Probability Theory and Interpretations

Prof. Yan Wang
Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.
yan.wang@me.gatech.edu

Learning Objectives

- To learn the history and evolvement of probability theory
- To understand the different interpretations of probability

Birth of Probability

- The calculus of probability is conventionally dated from July 1654
- when Blaise Pascal wrote Pierre de Fermat on the dice game problem posed by the Chevalier de Méré - the division of stakes between players when a game is interrupted before reaching conclusion
- Pascal and Fermat developed new mathematical techniques to calculate the odds in a number of card and dice games.

Blaise Pascal

Pierre de Fermat

Early Formation of Probability Calculus

- Christian Huyens derived similar results in his 1657 essay, *De Ratiociniis in Ludo Aleae*.
- Pierre Rémond de Montmort calculated the expected gains in several complex card and dice games published in 1708.
- Jacob Bernoulli worked on the subject from 1685 to his death in 1705. His *Ars Conjectandi* , the first systematic and analytically rigorous conspectus of the probability calculus was published in 1713, where the balls-and-urn model was introduced.
- Abraham De Moivre extended the mathematics, first defined and used the term 'probability', and incorporated improved combinatorial relations, series expansion and approximation of distributions in his gambler's bible, the *Doctrine of Chances*, published in 1718.
  - "the comparative magnitude of the number of chances to happen, in respect of the whole number of chances to either happen or to fail, is the true measure of probability"


Abraham De Moivre

Early Applications – Annuities and Insurance

- John Graunt published the first life tables in his *Actual and Political Observations* in 1662 based on the causes of death given by the London Bills of Mortality (instituted in 1562 as a warning of plague epidemics) with the estimation of the population of London.
- The table immediately interested the early probabilists.
- Johann De Witt used Huygen's formulae to calculate annuity rates in 1671.
- Edmund Halley published the first complete empirical mortality tables in 1693, which showed the annuities sold by the British government (6% returning on a seven-year lifetime – chance of death after the age of 6 was widely assumed to be constant) were undervalued.
- De Moivre's theoretical account, *Annuities Upon Lives*, was published in 1725.

Early Applications – Astronomy

- A planet or star has a certain position in the sky, yet changeable atmospheric conditions, together with inaccurate instruments and imperfect eyesight, will always introduce a scatter into a series of measurements.
- The face of the Moon visible from the Earth is not fixed but subject to a variation.
- Astronomers had long debated what to do about these.
- Kepler was in favoring the mean value, as a fair estimate.
- Galileo chose the mode, as the most witnessed.
- Roger Cotes recommended a weighted average according to the errors of single observations.
- Roger Boscovitch used probabilistic principles to combine length measurements of short arcs of the Earth's surface at different latitudes to estimate the shape of the Earth in 1755.
- Thomas Simpson applied De Moivre's series expansion methods to compare probabilities that the mean of observation errors lying in a specified range in 1755.
- Johann Lambert described the instruments' error distributions in 1760s.

Birth of Utility

- In early 1700's, Nicolaus Bernoulli studied the *St. Petersburg problem*, which is the fair price to participate in a game in which a coin is tossed until a head shows, with the prize starting at one dollar if the first toss is a head and doubling after each successive tail.
- The chance of a lengthy run halves with each extra tail, but the pay out on such a sequence is proportionately large.
- The mathematical expectation, and hence entry fee for a fair game, is apparently infinite.
- Yet as Bernoulli pointed out, a reasonable man would not pay any more than a few dollars to take part.

Birth of Utility – cont'd

- In the *St. Petersburg problem*, the probability calculus was not regarded as a direct description of rational judgement.
- In 1738, Daniel Bernoulli introduced the concept of "utility". He argued, the value of money is not constant for a rational man, unless you are particularly greedy or profligate.

Inverse Probability

- Inductive reasoning: effect -> cause
- Started with Jacob Bernoulli but no satisfactory solution, until…
- Thomas Bayes (1764, 3 yrs after his death) "Essay towards solving a problem in the doctrine of chances", published in *Philosophical Transactions of the Royal Society*.
- Inverse probability was popularized by Marquis de Laplace
  - ["Rule of Succession" - in Laplace's 1774 paper] from a infinitely large urn, the first (p+q) draws had yielded p black and q white balls, the probability of a further black ball on the $(p+q+1)^{th}$ draw, assuming that all possible fractions of black balls in the urn were initially equally likely, is $(p+1)/(p+q+2)$.
  - "What is the probability that the sun will rise tomorrow, assuming the probability of rising is fixed but initially unknown, and given that it has risen each day without fail in 5,000 years of recorded history?"

