Discrete Mathematics

4. Discrete Probability

Why Probability?

- In the real world, we often don't know whether a given proposition is true or false.
- Probability theory gives us a way to reason about propositions whose truth is *uncertain*.
- Useful in weighing evidence, diagnosing problems, and analyzing situations whose exact details are unknown.

Random Variables

- *Definition:*
  - A *random variable V* is a variable whose value is unknown, or that depends on the situation.
- Let the domain of *V* be **dom**[V]={v_1,…,v_n}
- The proposition $V=v_i$ may be uncertain, and be assigned a *probability*.
- *Example:*
  - The number of students in class today
  - Whether it will rain tonight (Boolean variable)

Experiments

- *Definition:*
  - A (stochastic) *experiment* is a process by which a given random variable gets assigned a specific value.
  - The *sample space S* of the experiment is the domain of the random variable.
  - The *outcome* of the experiment is the specific value of the random variable that is selected.

Events

- *Definition:*
  - An *event E* is a set of possible outcomes:

$$E \subseteq S = \mathbf{dom}[V]$$

Probability

- *Definition:*
  - The *probability* $p = \Pr[E] \in [0,1]$ of an event E is a real number representing the degree of certainty that E will occur.

    $\Pr[E]=|E|/|S| = n(E)/n(S)$

    - If $\Pr[E] = 1$, then E is absolutely certain to occur.
    - If $\Pr[E] = 0$, then E is absolutely certain *not* to occur.
    - If $\Pr[E] = ½$, then we are *completely uncertain* about whether E will occur.
  - What about other cases?

Probability of Complementary Events

- Let E be an event in a sample space S.
- Then, $\overline{E}$ represents the *complementary* event of E.

$$\Pr[\overline{E}] = 1 - \Pr[E]$$

Probability of Unions of Events

- Let $E_1, E_2 \subseteq S = \mathbf{dom}[V]$.
- Then:
  $\Pr[E_1 \cup E_2] = \Pr[E_1] + \Pr[E_2] - \Pr[E_1 \cap E_2]$
  - By the inclusion-exclusion principle.

Mutually Exclusive Events

- *Definition:*
  - Two events E_1, E_2 are called *mutually exclusive* if they are disjoint: $E_1 \cap E_2 = \emptyset$
- Note that two mutually exclusive events *cannot both occur* in the same instance of a given experiment.
- For mutually exclusive events,

$$\Pr[E_1 \cup E_2] = \Pr[E_1] + \Pr[E_2].$$

Independent Events

- *Definition:*
  - Two events E,F are *independent* if $\Pr[E \cap F] = \Pr[E] \cdot \Pr[F]$.
- Relates to product rule for number of ways of doing two independent tasks
- Example: Flip a coin, and roll a die.

  Pr[quarter is head ∩ die is 1] =

  Pr[quarter is head] × Pr[die is 1]

Conditional Probability

- *Definition:*
  - Let E,F be events such that Pr[F]>0. Then, the *conditional probability of E given F*, written Pr[$E|F$], is defined as

    Pr[$E \cap F$]/Pr[F] (=n($E \cap F$)/n(F)).

- This is the probability that E would turn out to be true, given just the information that F is true.

- If E and F are independent, Pr[$E|F$] = Pr[E].

Bayes's Theorem

- Allows one to compute the probability that a hypothesis *H* is correct, given data *D*:

$$\Pr[H \mid D] = \frac{\Pr[D \mid H] \cdot \Pr[H]}{\Pr[D]}$$

- Easy to prove from definition of conditional probability
- Extremely useful in artificial intelligence apps:
  - Data mining, automated diagnosis, pattern recognition, statistical modeling, evaluating scientific hypotheses.

Exhaustive Sets of Events

- *Definition:*
  - A set $\boldsymbol{E} = \{E_1, E_2, \ldots\}$ of events in the sample space S is *exhaustive* if $\bigcup E_i = S$
  - An exhaustive set of events that are all mutually exclusive with each other has the property that

$$\sum \Pr[E_i] = 1$$

Probability Distribution

- *Definition:*
  - Let p be any function p:$S\rightarrow[0,1]$, such that:
    - $0 \leq p(s) \leq 1$ for all outcomes $s\in S$.
    - $\sum p(s) = 1$.
  - Such a p is called a *probability distribution*.
  - Then, the probability of any event $E\subseteq S$ is just:

$$\Pr[E] = \sum_{s\in E} p(s)$$

Expectation Value

- *Definition:*
  - For a random variable V having a numeric domain, its *expectation value* or *expected value* $\mathbf{E}[V]$ is defined as $\sum_{v\in \mathbf{dom}[V]} v\cdot p(v)$.

Expectation Value (cont.)

- *Theorem:*
  - Let X_1, X_2 be any two random variables derived from the same sample space. Then,

$$\mathbf{E}[X_1+X_2] = \mathbf{E}[X_1] + \mathbf{E}[X_2]$$

$$\mathbf{E}[aX_1 + b] = a\mathbf{E}[X_1] + b$$

Independent Random Variables

- *Definition:*
  - Two random variables X and Y are **independent** if $p(X=r_1 \text{ and } Y=r_2) = p(X=r_1) \cdot p(Y=r_2)$ for every real numbers, r_1 and r_2
- *Theorem:*
  - If X and Y are independent random variables, then $E(XY) = E(X) \cdot E(Y)$

Variance

- *Definition:*
  - The *variance* **Var**[X] = $\sigma^2(X)$ of a random variable X is the expected value of the square of the difference between the value of X and its expectation value **E**[X]:

$$\mathbf{Var}[X] :\equiv E[(X - E[X])^2]$$

  - The *standard deviation* or *root-mean-square* (RMS) *difference* of X, $\sigma(X) :\equiv \mathbf{Var}[X]^{1/2}$.

Binomial Distribution

- *Theorem:*
  - The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure q=1-p, is

$$\frac{n!}{k!(n-k)!}p^k q^{n-k}$$

Binomial Distribution (cont.)

- *Definition:*
  - Binomial distribution is the discrete probability distribution of the number of successes (X) in a sequence of n independent experiments, each of which yields success with probability p.

$$p(X=k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

Binomial Distribution (cont.)

- *Theorem:*
  - Let X be a random variable with binomial distribution, then ***E[X]=np*** and ***Var[X]=np(1-p)***
  - Let $X_1, X_2, \ldots, X_n$ be pairwise independent random variables, then

    $$Var[X_1+X_2+\ldots+X_n] = Var[X_1]+Var[X_2]+\ldots+Var[X_n]$$

Excercise

1. Let A, B and C be events in a sample space and suppose $P(A \cap B) \neq 0$. Prove that $P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$

2. Let A and B be events with nonzero probability in a sample space.
   (a) Suppose $P(A|B) > P(A)$. Must it be the case that $P(B|A) > P(B)$?
   (b) Suppose $P(A|B) < P(A)$. Must it be the case that $P(B|A) < P(B)$?

Excercise

3. Let X and Y be two independent random variables.
   (a) Give the definition of variance, Var(X), of X and show that Var(X)= E(X^2)-E(X)^2.
   (b) Show that Var(X+Y)=Var(X)+Var(Y).