Prob. & Stats. - Week 1 Outline

Yi (Alice) Yin

September 15, 2017

1 The Ingredients of a probabilistic setup

There is an experiment whose outcomes are subject to uncertainty. A probabilistic description involves three ingredients.

1.1 First ingredient: Sample Space

- Sample Space
  - Definition:
    The sample space (or universe) is the set formed by all possible outcomes of the experiment.
  - Notation: Ω
- Element: $\omega \in \Omega$
- Two Cases:
  - Ω is Finite
  - Ω is Infinite

1.2 Second ingredient: Events

- Events:
  - An event A of a given sample space Ω is a subset of Ω.
  - Two Cases:
    - Ω is Finite: Events are all subsets of Ω.
    - Ω is Infinite: Events are all subsets of Ω that you will ever find.
  - Two ways of expressing:
    - $A = \{\text{list of element}\}$
    - $A = \{\text{property that must be satisfied}\}$
  - Occur:
    "An event A occurs": the outcome of the experiment is an element in A.

- Visualization: Venn Diagrams
- Set Relations & Operations
  - Union
    - $A \cup B = \{\omega \in \Omega : \omega \in A \text{ or } \omega \in B\}$
    - Generalization: $A_1 \cup A_2 \cup ... \cup A_n \cup ...$
  - Intersection
    - $A \cap B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}$
    - Generalization: $A_1 \cap A_2 \cap ... \cap A_n \cap ...$
  - Complement
    - $A^c = \{\omega \in \Omega : \omega \notin A\}$
- De Morgan's Laws:
  - $(A \cap B)^c = A^c \cup B^c$
  - $(A \cup B)^c = A^c \cap B^c$
  - Generalizations: $\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c$; $\left(\bigcap_i A_i\right)^c = \bigcup_i A_i^c$
- Disjoint / Mutually Exclusive
  - Empty Set: $\emptyset$
  - A and B are disjoint if $A \cap B = \emptyset$
  - Pairwise Disjoint:
    $A_1, A_2, ..., A_n, ...$ are pairwise disjoint if $A_i \cap Aj = \emptyset$ for all $i \neq j$.
- Useful identities:
  - $A \cap \Omega = A$, $A \cup \Omega = \Omega$, $A \cap \emptyset = \emptyset$, $A \cup \emptyset = A$
  - $(A \subset B) \Rightarrow (A \cup B = B,\ A \cap B = A)$.
  - $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.
  - In general, for algebraic purposes $\cup \sim +$, $\cap \sim \times$

Prob. & Stats. - Week 2 Outline

Yi (Alice) Yin

September 15, 2017

1 Probability

It is the third ingredient of a probabilistic description of an experiment.

1.1 Probability

- Definition:
  Given a sample space Ω and a set of events, a probability function / distribution / measure is a function P that maps the set of events to the interval $[0, 1]$ satisfying:
  - $P(\Omega) = 1$
  - If $A_1, A_2, \ldots, A_n, \ldots$ are pairwise disjoint, then $P\big(\bigcup_i A_i\big) = \sum_i P(A_i)$.
- Nomenclature
  - Impossible events: $P(A) = 0$.
  - Certain events: $P(A) = 1$.
- Properties
  - Monotonicity:
    $$A \subset B \Rightarrow P(A) \leq P(B)$$
  - Probability of complements:
    $$P(A^c) = 1 - P(A)$$
  - Probability of unions (not necessarily disjoint):
    $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
  - Dividing into Cases:
    If $B_1, B_2, \ldots, B_n, \ldots$ form a partition of Ω [that is, the B_i's are pairwise disjoint and their union is the whole of Ω], then
    $$P(A) = \sum_i P(A \cap B_i)$$

– Probability of intersections:
  If $A_1, A_2, \ldots, A_n, \ldots$ is a family of independent events

$$P(\bigcap_i A_i) = \prod_i P(A_i)$$

1.2 Combinatorics (Counting)

Combinatorics is useful for computing probabilities when all outcomes have the same probability.