Inverse Probability – cont'd

- Karl Friedrich Gauss in 1809 applied the inverse argument to the general problem of fitting a curve of a given form to a series of data.
- The most probable fit is that for which the sum of squares of the differences between each data point and the corresponding value of the curve is a minimum, given that the error follows the normal distribution.
- Similarly Adrien Marie Legendre published the least-squares coefficients in 1805.

Inverse Probability – cont'd

- Laplace read Gauss's work in 1810 and was provoked to rejoin the fray in the monumental *Théorie Analytique des Probabilités* of 1812
  - used his central limit theorem to generalize Gauss's justification of the normal distribution to any situation
  - presented a complete synthesis of the probability of causes and the analysis of errors
  - first associated a probability with the degree of a belief
  - formed the formal development of the inverse probability calculus as a method of rational judgement
  - illustrated by examples of orbit of comet, birth-rates, voting procedures, decision making processes, legal testimony, medical problems, etc.
  - "It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge… The most important questions of life are, for the most part, really only problems of probability." - Marquis de Laplace

Laplacean Probability

- Laplace defended the theory of probability as "good sense reduced to a calculus".
- Like Pascal, Descartes, and indeed most of the philosophers and scientists of that time, Laplace regarded the material world as strictly deterministic.
- Jacob Bernoulli had predicted that exact calculations would one day make gambling on dice obsolete.
- Laplace declared that a sufficiently vast intelligence, processing a complete knowledge of the state of the world at any instant, could predict its evolution with certainty.
- Although causes are a matter of certainty, human's imperfect knowledge of them is not. In other words, probability is not in the world, but reflects our incomplete knowledge of the world.
- Laplace's ***Principle of Insufficient Reason***: *when we have no knowledge of causes, or reason to favor one hypothesis over another, we should assign each the same probability*.
- As the connection between the epistemic and frequency interpretations, when backed by the law of large numbers, a probability could be measured as an objective long-run frequency.
- The success of Laplacean probability in early 1800's was largely due to the universal acceptance of the Gaussian error or normal curve and the method of least squares by the physicists.

Eclipse of Laplacean Probability

- Natural philosophers in 1800's dismissed outright the balls-and-urn model as illegitimate for natural events (sun rise is different from a well-defined game of chance).
- Assumption of probability providing a unique scale of rationality was also under fire. E.g. the stand-off over the *St.Petersburg* problem:
  - Daniel Bernoulli argued that rational thought was exemplified by the preferences of shrewd dealers and could be captured by the concept of utility
  - Nicolaus Bernoulli, in contrast, instead that reasonable behavior could only be underpinned by a concept of justice
- During the period of social upheaval of the France Revolution, the idea of probability as rationality was widely regarded as a dangerous or foolhardy extension of the theory.

Eclipse of Laplacean Probability – cont'd

- Philosophers and mathematicians in 1800's also began to attack the Laplacean, such as
  - Ignorance cannot, by definition, ground knowledge as in Laplace's *Principle of Insufficient Reason*)
  - John Stuart Mill (1843), Richard Leslie Ellis (1844) regarded the analysis as spurious that knowledge was based solely on experience since conclusions were drawn by relating matters of fact from ignorance was scandalous.
  - Gorge Boole (1854, in his book *An Investigation of the Laws of Thought*), similar to De Morgan, treated probability as a branch of logic and thus applicable to the relationship between propositions. He denied that every proposition could be assigned a definite numerical probability with respect to a body of data.
  - In applications of jury selection and conviction, Condorcet and Poisson (1830s) argued that the Laplacean way of selecting prior probabilities made it easier to convict the guilty.

