Preliminary Statistics

Lecture 2: Probability Theory

Sophie van Huellen (sv8@soas.ac.uk), September 2013

Outline

- Probability Theory
  - Definition
  - Terminology
  - Rules and Axioms
- Random Variables
  - Discret Random Variables
  - Continuous Random Variables
  - Distributions

What are Probabilities?

- Rolling a 6 with a fair dice?
- Selecting a heart card from a pack of 52 cards?
- Arsenal winning the Premiership this year?

- Classical definition of probability
  - Also known as *a priori*
  - It requires knowledge of all possible (and finite) outcomes of the unknown event which should also be mutually exclusive and equally likely.
  - It defines probabilities as:

$$P(A) = \frac{Number\ of\ Outcomes\ in\ favour\ of\ A}{Total\ number\ of\ Outcomes}$$

Approach to Probabilities

- Frequentist View
  - Probability = Proportion of times the event occurs in the Long Run (after "infinite" repetitions). I.e. Relative Frequency.
  - Collect data in many repetitions and use the proportion of one event happening relative to the number of repetitions as the probability.
- Subjective View
  - Foundation of Bayesian Statistics
  - If the event has not happened or happened only once, its likelihood of occurring can only be evaluated on the basis of individual subjectivity, i.e. Degrees of Belief .
  - Degree of Belief updated in the light of new information.

Terminology

- **An experiment or trial**:
  - A process where the outcome is not known until the process is completed. *E.g. rolling a die*
- **An outcome:**
  - The result of an experiment. *E.g. one of the numbers 1, 2, 3, 4, 5, 6*
- **An event:**
  - A set of outcomes defines an event. *E.g. rolling an odd number is the set of 3 outcomes {1,3,5}*
- **Sample space:**
  - The set of all possible outcomes, sometimes referenced as the Greek letter omega Ω. *E.g. for a dice the full set of possible outcomes is {1,2,3,4,5,6}*

Terminology (cont.)

- **Mutually exclusive:**
  - Two events A and B are *mutually exclusive* if the occurrence of the one prevents the occurrence of the other at the same time. (Two sets are mutually exclusive if they do not share any common elements). *E.g. A=Rain, B=Not Rain*
- **Collectively Exhaustive:**
  - Two events are *collectively exhaustive* if they exhaust all possible outcomes of an experiment. *E.g. MSc class: A = {Males}, B = {Females}*
- **Equally Likely:**
  - Two events are *equally likely* if we are confident that one event is as likely to occur as the other. *E.g. Flip a coin*
- **Independent Events:**
  - Two events are *independent* if the occurrence of the one does not depend on the other. *E.g.: Having Heads and Having Tails when tossing 2 coins*

Rules/Axioms

- General:
  - Probabilities lie between zero and one: $0 \leq P(A) \leq 1$.
  - Outcomes outside of the sample space do not occur, have a frequency of 0, and so have a probability of 0.
  - If event A will occur with certainty, then $P(A)=1$.
  - The sum of the probabilities of all outcomes in the sample space is 1.
  - The probability of A not happening is, $1-P(A)$. This is the complement of A.

Rules/Axioms (cont.)

- Intersection
  - Probability of both events happening, A and B, is denoted as $P(A \cap B)$.
  - If A,B are mutually exclusive: $P(A \cap B) = 0$.
  - If A,B are independent: $P(A \cap B) = P(A)*P(B)$.
- Union
  - Probability of event A or event B happening is $P(A \cup B)$. *E.g. Probability of getting a 3 or a 4 when throwing a die.*

Rules/Axioms (cont.)

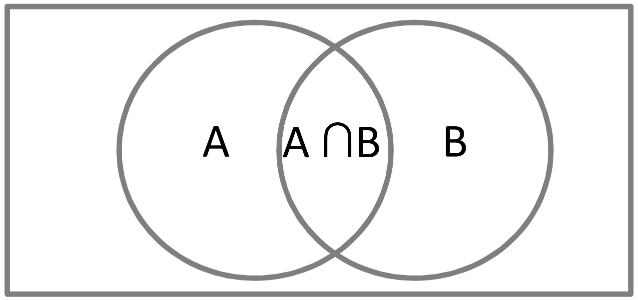
- The Addition Rule

$$\mathbf{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

  - i.e. the probability of A or B happening is the sum of their individual (marginal) probabilities minus the probability of their intersection
  - Where A and B are mutually exclusive, $P(A \cap B) = 0$ and the addition rule reduces to $P(A \cup B) = P(A) + P(B)$