- Product space:

$$\Omega_1 \times \Omega_2 \times ... \times \Omega_n \times ... \ = \ \left\{(\omega_1, \omega_2, ..., \omega_n, ...) : \omega_i \in \Omega_i\right\}$$

- Multiplication Principle:

  The number of elements of a product space is the product of the number of elements of the different components,

$$N(\Omega_1 \times \Omega_2 \times ... \times \Omega_n \times ...) = \prod_i N(\Omega_i)$$

- Permutations
  - Definition:
    A permutation of a set is an ordered arrangement of its elements.
  - N(permutations of a set with m elements) $= m!$

Prob. & Stats. - Week 3 Outline

Yi (Alice) Yin

September 30, 2017

1 Permutations (continued)

- The number of Permutations of k objects chosen out of n objects is

$$\frac{n!}{(n-k)!}$$

- The number of Permutations of n objects of which

  k_1 are identical objects of the 1st kind
  k_2 are identical objects of the 2nd kind
  $\vdots$
  k_m are identical objects of the m-th kind:

  is

$$\frac{n!}{\prod_{i=1}^{m}(k_i)!}$$

2 Combinations

- Definition:
  The Combination of n objects taken k at a time is a selection of k objects without regards for the order of selection.

- The number of combinations of k objects chosen out of n objects is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- By symmetry:

$$\binom{n}{k} = \binom{n}{n-k}$$

- Binomial Expansion:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

- Application:
$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

3 Conditional Probability

- Definition:
The conditional probability of A given B with $P(B) > 0$ is
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
    - Note: It corresponds to hanging the sample space/the universe Ω to B.
- Product Formula:
$$P(A \cap B) = P(A|B)P(B)$$
- Application:
    - To compute probability.
    - To see the effects of incorporating information.
(e.g. Prediction of the stocks)
    - Usually information counts through conditional probability.
(e.g. Statements with "if")
- Behavior of $P(A|B)$ with respect to the 1st slot:
    - $P(\cdot|B)$ is a probability.
    - Consequences:
        1. $P(A^c|B) = 1 - P(A|B)$
        2. $P(A|B) = P(A \cap C|B) + P(A \cap C^c|B)$
        3. $P(A \cup C|B) = P(A|B) + P(C|B) - P(A \cap C|B)$

- Behavior of $P(A|B)$ with respect to the 2nd slot: $P(A|\cdot)$ can be used to *divide & conquer*:
    - If $P(B), P(B^c) > 0$,
$$P(A) \;=\; P(A \mid B)\, P(B) + P(A \mid B^c)\, P(B^c)$$
    - Generalization: If $B_1, B_2, \ldots,$ form a partition of Ω,
$$P(A) \;=\; \sum_i P(A \mid B_i)\, P(B_i)\;.$$

- Bayes' Formula

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

Generalization: If $B_1, B_2, \ldots,$ form a partition of Ω,

$$P(B_i \mid A) \; = \; \frac{P(A \mid B_i)P(B_i)}{\sum_i P(A|B_i)P(B_i)}$$

- Successive Conditioning
  - $P(A \cap B_1 \cap B_2) \; = \; P(A|B_1 \cap B_2)\, P(B_2|B_1)\, P(B_1)$
  - Generalization:

    $P(A \cap B_n \cap B_{n-1} \cap \cdots \cap B_1) \; = \; P(A \mid \cap_{i=1}^n B_i)\, P(B_n | \cap_{i=1}^n B_i) \; \cdots \; P(B_2|B_1)P(B_1)$
  - Particular case: if B_i is decreasing, i.e. $B_n \subset B_{n-1} \subset ... \subset B_1$

$$P(A \cap B_n) = P(A|B_n) \prod_{i=2}^{n-1} P(B_i|B_{i-1})\, P(B_1)$$

4 Independence

- Definition:
  A and B are independent if $P(A \cap B) = P(A)P(B)$.
- If $P(B) > 0$, then A is independent of B if, and only if,

$$P(A|B) = P(A)$$

Prob. & Stats. - Week 4 Outline

Yi (Alice) Yin

September 30, 2017

1 Independent Events (continue)

- Definition:
  A and B are independent if $P(A \cap B) = P(A)P(B)$.
- If $P(B) > 0$, then A is independent of B, if and only if,

$$P(A|B) = P(A)$$

- Properties
  - Symmetry: A is independent of $B \Rightarrow B$ is independent of A.
  - If $B \cap C = \emptyset$, then
    (A is independent of B) and (A is independent of C)
    $\Rightarrow$ A is independent of $B \cup C$
- General Independence:
  We say that $A_1, A_2, \ldots, A_n$ are independent if for all $k \in \{2, 3, \ldots, n\}$

$$P(\cap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j})$$

for all different $i_1, i_2, \ldots, i_k$.