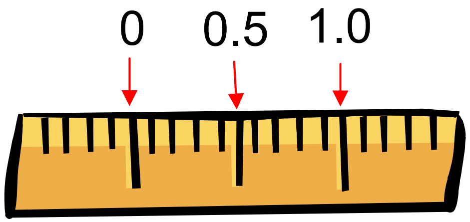
The Rise of Frequency Interpretation

- Since any rational man would form the same opinion when confronted with the same evidence, his subjective degree of belief could be identified with the objective aggregate or long-run frequency.
- Poisson in his *Recherches sur la Probabilité des Jugements* of 1837 distinguished between two forms of probability, using probability for the epistemic sense and chance for the objective frequencies of events.
- Shortly after that, empiricists including John Stuart Mill, Richard Leslie Ellis, Antoine Augustin Cournot, Jacob Friedrich Fries also promoted frequency theories of probability.
- Probabilities, the frequentists insisted, applied only to events or measurements, not hypotheses or causes.

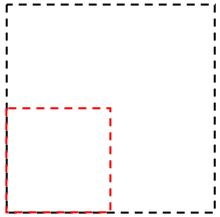
The Rise of Frequency Interpretation - cont'd

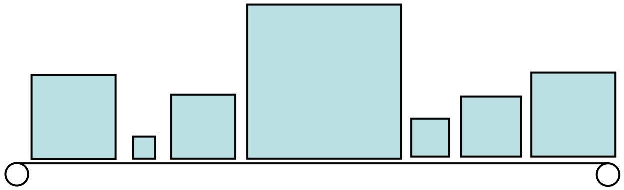
- Probabilities could be used for modeling and estimation, not inference or matters of judgment.
- As a frequency or ratio, a probability referred collectively to the population, and could not be transferred to an individual item or member (such as individual insurance premium).
- "Equal Distribution of Ignorance" was regarded as the major problem of Laplacean probability.
  - Prior probabilities based on belief were indeterminate unless we had knowledge or belief to distribute.
  - It leads to inconsistency when ignorance about correlated parameters, which cannot both be uniformly distributed.
  - The economist A.A. Cournot wrote that probabilities had little value other than the "charm of speculation" (for gambling) unless based on experience; Fries agreed.

van Fraassen's Cube Factory Paradox

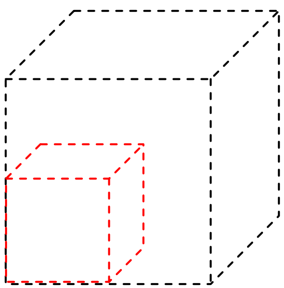


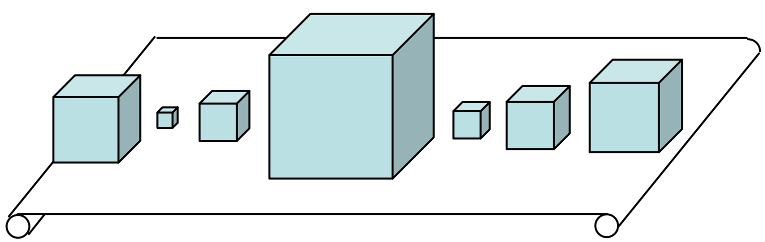
$$P(\textit{side length} \text{ of a randomly chosen cube} \in [0,1/2]) = ?$$





$$P(\textit{face area} \text{ of a randomly chosen cube} \in [0,1/4]) = ?$$





$$P(\textit{volume} \text{ of a randomly chosen cube} \in [0,1/8]) = ?$$

Interpretations & Debates in mid 1800's

- Social Sciences
  - Frequency interpretation drew much support through 19th century with the advancement of social statistics. Many sociologists were concerned more with aggregate behaviors.
  - Adolphe Quetelet stated that the law of large numbers showed the effects of free will, though important for an individual, would average out at the macro level. He was convinced that normal distribution was universal.
  - Auguste Comte however argued mankind could not be expressed simply as number as reductionism is inherent in quantification.

Interpretations & Debates in mid 1800's

- Medicine
  - Medical applications did not have large sample sizes. Practitioners (e.g. Risueño d'Amador, François Double) saw medicine as a matter of fine judgment and a trained intuition, not equations and rule books. "one size does not fit all."

Interpretations & Debates in mid 1800's – cont'd

- Physical Sciences
  - James Clerk Maxwell developed kinetic theory on statistical physics with his argument that since knowledge of the physical world was always statistical in character, it could never be complete.
  - Poincaré: "what is chance for the ignorant is not for the scientist. Chance is only the measure of our ignorance." He believed there was nothing inherently random about the universe. Physical causes determined all events.
  - Ludwig Boltzmann was also a strict determinist. He tried to reconcile his mechanistic philosophy with the irreversibility of the second law of thermodynamics. His use of probability did not imply uncertainty but a description of a large number of molecules.

