A review of Chapter 2 and some of Chapter 3

STA 320, Fall 2008

Ruriko Yoshida
Dept. of Statistics University of Kentucky

www.ms.uky.edu/~ruriko

Basic on Set Theory

Let $A, B \subset S$. Then,

$$A \cup B = \{x | x \in A \text{ or } x \in B\}.$$
$$A \cap B = \{x | x \in A \text{ and } x \in B\}.$$
$$A - B = \{x | x \in A \text{ and } x \notin B\}.$$
$$A \subset B \text{ means } x \in A \Rightarrow x \in B.$$
$$A = B \text{ if and only if } A \subset B \text{ and } B \subset A.$$

Let $A \subset S$. Then,

$$(A^c)^c = A.$$
$$\emptyset^c = S.$$
$$S^c = \emptyset.$$
$$A \cup A^c = S.$$
$$A \cap A^c = \emptyset.$$

Let $A, B, C \subset S$. Then,

$$A \cup B = B \cup A.$$
$$A \cup (B \cup C) = (A \cup B) \cup C.$$
$$A \cap B = B \cap A.$$
$$A \cap (B \cap C) = (A \cap B) \cap C.$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$
$$(A \cap B) = A^c \cup B^c$$
$$(A \cup B) = A^c \cap B^c$$

Definition of Probability

Definition Suppose $A_1, A_2, \cdots \subset S$ are infinite sequence of events. Then we say $A_1, A_2, \cdots$ are disjoint iff

$$A_i A_j = \emptyset, \forall i, j \text{ with } i \neq j.$$

Definition A probability P is a function from the set of all possible events in S to $\mathbb{R}$ such that

$$P(A) \geq 0 \ \forall A \subset S,$$

$$\text{if } A_1, A_2, \cdots \subset S \text{ are disjoint, } P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i),$$

$$P(S) = 1.$$

Some Theorems

Thm

$$P(\emptyset) = 0.$$

Thm Suppose $A_1, A_2, \cdots A_n \subset S$ are finite sequence of disjoint events. Then,

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

Thm $\forall A \subset S$,

$$P(A^c) = 1 - P(A) \text{ and } 0 \leq P(A) \leq 1.$$

Thm $\forall A, B \subset S$ such that $A \subset B$,

$$P(A) \leq P(B).$$

Combinatorial Methods

Definition A permutation of order n, S_n, is an arrangement or ordering of n objects.

Definition An r permutation of order n, S_n^r, is an arrangement using r out of n objects.

Definition An r combination of n distinct objects is an unordered selection or subset of r out of n objects.

We write

$$P_{n,r} = \# \text{ of } S_n^r = \frac{n!}{(n-r)!},$$

$$C_{n,r} = \# \text{ of } r \text{ combinations of } n \text{ distinct objects } = \frac{n!}{r!(n-r)!} = \binom{n}{r}.$$

Binomial Coefficients

Definition $C_{n,r}$ are called binomial coefficients.

Thm (Binomial Theorem) $C_{n,i}$ are coefficients of x^i in the polynomial $(1+x)^n$. In other words,

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n-1}x^{n-1} + \binom{n}{n}x^n.$$

Note $C_{n,0} = C_{n,n} = 1$.

Binomial Identities

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

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1},$$

$$\sum_{i=0}^{n}\binom{n}{i} = 2^n,$$

$$\sum_{i=0}^{r}\binom{n+i}{i} = \binom{n+r+1}{r},$$

$$\sum_{i=0}^{n}\binom{n}{i}^2 = \binom{2n}{n},$$

Binomial Identities cont....

$$\sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r},$$

$$\sum_{k=0}^{r} \binom{m}{k}\binom{n}{r+k} = \binom{m+n}{m+r},$$

$$\sum_{k=s-n}^{m-r} \binom{m-k}{r}\binom{n+k}{s} = \binom{m+n+1}{r+s+1}.$$

Multinomial Coefficients

Definition A multinomial coefficient is defined by $\frac{n!}{n_1!n_2!\cdots n_k!}$ where $n_1 + n_2 + \cdots + n_k = n$ and $n_i \geq 0$ integer for all $i = 1, 2, \cdots k$. It is denoted by

$$\binom{n}{n_1, n_2, \cdots, n_k}.$$

Thm (Multinomial Theorem)

For all numbers $x_1, x_2, \cdots, x_k$ and each positive integer n, we have

$$(x_1 + x_2 + \cdots + x_k)^n = \sum \binom{n}{n_1, n_2, \cdots, n_k} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}$$

where the summand extends over all possible combinations of nonnegative integers $n_1, n_2, \cdots, n_k$ such that $n_1 + n_2 + \cdots + n_k = n$.

