COMP3702/7702 Artificial Intelligence
Probability Review

Hanna Kurniawati

How to quantify uncertainty ? Probability to the rescue…

Bertsekas & Tsitsiklis, Introduction to Probability.

- FATHER(F): Nurse, what is the probability that the drug will work?
- NURSE (N): I hope it works, we'll know tomorrow.
- F: Yes, but what is the probability that it will?
- N: Each case is different, we have to wait.
- F: But let's see, out of a hundred patients that are treated under similar conditions, how many times would you expect it to work?
- N (somewhat annoyed): I told you, every person is different, for some it works, for some it doesn't.
- F (insisting): Then tell me, if you had to bet whether it will work or not, which side of the bet would you take?
- N (cheering up for a moment): I'd bet it will work.
- F (somewhat relieved): OK, now, would you be willing to lose two dollars if it doesn't work, and gain one dollar if it does?
- N (exasperated): What a sick thought! You are wasting my time!

Probabilistic Modeling

- View:
  - Experiments with random outcome.
  - Quantifiable properties of the outcome.
- Three components:
  - Sample space: Set of all possible outcomes.
  - Events: Subsets of sample space.
  - Probability: Quantify how likely an event occurs.

Probability

- Probability: A function that maps events to real numbers satisfying these axioms:
  1. Non-negativity: $P(E) \geq 0$, where E is an event.
  2. Normalization: $P(S) = 1$, where S is the sample space.
  3. Additivity of finite / countably infinite events.

$$P\left(\bigcup_{i=1}^{\infty/n} E_i\right) = \sum_{i=1}^{\infty/n} P\left(E_i\right),$$

where $\mathrm{E_i}$ are disjoint / mutually exclusive, i: natural number.

Use of Probability Axioms

- Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which is more probable?
    1. Linda is a bank teller.
    2. Linda is a bank teller and is active in the feminist movement.

Tversky & Kahneman.

Conditional Probability

- Sometimes, knowing something makes a difference.
- Model/reason about the outcome of an experiment, based on partial information.
- Given that event A occurs.
- The probability that event B also occurs:

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

- Extendable to knowing multiple events.

Chain Rule

- Probability that two events occur:

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

- In general,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(A_1 \middle| \bigcap_{j=2}^{n} A_j\right) P\left(\bigcap_{j=2}^{n} A_j\right)$$

$$= P\left(A_1 \middle| \bigcap_{j=2}^{n} A_j\right) P\left(A_2 \middle| \bigcap_{k=3}^{n} A_k\right) \ldots P\left(A_{n-1} \middle| A_n\right) P\left(A_n\right)$$

Bayes Rule

- Knowing P(AlB) maybe easier than knowing P(BlA), e.g.,
  - Knowing symptoms of a disease is easier than figuring out the disease given the symptoms.

posterior

prior

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Example

- A cab was involved in a hit & run accident at night.
- Only 2 cabs company operate in the city, the Blue & the Green.
- 85% of the cabs in the city are Green.
- The court tested the reliability of the witness under the same circumstances that existed on the night of the evidence & concluded that the witness correctly identify the color of the taxi 80% of the time.
- A witness identified the cab is Blue.
- Is Blue cab more likely to be the one involved in the accident ?

Tversky & Kahneman.

B: Blue cab was involved in the accident.

G: Green cab was involved in the accident.

W_B: The witness says blue is the one involved in the accident.

W_G: The witness says green is the one involved in the accident.

$$P(B) = 0.15 \quad ; \quad P(G) = 0.85$$

$$P(W_B|B) = P(W_G|G) = 0.8 \quad ; \quad P(W_B|G) = P(W_G|B) = 0.2$$

$$P(B|W_B) = \frac{P(W_B|B)P(B)}{P(W_B)}$$

$$= \frac{P(W_B|B)P(B)}{P(W_B \cap B) + P(W_B \cap G)}$$

$$= \frac{P(W_B|B)P(B)}{P(W_B|B)P(B) + P(W_B|G)P(G)}$$

The same derivation applies to $P(G|W_B)$. The answer to the question is blue if $P(B|W_B) > P(G|W_B)$ and vice versa.

Independence

- When knowing something doesn't make a difference.
- Knowing event A occurs does not change the probability that event B occurs:

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

- Chain rule becomes:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

Random Variables

- Interest is on numerical values associated w. samples, e.g.:
  - Sample 50 students enrolled in COMP3702/7702, the number of students from mechatronics, IT, s/w eng., bioinf.
  - Roll a fair dice, get \$5 if the outcome is even, & loose \$5 if the outcome is odd.
- Random variable X is a function $X : S \rightarrow Num$.
  - Num: countable set (e.g., integer) → discrete random variable.
  - Num: uncountable set (e.g., real) → continuous random variable.

Characterizing Random Variables

- Cumulative distribution function (cdf)

$$F_X(x) = P(X \le x) = P\left(\left\{s \middle| X(s) \le x, s \in S\right\}\right)$$

- Discrete: Probability mass function (pmf)

$$f_X[x] = P(X = x)$$

- Continuous: Probability density function/probability distribution function (pdf)

$$f_X(x) = \frac{dF_X(x)}{dx} \quad ; \quad P(a \le X \le b) = \int_a^b f_X(x)\,dx$$

More Compact Characterization of Random Variables

- Expectation: Weighted average of possible values of X, weight: probability.

$$E[X] = \sum_x x f_X[x] \quad ; \quad E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx$$

$$E[g(X)] = \sum_x g(x) f_X[x] \quad ; \quad E[g(X)] = \int_{-\infty}^{\infty} g(X) f_X(x)\,dx$$

$$\text{Linearity of Expectation}: E[aX + b] = aE[X] + b$$

More Compact Characterization of Random Variables

- Variance: A measure of dispersion around the mean.

$$\mathrm{var}(X) = E\left[\left(X - E[X]\right)^2\right] = E\left[X^2\right] - \left(E[X]\right)^2$$

$$\mathrm{var}\left(g(X)\right) = \sum_x g(X) f_X[x] \quad ; \quad g(X) = \left(X - E[X]\right)^2$$

$$\text{When g is linear}: \quad \mathrm{var}(aX + b) = a^2\,\mathrm{var}(X)$$

- Standard deviation: .

$$\sigma(X) = \sqrt{\mathrm{var}(X)}$$

For more on probability…

- Khan Academy (linked from http://robotics.itee.uq.edu.au/~ai/doku.php?id=ai:resources)
- STAT2202/STAT2203.
- ENGG7302.
- Introduction to Probability by Bertsekas & Tsitsiklis.