Interpretations & Debates in mid 1800's – cont'd

- Biological Sciences
  - Different from Charles Darwin's variation theory about species where the cause was unknown rather than random, Francis Galton (Darwin's cousin) instituted the 'biometric' school and studied inheritance and variation from data gathered from large biological surveys.
  - French criminologist Alphonse Bertillon recommended using correlation to identify criminals according to the sizes of their body parts, assuming that finger, foot, and arm measurements were independent.

Interpretations & Debates in late 1800's

- Karl Pearson was convinced that "evolution was a statistical problem".
- Sharing Galton's reverence for the normal curve as representing stability, Pearson also saw multinomial distributions (crabs in Plymouth Sound) with his famous χ^2 goodness-of-fit test.
- His positivism view was that statistics – not religion or philosophy – was the sole route to reliable knowledge. Probability is a 'degree of credibility'.

Interpretations & Debates in late 1800's – cont'd

- John Venn, in his *The Logic of Chance*, insisted that probabilities must only apply to large ensembles, since a single event could fall into several different categories.
- He devoted a chapter to ridiculing conclusions draw from the *Principle of Insufficient Reason* and *Laplace's Rule of Succession*. As he wrote,
  - "the subjective side of Probability, therefore, though very interesting and well deserving of examination, seems a mere appendage of the objective, and affords in itself no safe ground for science of inference."

Interpretations & Debates in late 1800's – cont'd

- William Donkin defended degree-of-belief probability that with the *Principle of Insufficient Reason* as an axiom, the equation of inverse probability gave a satisfactory fit with several observed features of the scientific method.
- George Boole and William Stanley Jevons, on the other hand, deplored naïve applications of the *Principle of Insufficient Reason* and preferred to think of probability as a relation in logic and a 'quantity of knowledge'.
- John Herschel demonstrated the applications of inverse probability for real scientific predictions in his textbook *Outlines of Astronomy*.
- William Sealy Gosset too defended inverse probability, occasionally guessing at forms of prior (polynomial priors) that could reasonably represent experience.

Interpretations in early 1900s

- Determinism was still a popular view among physicists at the turn of the century, e.g.
  - In 1900, Rutherford showed that Nuclei decayed at random and the rate was proportional to the size of a sample, consistent with each atom having a fixed chance of decay per time interval.
  - In 1900, Max Planck used probability to derive the law of black body radiation, the distribution across wavelengths of the energy radiated from an idealized black surface.
  - In 1905, Albert Einstein gave a complete statistical account of the Brownian motion effect.
- until …

Interpretations in early 1900's – cont'd

- In 1926, Max Born interpreted Erwin R.J.A. Schrödinger's wave account in quantum mechanics as a sort of probability.
- Probability, Born declared, was fundamental, and "from the point of view of quantum mechanics there exists no quantity which in an individual case causally determines the effect of collision."
- At Copenhagen, He argued that observables such as position and momentum were not fundamental, but emergent properties of the irreducible probability distribution described by the Schrödinger equation.

Interpretations in early 1900's – cont'd

- Arthur Edington quickly adopted the acausal implications of quantum mechanics and proposed a connection between physics and consciousness:
  - Determinism was a mental artifact imposed on the world;
  - Freedom from it would reveal a deep connection between the elementary particles and animating spirit of the mind.
- Edington's 'selective subjectivism' held that it was the power of our minds and sensory equipment that decomposed the universe into types and numbers of particles.

Interpretations in early 1900's – cont'd

- Paul Dirac wrote in 1930 that when "*an observation is made on any atomic system … the result will not in general be determinate, i.e., if the experiment is repeated several times under identical conditions several results may be obtained. If the experiment is repeated a large number of times it will be found that each particular result will be obtained a definite fraction of the total number of times, so that one can say there is a definite probability of its being obtained any time the experiment is performed.*"
- Even those dubious about indeterminism, such as Norman Campbell, recognized quantum mechanical probabilities were more objective and foundational than the old degree-of-belief.
- Von Neumann wrote in 1955 that probability is fundamental, and statistical ensembles are necessary for "establishing probability theory as the theory of frequency."