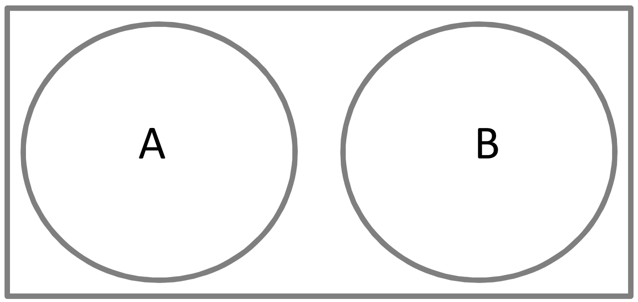
Rules/Axioms (cont.)

- Suppose the two events are not mutually exclusive



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Suppose the two events are mutually exclusive



$$P(A \cup B) = P(A)+P(B)$$

Rules/Axioms (cont.)
Conditional Probability

- The conditional probability of A given B is written $P(A|B)$:

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

- When A and B are independent, B occurring will not effect the probability of A occurring. Hence, $P(A|B) = P(A)$

Rules/Axioms (cont.)
Multiplication Rule

- The conditional probability rule provides a method of obtaining the joint probability of two events, i.e. the intersection, when the events are not independent:

$$\mathbf{P(A \cap B) = P(A|B)P(B)}$$

Rules/Axioms (cont.)

Bayes Theorem

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

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

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

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

Rules/Axioms (cont.)

Probability Rules

- **Additional Rule**:
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - If A, B are mutually exclusive: $P(A \cup B) = P(A) + P(B)$
  -
- **Conditional Probability**:
  - $P(A|B) = P(A \cap B)/P(B)$
  - If A and B are independent: $P(A|B) = P(A)$
- **Multiplication Rule**:
  - $(A \cap B) = P(A|B)P(B)$
  - If A and B are independent: $P(A \cap B) = P(A)P(B)$
- **Very Useful Rule**:
  - We can always split a marginal probability as such:
  - $P(A) = P(A \cap B) + P(A \cap notB) = P(A|B)P(B) + P(A|notB)*P(notB)$

Rules/Axioms (cont.)

Recap

Probability Theory	
Are A and B mutually exclusive?	
Yes	No
$P(A \cap B) = 0$	$P(A \cap B) > 0$
$P(A \cup B) = P(A) + P(B)$	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Are A and B independent?	
Yes	No
$P(A\|B) = P(A)$	$P(A\|B) = P(A \cap B) / P(B)$
$P(A \cap B) = P(A)P(B)$	$P(A \cap B) = P(A\|B)P(B)$

Random Variables

- A random variable is a variable whose outcome or value is the result of chance, and hence unpredictable. Range of possible values and their probabilities might be known.

1. Discrete Random Variables
2. Continuous Random Variables

Random Variables (cont.)

Discrete Random Variables

- X takes a number of distinct possible values, x_i with probabilities p_i. *E.g. X=Total obtained by throwing 2 dice, X=2, …12.*

- Probability Distribution:

$$p_i = f(x_i)$$

- Gives the probability of obtaining each of the possible values:

$$f(x_i) = P(X = x_i)$$

- Cumulative prob distribution:

$$F(x_i) = \sum_{i=1}^{j} f(x_i) = P(X \le x_i)$$

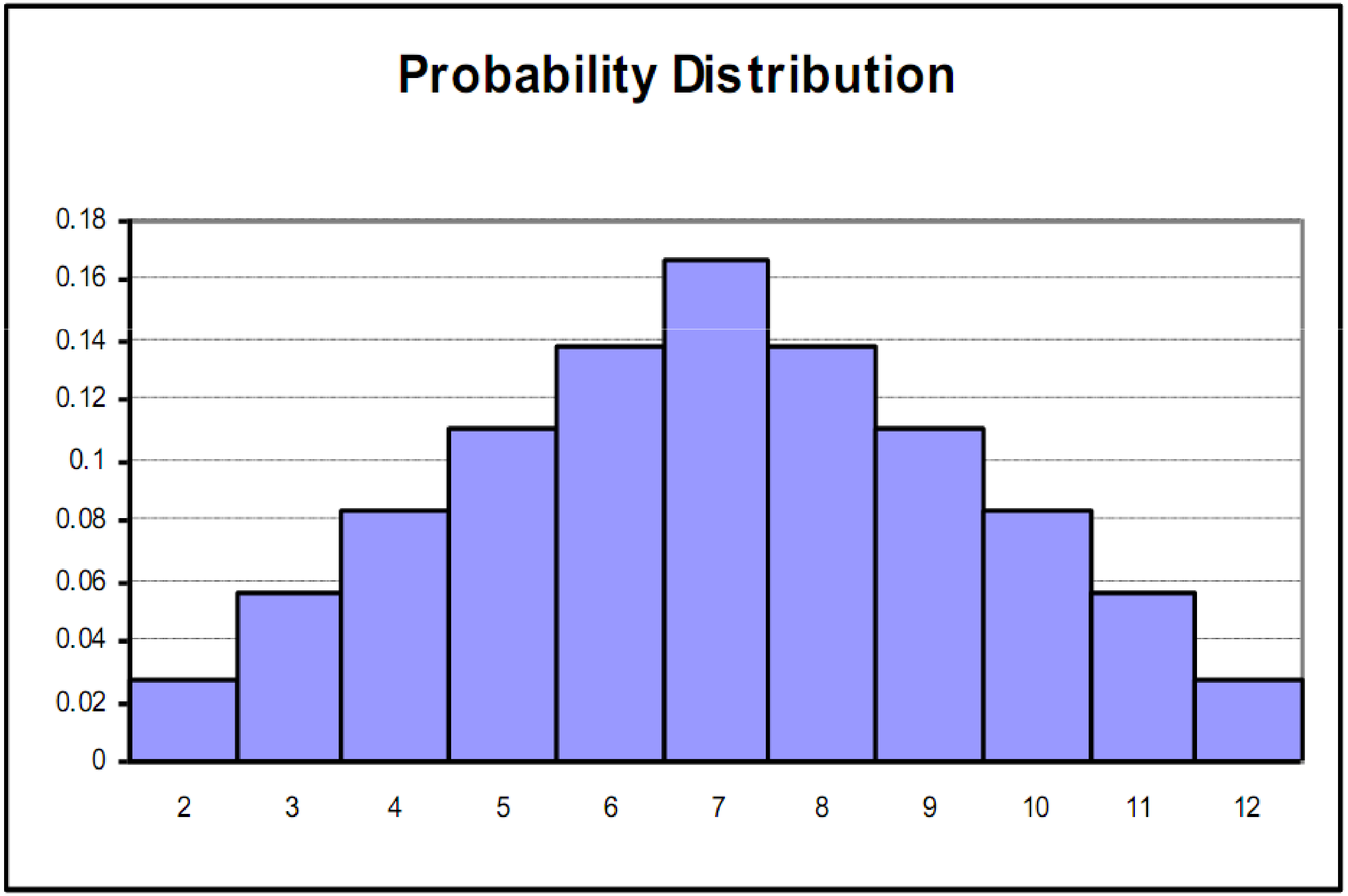
Random Variables (cont.)
Discrete Random Variables

Probability distribution, with X = Total obtained from throwing two dice

x	f(x)	F(x)
1	0	0
2	1/36	1/36
3	2/36	3/36
4	3/36	6/36
5	4/36	10/36
6	5/36	15/36
7	6/36	21/36
8	5/36	26/36
9	4/36	30/36
10	3/36	33/36
11	2/36	35/36
12	1/36	36/36

Random Variables (cont.)

Discrete Random Variables



Random Variables (cont.)

Discrete and Continuous Random Variables

X (discrete)	Y (continuous)
Probability Distribution	Probability Density Function (PDF)

