F71SB2 Statistics II Handout 1

1 Definition of Probability [Statistics I review]

1.1 Experiment

An experiment is any procedure whose outcome is uncertain

- toss a coin
- throw a die
- buy a lottery ticket
- measure an individual's height h

1.2 Sample Space

A sample space, S, is the set of all possible outcomes of a random experiment

- $S = \{H, T\}$
- $S = \{1, 2, 3, 4, 5, 6\}$
- $S = \{\mathit{WIN},\ \mathit{LOSE}\}$
- $S = \{h : h \geq 0\}$

1.3 Events

An event, A, is an element, or appropriate subset of S
e.g. roll a die, $S = \{1, 2, 3, 4, 5, 6\}$

- event A, that an even number is obtained is given by $A = \{2, 4, 6\}$
- event B, that a number no greater than 4 is obtained is described by $B = \{1, 2, 3, 4\}$

S can be <u>discrete</u> (e.g. coin, die, lottery), or <u>continuous</u> (e.g. height).

Set operations
If A, B are events, according to Set Theory, we define the operations:

- $A \cup B$: union, 'A <u>or</u> B happening'
- $A \cap B$: intersection, 'A <u>and</u> B happening'
- A^c (or $A', \bar{A}$): complement of A wrt S, '<u>not</u> A'
  Notice that: $A \cup A^c = S$, $A \cap A^c = \emptyset$

1.4 Probability function

A probability function, $P(s)$, is a real-valued function of a collection of events from a sample space, S, attaching a value in $[0,1]$ to each event, satisfying the axioms:

A1. $P(A) \geq 0$ for any event A

A2. $P(S) = 1$

A3. If $A_1, A_2, A_3, \ldots$ is a countable collection of mutually exclusive events ($A_i \cap A_j = \emptyset$ for $i \neq j$, i.e. they have no common elements), then

$$P(\cup_i A_i) = \sum_i P(A_i)$$

Important properties of $P(s)$

- $P(\emptyset) = 0$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  This can be generalised for n events.
  Notice: $P(A \cup B) = P(A) + P(B)$ not always true!
  It only holds for disjoint (mutually exclusive) events.
- $P(A^c) = 1 - P(A)$
  In many practical situations it is easier to calculate $P(A^c)$ first and use this property to calculate $P(A)$.

1.5 Independence and Conditional Probability

For any event A with $P(A) > 0$ we can define the conditional probability

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \tag{1}$$

If A and B are independent then by definition $P(A \cap B) = P(A) \times P(B)$ and therefore

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

Note that rearrangement of (1) gives the chain (multiplication) rule:

$$P(A \cap B) = P(A)P(B/A) = P(B)P(A/B). \tag{2}$$

This can also be generalised for n events.

1.6 Total Probability Rule

If $A_1, A_2, \ldots, A_k$ are mutually exclusive events and form a partition of a sample space S (i.e. $\cup_{i=1}^{k} A_i = S$), with $P(A_i) > 0 \qquad \forall i$,

then for an event $B \in S$:

$$P(B) = \sum_{i=1}^{k} P(B \cap A_i) = \sum_{i=1}^{k} P(A_i)P(B/A_i) \tag{3}$$

Example:
An insurance company covers claims from 4 different motor portfolios, A_1, A_2, A_3 and A_4. Portfolio A_1 covers 4000 policy holders; A_2 covers 7000; A_3 covers 13000; and A_4 covers 6000 of the total 30000 policy holders insured by the company. It is estimated that the proportions of policies that will result in a claim in the following year in each of the portfolios are $8\%, 5\%, 2\%$ and 4% respectively.
What is the probability that a policy chosen randomly from one of the portfolios will result in a claim in the following year?

Solution ...

1.7 Bayes Theorem

If $P(A) > 0$ and $P(B) > 0$, from (1)

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B/A)}{P(B)} \tag{4}$$

or, if $A_1, A_2, \ldots, A_k$ form a partition of a sample space S with $P(A_i) > 0, \forall i$, and $P(B) > 0$, then from (3)

$$P(A_j/B) = \frac{P(A_j)P(B/A_j)}{\sum_{i=1}^{k} P(A_i)P(B/A_i)}, \qquad j = 1, 2, \ldots, k. \tag{5}$$

Example: *(previous cont.)*
If a claim rises from a policy in the year concerned, what is the probability that the policy belongs to portfolio A_3?

Solution ...

1.8 Random variables

A random variable (r.v.) X is a function from a sample space S to $\mathbb{R}$ (the real numbers). That is, to any outcome, s, we associate a real number $X(s)$.

The range of a r.v. X, denoted as S_x, is the set $\{r \in \mathbb{R} : r = X(s) \text{ for some } s \in S\}$, i.e. the set of all real numbers r which are equal to $X(s)$ for some outcome s.
We will usually denote random variables using capital letters, e.g. X, and their realisations (values) using small letters, e.g. x.

If S is finite or countable then the range of X is a set of discrete numbers. We say X is a discrete random variable.

Some simple examples:

(i) Experiment: Toss a balanced coin 3 times.
Sample space S: All possible sequences of T, H of length 3.
Let $X(s) =$ no. of Heads in outcome (sequence) s. Then X is a discrete random variable.

(ii) Experiment: Toss a balanced coin until a 'H' is thrown.
Sample space: $S = \{H, TH, TTH, \ldots\}$.
Let $X(s) =$ no. of throws (i.e. length of s). Then X is a discrete random variable.

1.9 Probability function of a discrete r.v.

Any probability function on S leads naturally to a probability function on the range of X, defined by

$$f_X(x) = P(X = x) = \sum_{s:X(s)=x} P(s).$$

Axioms of probability
For a probability function of a discrete r.v. to satisfy the axioms of probability, it is sufficient that

i. $f_X(x) \geq 0$ for all $x \in S_x$

ii. $\sum_{\text{all } x} f_X(x) = 1$