Ronald Aylmer Fisher's Interpretation

- Probability was regarded a frequency ratio in a "hypothetical infinite population", whereas "likelihood" of an event is a numerical measure of rational belief.
- Fisher in his 1915 paper proposed the *Method of Maximum Likelihood*, although numerically similar to an application of inverse probability with the assumption of uniform priors, which was a new concept.
- He argued probability can be applied to small samples, and the inverse probabilists confused probabilities of the sample with those of the parent population.
- "*If the population of interest is itself drawn from a known super-population, we can deduce, using perfectly direct methods, the probability of a given population and hence of the sample. But if we do not know the function specifying the super-population, we are hardly justified in simply taking it to be constant. Not only is this choice of* ***a priori*** *distribution completely baseless, but the restatement of our population using different parameters would lead to a different function. A* ***prior*** *probability distribution can only be verified by sampling from reference set.*"
- The *Principle of Insufficient Reason* is arbitrary and inconsistent.

Harold Jeffreys' Interpretation

- The *Principle of Insufficient Reason* should be decoupled from the rest of probabilistic machinery ("similar individuals are likely to be associated.").
- Assuming a uniform prior probability to the absolute value of precision constant h is inappropriate. Rather, it should be applied to the order of magnitude as dh/h.
- With a large enough sample, the precise choice of prior probability had little influence on the calculated posterior probability.
- Probability should be regarded as a model of scientific inquiry. "Simplicity Postulate": the simpler form the scientific laws have, the greater its prior probability.
- Inverse probability was the only way to account for learning from experiences.
- "error" is a verbal matter and observations can never be 'wrong'. Inverse probability is to seek a most probable 'true' value and precision from a series of measurements drawn from an unknown distribution.

Fisher-Jeffreys Debate (1932-1934)

- Fisher (F) studied agriculture and genetics by *experiments*, whereas Jeffreys (J) studied geophysics and astronomy with *observations*.
- Prior probability
  - **J**: **Prior** probability was a rough way to encode previous knowledge or information as a convenient starting point for calculation. It is not worth assessing precisely. An ***a priori*** distribution, in contrast, was a unique measure constructed independent of experience.
  - **F**: **Prior** probabilities were just objective statements about frequencies and could be evaluated, whereas ***a priori*** probabilities could not.
  - **J**: complete ignorance is a state of knowledge, just same as the statement that a vessel is empty.
  - **F**: Ignorance is a lack of information concerning the population of all possible measuring conditions. Supplying an ***a priori*** distribution from ignorance was a contradiction.

Fisher-Jeffreys Debate (1932-1934) - cont'd

- <u>The Principle of Insufficient Reason</u>
  - **F**: The Principle is not only subjective and impossible to verify experimentally, but arbitrary and inconsistent too. It was too flimsy a base for the entire weight of a numerical theory of probability. "*A man is as likely to be an inhabitant of Ireland as of France. On the same principle he is as likely to be an inhabitant of the British Isles as of France. Inconsistency arises with the two that he is twice as likely to be an inhabitant of the British Isles as of France.*"
  - **J**: The Principle was merely a starting point for calculation and not to rule out any new possibility that subsequent evidence might suggest. He introduced 'the Principle of Non-sufficient Reason' that was often used in practice to input numbers.

Fisher-Jeffreys Debate (1932-1934) - cont'd

- Definition of Probability
  - **F**: Epistemic interpretation forms no basis for science. To be useful, a probability must admit of experimental verification, as can frequencies in gaming and genetic applications. This is not possible in Jeffreys's "subjective and psychological" version. Second, the epistemic interpretation relies on the unjustified equation of a logical relationship between propositions with the relative degree of belief in those propositions. Third, inverse probability is too blunt a tool for the many and various sorts of uncertainty that occur in practice. In contrast, frequency definition in his papers was rigorous and unambiguous.
  - **J**: Hypothetical infinite population does not exist. Its properties would have to be inferred from the finite facts of experiences. Probability must precede any judgement about frequencies.