x_i	f(x_i)
1	0
2	1/36
…	…
12	1/36

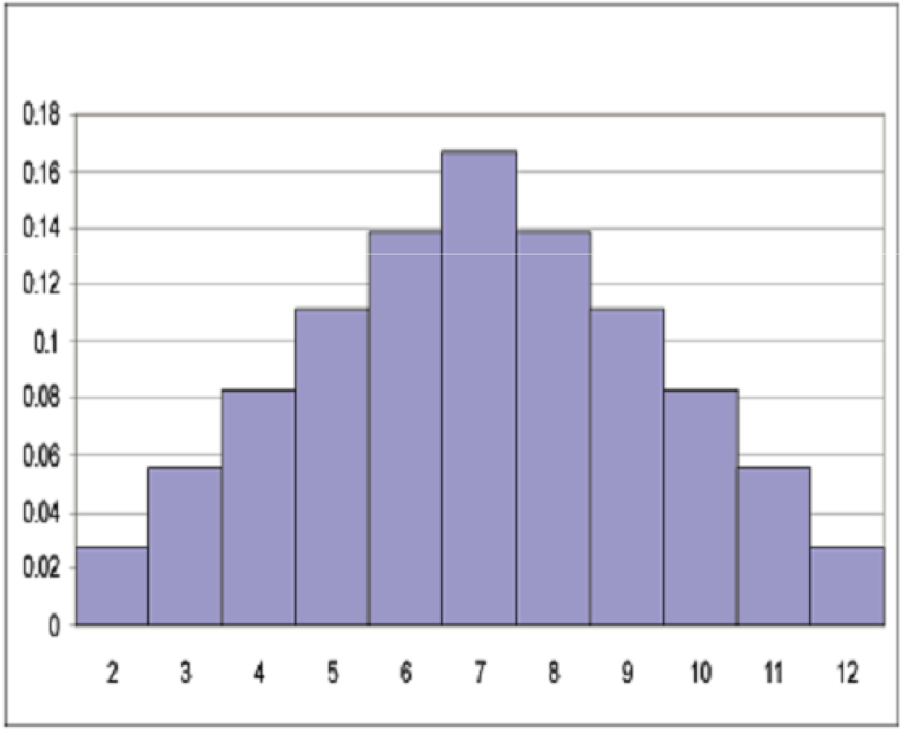
$$f(y_i) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \mu}{\sigma}\right)^2\right)$$

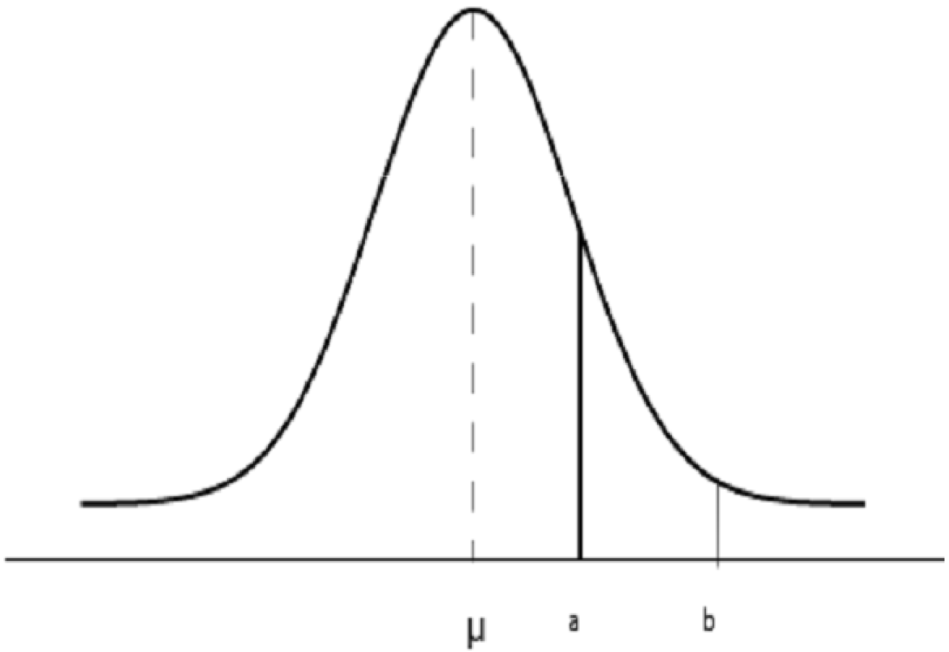
Cummulative Distribution	Cummulative Distribution Function

$$F(x_j) = prob(X \le x_j) = \sum_{x_i < x_j} f(x_i)$$

$$F(y_j) = prob(Y \le y_j) = \int_{\infty}^{y_j} f(y_i)\,dy_i$$

Random Variables (cont.)
Discrete and Continuous Random Variables





Random Variables (cont.)

Distributions

- **Marginal Distribution** is like the individual probability: $f(x_i)$
- **Joint Distribution** is like the intersection of two events: $f(x_i, y_i)$
- **Conditional Distribution** is like the conditional probability
- **Statistical Independence**: $f(x_i, y_i) = f(x_i) f(y_i)$

Characteristics of Probability Distributions

- Expected Value
- Variance
- Covariance

- Additional Readings:
  - Gujarati (2003) Appendix A.
  - Stock and Watson (2007) Chapter 2.
  - Thomas (2004) Chapter 1.
  - Wooldridge (2006) Appendix B.