X}}$$

Second-order approximate variance:

$$Var[Y] \approx Var[X] \left(\frac{dg}{dX}\bigg|_{\bar{X}} \right)^2 - \frac{1}{4} Var^2[X] \left(\frac{d^2 g}{dX^2}\bigg|_{\bar{X}} \right)^2 + E[X - \bar{X}]^3 \frac{dg}{dX}\bigg|_{\bar{X}} \frac{d^2 g}{dX^2}\bigg|_{\bar{X}} + \frac{1}{4} E[X - \bar{X}]^4 \left(\frac{d^2 g}{dX^2}\bigg|_{\bar{X}} \right)^2$$

Example of functions of multiple random variables

1. $Z = \sum_{i=1}^{n} X_i$
   a) $X_i \sim \text{Poisson}(\lambda_i)$,
   b) X_i-independent

   $\Rightarrow Z \sim \text{Poisson}\left(\lambda = \sum_{i=1}^{n} \lambda_i \right)$

2. $Z = X + Y$, $X \sim N(\mu_X, \sigma_X)$, $Y \sim N(\mu_Y, \sigma_Y)$

   $\Rightarrow Z \sim \text{N}\left(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2} \right)$

$Z = aX + bY + c$ $\qquad \Rightarrow Z \sim \text{N}\left(a\mu_X + b\mu_Y + c, \sqrt{a^2\sigma_X^2 + b^2\sigma_Y^2} \right)$

$Z = X - Y$ $\qquad \Rightarrow Z \sim \text{N}\left(\mu_X - \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2} \right)$

Any linear function of normal variates is also a normal variate

Example of functions of multiple random variables

3. $Y = aX^b$, $X \sim \text{log-normal}(\lambda, \xi)$

   $\Rightarrow \ln Y = \ln a + b \ln X$
   $\ln Y \sim N\left(\ln a + b\lambda, \sqrt{b^2 \xi^2} \right)$

   $\Rightarrow Y \sim \text{log-normal}\left(\ln a + b\lambda, \sqrt{b^2 \xi^2} \right)$

4. $Z = XY$, $X \sim \text{log-normal}(\lambda_X, \xi_X)$, $Y \sim \text{log-normal}(\lambda_Y, \xi_Y)$

   $\Rightarrow \ln X = \ln X + \ln Y$
   $\ln Z \sim N\left(\lambda_X + \lambda_Y, \sqrt{\xi_X^2 + \xi_Y^2} \right)$

   $\Rightarrow Z \sim \text{log-normal}\left(\lambda_X + \lambda_Y, \sqrt{\xi_X^2 + \xi_Y^2} \right)$

THE END (probably)

Thank you for the attention!!!!