Fisher-Jeffreys Debate (1932-1934) - cont'd

- Role of Science: inference and estimation
  - **F**: Epistemic interpretation is subjective and forms no basis for science.
  - **J**: Science is just a matter of individual experience and is subjective. Different people interpret their essential private sensations in terms of the same 'reality'.
  - **F**: if the prior probability is truly unimportant for large samples, it is irrelevant and should have no part in our reasoning; if it affects our reasoning, then it is misleading.
  - **J**: The point of sampling is to move from the sample to a statement of the class, and this cannot be done without background assumptions.

Kolmogorov's Axiomatic Probability

- The axioms of Probability $p : \mathcal{A} \mapsto [0,1]$
  - (Non-negativity) $p(E) \geq 0 \; (\forall E \in \mathcal{A})$
  - (Normalization) $p(\Omega) = 1$
  - (Finite additivity) For countable mutually disjoint events E_i's $p(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} p(E_i)$

Categories of Modern Interpretations

- Subjective Bayesianism
- Frequency Interpretation
- Logical Interpretation
- Propensity Interpretation

Subjective Interpretation

- Probability is the 'degree of belief'
- Probability is associated with individuals and is personal.
- In order to make rational decisions, the belief should obey the axioms of probability.
- Your degree of belief in an event E is p, iff p units of utility is the fair price at which you would buy or sell a bet that pays 1 unit of utility if E, 0 if not E.

Dutch Book Arguments

- Championed by Bruno de Finetti with his coherent prevision interpretation
- For any proposition A, there is a number p(A) such that you are willing to accept any bet with betting quotient p(A).
- Dutch book arguments show that the rationality requires p to satisfy the axioms of probability.
- A set of bets, in which p does not satisfy the axioms of probability thus will always lead to a loss, is called a *Dutch book*.
- Rational agents will never accept a Dutch-book bet.

Dutch Book Arguments – cont'd

- Ex1. Let H denote that a coin will land heads on the next toss, and suppose that for you $p(H)$=.6 and $p(\neg H)$=.5.
  - You are willing to pay \$0.6 for a bet on H that pays you \$1 if H.
  - Similarly, you are willing to pay \$0.5 for a bet against H that pays you \$1 if $\neg H$.
  - A bookie sells you one bet on H and another bet on $\neg H$, who collects \$1.1 from you, and immediately hands \$1 back to you. No matter whether H or $\neg H$ occurs, you have a sure loss.

Dutch Book Arguments – cont'd

- Ex2. Let N denote that Netherlands wins the next world cup and S that Spain wins the next world cup. You post your betting quotients as: $p(N)=.5$, $p(S)=.2$, $p(N \text{ or } S)=.6$
  - A bookie sells you a bet on N for \$1, a bet on S for \$1, and buys a bet from you on (N or S) for \$1. You gain

(-\$0.5+\$1)+(-\$0.2)+(\$0.6-\$1)=-\$0.1	if N occurs
(-\$0.5)+(-\$0.2+\$1)+(\$0.6-\$1)=-\$0.1	if S occurs
(-\$0.5)+(-\$0.2)+(\$0.6)=-\$0.1	if other

Dutch Book Arguments – cont'd

- Dutch book argument assumptions
  - You must post betting quotients of all events
  - You must accept all bets anyone wants to make at your posted quotients.
  - You are risk neutral.
  - There is a bookie who will bankrupt you if your posted quotients do not satisfy the axioms of probability
- Suppose that you know if you post betting quotients of .6 on H and .5 on $\neg H$, the bookie will make a bet with \$1. You also know if you post quotients of .5 on H and .5 on $\neg H$, the bookie will make the bet \$100.
- If you are *risk averse*, and prefer to accept the sure loss of \$0.1 than to gamble on \$100, it would be rational for you to post quotients that violate the probability axioms.

Dutch Book Arguments – cont'd

- Dutch Book arguments of rationality also were extended to include other principles
  - Conditionalization (Paul Teller, 1973; 1976): P(A)=P(A|E)P(E)

  If your current probability function p, and if q is the probability function you would have if you learned E and nothing else, then q should be identical to p.

  - Inclusion of utility (Shimony 1955, Horwich 1982)

  Instead of monetary amounts, betting quotients are in terms of utilities.

Frequency Interpretation

- Probabilities should only deal with experiments that are random and well-defined.
- The probability of a random event denotes the relative frequency of occurrence of an experiment's outcome, when repeating the experiment.
- Probability is the relative frequency ‘in the long run’ of outcomes for population or samples.
- Probability does not make sense when applying to an individual.

