Lecture 1. Probabilities - Definitions, Examples and Basic Tools

Igor Rychlik

Chalmers
Department of Mathematical Sciences

Risk *is a quantity derived both from the probability that a particular hazard will occur and the magnitude of the consequence of the undesirable effects of that hazard. The term risk is often used informally to mean the probability of a hazard occurring.* *Example*
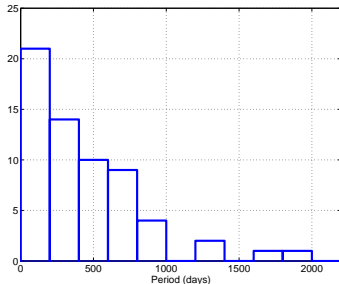
Probabilities *are numbers, assigned to (events) statements about outcome of an experiment, that express the chances that the statement is true.*

Statistics *is the scientific application of mathematical principles to the collection, analysis, and presentation of numerical data.*

Common usages of the concept of probability:

- To describe variability of outcomes of repeatable experiments, e.g. chances of getting "Heads" in a flip of a coin, chances of failure of a component (mass production) or of occurrence of a large earthquake worldwide during one year.

- To quantify the uncertainty of an outcome of a non-repeatable event. Here the probability will depend on the available information. How many years can the ship be safely used?

- To measure the present state of knowledge, e.g. the probability that the detected tumor is malignant.
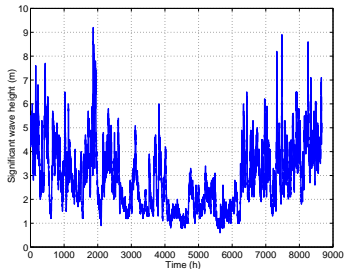
Examples of data

Histogram: Periods in days between serious earthquakes 1902–1977.

Probabilistic questions?

Probabilistic models?

Measurements of Significant wave height,

Jan 1995 – Dec 1995 100-years significant wave?

Probabilities

Term *experiment* is used to refer to any process whose outcome is not known in advance. Consider an experiment.

- ▶ Sample space $\mathcal{S}$: A collection of all possible outcomes.

- ▶ Sample point $s \in \mathcal{S}$: An element in $\mathcal{S}$.

- ▶ Event A: A subset of sample points, $A \subset \mathcal{S}$ for which a statement about an outcome is true.

Rules for probabilities:

$$P(A \cup B) = P(A) + P(B), \quad \text{if} \quad A \cap B = \emptyset.$$

For any event A,

$$0 \le P(A) \le 1.$$

Statements which are always false have **probability zero**, similarly,

always-true statements have **probability one** Example 1 .

Probabilities - Kolmogorov's axioms.

Definition. Let $A_1, A_2, \ldots$ be an infinite sequence of statements such that at most one of them can be true (A_i are mutually excluding); then

$$P(\text{"At least one of } A_i \text{ is true"}) = P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i). \qquad (*)$$

Any function P satisfying (*) taking values between zero and one and assigning

- ▶ value zero to never-true statements (impossible events)
- ▶ value one to always-true statements (certain events)

is a correctly defined probability.

Example 2

A.N. Kolmogorov

A.N. Kolmogorov (1903-1987).

Grundbegriffe der Wahrscheinlichkeitsrechnung.

How to find "useful" probabilities:

- Classically for finite sample spaces $\mathcal{S}$, if all outcomes are equally probable then

  $P(A) =$ number of outcomes for which A is true/number of outcomes

- Employ a concept of independence.
- Employ a concept of conditional probabilities.

If everybody agrees with the choice of P, it is called an **objective probability**. (If a coin is "fair" the probability of getting tails is 0.5.)

For many problems the probability will depend on the information a person has when estimating the chances that a statement A is true. One then speaks of **subjective probability**.

Independence

For a sample space $\mathcal{S}$ and a probability measure P, the events $A, B \subset \mathcal{S}$ are called **independent** if

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

Two events A and B are **dependent** if they are not independent, i.e.

$$P(A \cap B) \neq P(A) \cdot P(B).$$

Example 3

Conditional probability

Conditional probability:

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

The chances that some statement B is true when we *know* that some statement A is true.

Example 4

Law of total probability

Let $A_1, \ldots, A_n$ be a partition of the sample space.

Then for any event B

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)$$

Example 5

Bayes' formula

Again let $A_1, \ldots, A_n$ be a partition of the sample space, i.e. we have n excluding hypothesis and only one of them is true. The evidence is that B is true. Which of alternatives is most likely to be true?

Bayes' formula:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(A_i)}{P(B)}P(B|A_i)$$

Name due to *Thomas Bayes* (1702-1761)

Likelihood: $L(A_i) = P(B|A_i)$ (How likely is the observed event B under alternative A_i?)

Example 6

Odds (fractional odds) [1] for events A_1 and A_2: Any positive numbers q_1 and q_2 such that

$$\frac{q_1}{q_2} = \frac{P(A_1)}{P(A_2)}$$

Let $A_1, A_2, \ldots, A_n$ be a partition of the sample space having odds q_i, i.e. $P(A_j)/P(A_i) = q_j/q_i$. Then

$$P(A_i) = \frac{q_i}{q_1 + \cdots + q_n}$$

The language of odds such as "ten to one" for intuitively estimated risks is found in the sixteenth century, well before the invention of mathematical probability.[1] Shakespeare writes in Henry IV:

> Knew that we ventured on such dangerous seas that if we wrought out life 'twas ten to one.

---

[1]European odds: $q_i = 1/P(A_i)$ - if you bet 1 SEK on A_i then you get q_i SEK if A_i is true (occurs) or loos your bet if A_i is false. Odds 1:4 corresponds to Europiean odds 5.

Bayes' formula formulated using odds

Bayes' formula can be conveniently written by means of *odds*:

$$q_i^{\text{post}} = P(B \mid A_i) q_i^{\text{prior}}$$

Prior odds: q_i^{prior} for A_i **before** B is known.

Posterior odds: q_i^{post} for A_i **after** it is known that B is true.

Example 7

Conditional independence

Events B_1, B_2 are **conditionally independent** given a partition $A_1, A_2, \ldots, A_k$ of the sample space if

$$P(B_1 \cap B_2 | A_i) = P(B_1 | A_i)P(B_2 | A_i).$$

Events $B_1, \ldots, B_n$ are conditionally independent if all pairs B_i, B_j are conditionally independent.

Odds: Recursive updating

Let $A_1, A_2, \ldots, A_k$ be a partition of the sample space, and $B_1, \ldots, B_n, \ldots$ a sequence of true statements (evidences).

If the evidences B are conditionally independent of A_i then the **posterior odds** can be computed recursively using the induction

$$q_i^0 = q_i^{\text{prior}}$$

$$q_i^n = P(B_n \mid A_i) q_i^{n-1}, \qquad n = 1, 2, \ldots$$

Example 8: Waste-water treatment:

Check your intuition:

Are independent events B_1, B_2 always conditionally independent? see for help "click".

Solve problem 2.6.

Examples in this lecture "click"