Purdue-NCKU program

Lecture 0
Probability Theory: The Logic of Sciences

Dr. Qifan Song

Deductive and plausible reasoning

- Deductive reasoning
  - "if A is true, then B is true" is equivalent to

    "if B is false, then A is false"
  - Deterministic reasoning
- plausible reasoning
  - "if A is true, then B is true" implies that

    "B is true, therefore, A becomes more plausible.", or

    "A is false, therefore, B becomes less plausible."
  - The evidence does not prove that A is true, but verification of one of its consequences does give us more confidence in A.

Deductive and plausible reasoning

A dark night a policeman hears a burglar alarm, and sees a jewelry store with a broken window. Then a gentleman wearing a mask comes crawling out through the broken window, carrying a bag. The policeman immediately decides that this gentleman is dishonest. What is the reasoning process

- "If A is true, then B becomes more plausible" implies

  "If B is true, therefore, A becomes more plausible"

- In spite of the apparent weakness of this argument, we recognize that the policeman's conclusion has a very strong convincing power.

- The brain, in doing plausible reasoning, not only decides whether something becomes more plausible or less plausible, but that it evaluates the degree of plausibility in some way.

Numerical Rules

Boolean algebra

- $A, B, C, \ldots$ means some proposition which can be evaluated (True of False)
- AB means logical product, denoting the proposition "both A and B are true".
- $A + B$ called the logical sum, standing for "at least one of the propositions, A, B is true". By def, $B + A$ and $A + B$ are the same
- the equal sign is used to denote equal truth value: $A = B$ means A and B are either both True or both False.
- A^c means the denial (opposite) of A, i.e., A is false.

Boolean algebra

- $(AB)^c$ means AB is false
- A^cB^c means both A and B are false.
- $AA = A + A = A$
- $AB = AB$, $A + B = B + A$
- $A(BC) = (AB)C = ABC$, $A + (B + C) = (A + B) + C = A + B + C$
- $A(B + C) = AB + AC$; $A + BC = (A + B)(A + C)$
- $(AB)^c = A^c + B^c$, $(A + B)^c = A^cB^c$

Boolean algebra

Implication

- $A \Rightarrow B$ means "A implies B".
- It does not assert that either A or B is true
- It means only that AB^c is false, or, what is the same thing, $(A^c + B)$ is true.
- This can be written also as the logical equation $A = AB$. That is, given $A \Rightarrow B$, if A is true then B must be true; or, if B is false then A must be false.
- On the other hand, if A is false, it says nothing about B: and if B is true, it says nothing about A.

Adequate sets of operations

Q: Is logic sum/product/denial sufficient for all deductive logic?

For example

- Given an A whose value can be 0 or 1.
- f maps A to logic values $\{0, 1\}$
- There is 4 possible f (maps 0 to 0 or 1, maps 1 to 0 or 1)
- All 4 functions can be represented by these three logic operation (i.e., A, A^c, $A + A^c$, AA^c)

Adequate sets of operations

- Given multiple propositions, say A, B, C, D, there should be 2^{2^4} different deductive logic function f's.

- These functions can be constructed by

1. basic conjunctions $ABCD, A^cBCD, AB^cCD, ..., A^cB^cC^cD^c$;

2. their logic sum;

3. and "always False": AA^c

{AND, OR, NOT} suffice to generate all possible logic functions;
OR can be derived from AND and NOT: $A + B = (A^cB^c)^c$;
Or can be derived from only one logic operation

Degrees of plausibility

- Degrees of plausibility are represented by real numbers (quantitative study)
- $A|B$ means "the conditional plausibility that A is true, given that B is true" or just "A given B". It stands for some real number.

  If B is an always true statement $A|B = A$
- $(A + B)|CD$ represents the plausibility that at least one of the propositions A and B is true, given that both C and D are true
- $(A|B) > (C|B)$ says that, given B, A is more plausible than C.
- We need to define some quantitative rule of reasonable plausibility

Quantitative Rule of common sense

- If $A|C' > A|C$ but $B|AC' = B|AC$ then

  $AB|C' \geq AB|C$ and $A^c|C' < A^c|C$

This qualitative requirement simply gives the 'sense of direction' in which the reasoning is to go; it says nothing about how much the plausibilities change, except that our continuity assumption (which is also a condition for qualitative correspondence with common sense) now requires that if $A|C$ changes only infinitesimally, it can induce only an infinitesimal change in $AB|C$ and $A|C$.