Probability of a union of events

Thm Suppose $A_1, A_2, \cdots, A_n \subset S$ are finite sequence of events. Then

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_iA_j) + \sum_{i<j<k} P(A_iA_jA_k)$$

$$-\sum_{i<j<k<l} P(A_iA_jA_kA_l) + \cdots (-1)^{n+1} P(A_1A_2\cdots A_n).$$

Conditional Probability

Definition Suppose $A, B \subset S$. The conditional probability of A given B, $P(A|B)$, is a proability that A occurs after B occurs.

Note If $A, B \subset S$ such that $P(B) > 0$, then

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

Note (Multiplication Rule)

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

Independent Events

Definition $A, B \subset S$ are independent iff $P(A)P(B) = P(AB)$.

Thm If $A, B \subset S$ are independent, then A, B^c are independent.

Definition $A_1, A_2, \cdots A_n \subset S$ are independent iff for every subsets $A_{i_1}, A_{i_2}, \cdots, A_{i_j}$ of j of these events,

$$P(A_{i_1}A_{i_2}\cdots A_{i_j}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_j}).$$

Definition $A_1, A_2, \cdots A_n \subset S$ are pairwise independent iff for every i, j with $i \neq j$

$$P(A_iA_j) = P(A_i)P(A_j).$$

Independent Events and Conditional Prob

Note $A, B \subset S$ are independent iff $P(A|B) = P(A)$.

Definition Let $A_1, A_2, \cdots, A_n, B \subset S$. We say Let $A_1, A_2, \cdots A_n$ are conditionally independent given B iff for every subset $A_{i_1}, A_{i_2}, \cdots, A_{i_j}$ of j of these events,

$$P(A_{i_1} A_{i_2} \cdots A_{i_j} | B) = P(A_{i_1}|B) P(A_{i_2}|B) \cdots P(A_{i_j}|B).$$

Thm Suppose that $A_1, A_2, B \subset S$ such that $P(A_1 B) > 0$. Then A_1, A_2 are conditionally independent given B iff $P(A_2|A_1 B) = P(A_2|B)$.

Law of Total Probability

Definition A collection $\{B_i\}_{i=1}^{\infty}$ of disjoint events for which $\cup_{i=1}^{\infty} B_i = S$ is called a partition of the sample space S.

Thm (Law of Total Probability)

For any partition of S, $\{B_i\}_{i=1}^{\infty}$, for any event $A \subset S$, we have

$$P(A) = \sum_{i=1}^{\infty} P(AB_i) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i).$$

Law of Total Probability

Thm (Conditional version of Law of Total Probability)

For any partition of S, $\{B_i\}_{i=1}^{\infty}$, for any event $A, C \subset S$, we have

$$P(A|C) = \sum_{i=1}^{\infty} P(AB_i|C) = \sum_{i=1}^{\infty} P(A|B_iC)P(B_i|C).$$

Bayes' Theorem

Thm (Bayes' Theorem)

Suppose $\{B_i\}_{i=1}^{\infty}$ is a partition of S and $A \subset S$ for which $P(A) > 0$. Then, for any event B_i with $P(B_i) > 0$, we have:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_j P(B_j)P(A|B_j)}.$$

Definition $P(B_i)$ in the equation above is called a prior probability and $P(B_i|A)$ in the equation above is called a posterior probability.

Discrete random variables

Definition A random variable (r.v.) X is a function from S to $\mathbb{R}$. A discrete random variable X is a function from S to $\mathbb{Z}$.

Definition A probability function (p.f.) $f(x)$ of a discrete r.v. X is a function defined over $\mathbb{R}$ such that

$$f(x) = P(X = x) \text{ where } x \in \mathbb{Z}.$$

Note Since $P(S) = 1$, we have

$$\sum_{x=-\infty}^{\infty} f(x) = 1.$$

Discrete distributions

Definition A bionomial distribution is a distribution of a discrete r.v. X with a p.f. $f(x)$ with given p and n such that

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Definition A hypergeometric distribution is a distribution of a discrete r.v. X with a p.f. $f(x)$ with given N, m and n such that

$$f(x) = P(X = x) = \binom{m}{x}\binom{N-m}{n-x} \Big/ \binom{N}{n}.$$