Logical Interpretation

- Probability is the 'degree of implication'.
  - the degree to which hypothesis H is confirmed by evidence E
- As an extension of classical probability, the major contributors include Carnap, Jeffreys, Keynes, and W.E. Johnson.
- Probability is part of inductive logic.
  - A language consists of a finite number of logically independent monadic *predicates* (*properties*) applied to countably many individual *constants* (*individuals*) or *variables*.
  - Every sentence H is equivalent to a disjunction of mutually exclusive state descriptions, and its a priori probability measure $m(H)$ is thus determined.
  - m in turn will induce a confirmation function $c(H,E)$ according to the conditional probability $c(H,E) = m(H\&E)/m(E)$.

Logical Interpretation – cont'd

- e.g. A language has three individuals, *a*, *b* and *c*, and one predicate *F*.

State Description	*Structural Description*	*weight*	*m*
Fa&Fb&Fc	"everything has property F"	1/4	1/4
¬Fa&Fb&Fc	"two F's, one ¬F"	1/4	1/12
Fa&¬Fb&Fc			1/12
Fa&Fb&¬Fc			1/12
¬Fa&¬Fb&Fc	"one F, two ¬F's"	1/4	1/12
¬Fa&Fb&¬Fc			1/12
Fa&¬Fb&¬Fc			1/12
¬Fa&¬Fb&¬Fc	"nothing has property F"	1/4	1/4

Logical Interpretation – cont'd

- Hypothesis H: Fc - "c always has property F"
  - Prior probability $m(H) = 1/4+1/12+1/12+1/12=1/2$
- Evidence E: exam individual "a" and find the property F
  - $m(H\&E)=1/4+1/12=1/3$
  - $m(E)=1/4+1/12+1/12+1/12=1/2$
- Confirmation
  - $c(H,E) = m(H\&E)/m(E) = 2/3$

❑ Carnap's continuous confirmation function:

- for a family of predicates $\{P_n\}$, $n = 1, \ldots, k$ ($k > 2$)

c_λ(individual s+1 is P_j, s_j of the first s individuals are P_j)

$= (s_j + \lambda/k)/(s + \lambda)$

where λ is a positive real number. The higher the value of λ is, the less impact the evidence has.

Propensity Interpretation

- Championed by Popper, Miller, Giere, Gillies, and others.
- Probability is regarded as a physical propensity, or disposition, or tendency of a given type of physical situation to yield an outcome of a certain kind, or to yield a long run relative frequency of such an outcome.
- Explicitly intended for single case: "the probability that *this* radium atom decays in 1500 years is ½".
- long-run propensity vs. single-case propensity
  - long-run propensities are associated with repeatable conditions, and are regarded as propensities to produce in a long series of repetitions of these conditions frequencies which are approximately equal to the probabilities.
  - Single-case propensities are the ones to produce a particular result on a specific occasion

Propensity Interpretation – cont'd

- Long-run propensities are tendencies to produce relative frequencies with particular values. The propensities are not the probability values themselves.
- Single-case propensities *are* the probability values.
- The main challenge facing propensity theories is to say exactly what propensity means.
- Yet, it was argued that propensity is defined as the theoretical role of which it plays in science.
- Similar to physical magnitudes such as electrical charge cannot be explicitly defined either, in terms of more basic things, but only in terms of what they do (such as attracting and repelling other electrical charges), propensity fills the various roles that physical probability plays in science.
- 'Principal Principle': When a propensity is known, it constrains rational belief to take the same numerical value.

Summary

- The same mathematical form of Probability theory has multiple interpretations.
- There have been continuous debates between schools, particularly between subjectism and frequentism, since the start of the theory four hundred years ago.

Interpretations

Further Readings

- David Howie (2002) *Interpreting Probability: Controversies and Developments in the Early Twentieth Century*. Cambridge University Press, Cambridge, UK.
- Hájek A. (2001) Probability, logic, and probability logic. In L. Goble (ed.) *The Blackwell Guide to Philosophical Logic*. Blackwell, Oxford. Ch.16. pp. 362-384.
- Patrick Maher (1993) *Betting on Theories*. Cambridge University Press, Cambridge, UK.
- Stanford Encyclopedia of Philosophy, Interpretations of Probability http://plato.stanford.edu/entries/probability-interpret/