Quantitative Rule of consistency

- If a conclusion can be reasoned out in more than one way, then every possible way must lead to the same result.
- Always takes into account all of the evidence it has relevant to a question. It does not arbitrarily ignore some of the information, basing its conclusions only on what remains.
- Always represents equivalent states of knowledge by equivalent plausibility assignments. That is, if in two problems our state of knowledge is the same, then it must assign the same plausibilities in both.

Mathematical Goals

- Representation of degrees of plausibility by real numbers
- Qualitative correspondence with common sense
- Consistency

The product rule

We want to relate the plausibility of the logical product AB to the plausibilities of A and B separately. $((AB)|C)$ should be determined by either

- $(B|C)$ and $(A|BC)$; or
- $(A|C)$ and $(B|AC)$

The product rule

A function F combines the marginal and conditional plausibility, i.e., $(AB|C) = F((B|C),(A|BC))$, $(AB) = F(A,(B|A))$ etc

- The rule of common sense implies that $F(x, y)$ is increasing w.r.t. x or y
- Let $M \in (0, +\infty)$ be the maximum plausibility, then $F(M, M) = M$.
- The consistency implies that $F((B|C),(A|BC)) = F(A|C,(B|AC))$, $F(A,(B|A)) = F(B,(A|B))$ etc.

  $(ABC|D) = F\{F[(C|D),(B|CD)],(A|BCD)\}$ and $(ABC|D) = F[(C|D),(AB|CD)] = F\{(C|D), F[(B|CD),(A|BCD)]\}$, i.e.,

$$F[F(x, y), z] = F[x, F(y, z)]$$

Solving F

$$F[F(x,y),z] = F[x,F(y,z)]$$

- A trivial solution is $F \equiv$ Constant, but this solution is not increasing

- Another easy solution is that $F(x,y) = w^{-1}(w(x) + w(y))$ (or $w(F(x,y)) = w(x) + w(y)$), where w is any increasing function. But it cannot meet $F(M,M) = M$.

- The third solution is $F(x,y) = w^{-1}(w(x) \times w(y))$ (or $w(F(x,y)) = w(x) \times w(y)$), where w is any increasing function. To meet $F(M,M) = M$, we must have $w(M) = 1$.

  Product rule: $w(AB|C) = w(A|BC)w(B|C) = w(B|AC)w(A|C)$, with $w \in [0,1]$

Sum rule

We want to relate A and A^c. More specifically, let $u = w(A|B), v = w(A^c|B)$, and we want to find the function S, such that $S(u) = v$

- By common sense S is deceasing, $S(0) = 1$ and $S(1) = 0$
- By product rule: $w(AB|C) = w(A|C)w(B|AC)$, $w(AB^c|C) = w(A|C)w(B^c|AC)$.

  $w(AB|C) = w(A|C)S[w(B^c|AC)] = w(A|C)S[w(AB^c|C)/w(A|C)]$
- Similarly

$$w(A|C)S\left[\frac{w(AB^c|C)}{w(A|C)}\right] = w(B|C)S\left[\frac{w(A^cB|C)}{w(B|C)}\right]$$

- We set $B^c = AD$ for any statement D, then $w(AB^c|C) = w(B^c|C) = S[w(B|C)]$, $w(BA^c|C) = w(A^c|C) = S[w(A|C)]$
- Let $x := w(A|C), y := w(B|C)$, it leads to

$$xS[S(y)/x] = yS[S(x)/y]$$

Solve S

$$xS\left[\frac{S(y)}{x}\right] = yS\left[\frac{S(x)}{y}\right]$$

- Taking $y = 1$ leads to $S[S(x)] = x$ (by the facts $S(0) = 1$ and $S(1) = 0$)
- $S = S^{-1}$.
- It can be show that the solution must be $S(x) = (1 - x^m)^{1/m}$ for some positive m.
- Sum rule: $w^m(A|B) + w^m(A^c|B) = 1$