2 Random Variables

- Definition:
  Given Ω, the set of all possible events, and a probability P, a **random variable** is a function:

$$X : \Omega \to \mathbb{R}$$

- Comments:

- If Ω is finite,and all subsets of Ω are events, all functions are random variables. In the general case there is some further condition that will not be discussed in detail in this course. Rather, we shall consider two very popular types of random variables for which this additional condition is satisfied, the discrete and continuous random variables.
- In the sequel, when we say "X is a random variable", it is understood that there is some Ω, set of events and P even if we do not write them explicitly.

2.1 Discrete Ramdom Variables

- Definition:
  A random variable X is a **discrete random variable** if it can take only finitely many values $\{a_1, a_2, \ldots, a_n\}$ or infinitely but countably many (that is, the possible values can be labelled by the integers as $a_1, a_2, \ldots, a_k, \ldots$).

- Comments:
  - The set of values that X can take are called the *results* of X or, more formally, the **image** of X. Hence, a random variable X is discrete if

$$\text{Image of } X \;=\; \big\{a_1, a_2, \ldots, a_k, \ldots\big\} \;.$$

  - Note that the requirement is that the *image* be discrete. The sample space Ω need not be.

- Probability Mass/Weight Function:
  The **probability mass function** ($p.m.f$) of a discrete random variable X with possible results $\{a_1, a_2, \ldots, a_n\}$ is the function:

$$p : \mathbb{R} \to [0, 1]$$

$$p_X(a) \;:=\; P(\{X = a\}) \;=\; P(\{\omega : X(\omega) = a\})$$

  To simplify the notation we will denote $P(X = a)$ instead of the more precise, but harder to read, $P(\{X = a\})$.

- Cumulative Distribution Function:
  The **cumulative distribution function** ($c.d.f$) of a random variable X is the function:

$$F_X : \Omega \to [0, 1]$$

$$F_X(a) = P(X \leq a)$$

- Relation c.d.f. $\leftrightarrow$ p.m.f.

  F_X can be obtained from p_X and viceversa. Indeed, if the values of the image of X are ordered so that $a_{i-1} \leq a_i$, then

$$F_X(a) \;=\; \sum_{i : a_i \leq a} p_X(a_i)$$

for any $a \in \mathbb{R}$, and

$$p_X(a_i) \;=\; F_X(a_i) - F_X(a_{i-1})$$

for any i.

- Properties of $c.d.f$:

(P1) $F(a) = 0$ if $a < a_{min}$. More generally,

$$F(a) \xrightarrow[a \to -\infty]{} 0$$

(P2) $F(a) = 1$ if $a \geq a_{max}$. More generally,

$$F(a) \xrightarrow[a \to \infty]{} 1$$

(P3) F is an increasing function:

$$a \leq b \implies F(a) \leq F(b)$$

(P4) F is right-continuous.

$$F(X + \epsilon) \xrightarrow[\epsilon \downarrow 0]{} F(X)$$

(P5) If F has a jump at a, its size equals $p_X(a)$.

- Laws:
  - Two discrete random variables X, Y have the same **law** if:

$$p_X = p_Y \quad \text{or, equivalently,} \quad F_X = F_Y$$

  - Comments:
    * The laws contain all the information needed to compute probabilities involving random variables.
    * If the law is known, there is no need to know the initial Ω or P used to define the random variable.

- Catalogue of important discrete Laws

Law	**Image of X**	p_X
Ber(p)	$\{0, 1\}$	$p_X(1) = p$
Uni(N)	$\{1, 2, \dots, N\}$	$p_X(i) = \frac{1}{N}$
Bin(n, p)	$\{0, 1, \dots, n\}$	$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$
Geo(r)	$\mathbb{N}^*$	$p_X(r) = (1-p)^{r-1} p$
Poi(λ)	$\mathbb{N}$	$p_X(n) = e^{-\lambda} \frac{\lambda^n}{n!}$