Probability

- Define probability p as w^m
- Product rule $p(AB|C) = w^m(AB|C) = w^m(A|BC)w^m(B|C) = p(A|BC)p(B|C)$, or by setting $C =$ an always true statement, $p(AB) = p(A|B)p(B)$
- Sum rule: $p(A|B) = w^m(A|B) = 1 - w^m(A^c|B) = 1 - p(A^c|B)$, or by setting $C =$ an always true statement, $p(A) = 1 - p(A^c)$
- $p(A) \in [0, 1]$
- $p(A + B|C) = 1 - p(A^cB^c|C) = 1 - p(A^c|C)p(B^c|A^cC) = 1 - p(A^c|C)[1 - p(B|A^cC)] = p(A|C) + p(A^cB|C)$
  $= p(A|C) + p(B|C)p(A^c|BC) = p(A|C) + p(B|C)[1 - p(A|BC)]$
  $= p(A|C) + p(B|C) - P(B|C)p(A|BC) = p(A|C) + p(B|C) - p(AB|C)$

  or $p(A + B) = p(A) + p(B) - p(AB)$

Subjective and Objective probability

- Subjective probability: any measure of uncertainty satisfying the probability atoms in the above slide.

  My pregnant sister, with 60% chance, is carrying a boy, with 40% chance is carry a girl.

- Objective probability: Predictive statement for future or random events. It should natural satisfies the probability atoms

  If I flip a dice, the chance to get 2 or 4 is 33.33%; the chance not to get 2 and 4 is 66.66%, then chance of getting a 2 = (the chance of getting a 2 or 4) $\times$ (the chance of getting a 2, given that I already get a 2 or 4)

Representation under the Set theory

- All proposition A, B, C can be view as a subset of whole space S
- $S = A + A^c = B + B^c = ...$ means "always true" or logic sum of all possible proposition
- $S = \{e_1, e_2, \ldots ...\}$ where e_i's enumerate all possible outcomes. S means one of the outcome is true; i.e., always true
- $A = \{e_j, e_k, \ldots ..\} \subset S$ means one of e_j, $e_k, \ldots$ is true
- logic sum means union of sets
- logic product means intersection of sets
- denial means complement of sets
- Example: $p(A \cup B) = p(A + B) = p(A) + p(B)$ if sets A and B are disjoint.

Outcome with equal probability

- $S = \{e_1, e_2, \ldots\ldots e_n\}$ and $A_i = \{e_i\}$
- Equal chance for all: $p(A_i) = 1/n$
- By prob atoms: $p(A) = |A|/n$ where $|A|$ means the number of elements in A
- Counting problem

Exercise

- Q: what is the chance of a Full House hand (three cards of one rank and two cards of another rank) in a poker game (13 ranks of diamond/heart/spade/club cards, no Jokes)
- Ans: $n = \binom{52}{5}$, and $|A| = 13 \cdot 12 \cdot \binom{4}{3} \cdot \binom{4}{2}$

Bayes rule

- B_i's are disjoint and $\cup B_i = S$.
- $p(A|B_i)$ is known, how to compute $p(B_i|A)$
- $p(B_i|A) = p(B_iA)/p(A) = [p(B_i) * p(A|B_i)]/p(AS) = [p(B_i) * P(A|B_i)]/\sum_j p(AB_j) = [p(B_i) * p(A|B_i)]/[\sum_j p(B_j) * p(A|B_j)]$
- Q: The proportion of some disease among population is 1%. A test has a chance of 99% to give a correct diagnose. If a random person is tested positive, what is the chance that s/he indeed have the disease?
- Ans: $p(+|D) = p(-|D^c) = .99$, $p(D) = .01$, $p(D|+) = p(D)p(+|D)/[p(D)p(+|D) + p(D^c)p(+|D^c)] = 0.5$

Independence

- A and B are independent is $p(A|B) = p(A)$ or $p(B|A) = p(B)$, or $p(AB) = p(A)p(B)$
- A and B are independent, then $p(A + B) = p(A) + p(B) - p(A) * p(B)$.
- Q: Independent trials with prob p of success (i.e., Bernoulli trial). What is probability to have the first success with in 3 trials
- Ans: $p + p(1 - p) + p(1 - p)^2$
- Q: One Tien Kung missile has 70% interception rate, how about 3 missiles
- Ans: $p(\text{all 3 miss}) = 0.3*0.3*0.3$, $p(\text{at least one hits on target}) = 1 - 0.3^3 = 0.973$

one famous exercise

A reward is behind one unknown door among Doors A, B and C. You make a random guess, then the game host will point one door which is not your choice and behind which there is no reward.

- Q1: what is probability that your original guess is correct?
- Q2: what is probability that your change your guess and are correct?