ECON 214
Elements of Statistics for Economists
2016/2017

Topic – Probability

Lecturer: Dr. Bernardin Senadza, Dept. of Economics
bsenadza@ug.edu.gh

Overview

- For the past 4 weeks or so we have been discussing descriptive statistics.
- We begin the discussion of inferential statistics, the drawing of conclusions from sample data.
- Probability, or computing the chance that something will occur in the future, underlies statistical inference.
- If samples are drawn at random from a population, their characteristics (such as the sample mean) depend upon chance.
- Hence to understand how to interpret sample evidence, we need to understand chance, or probability.

Overview cont'd

- At the end of this lecture, you should be able to:
  - Define probability and describe the classical, empirical and subjective approaches to probability.
  - Explain terms such as experiment, event, outcome, and sample space.
  - Calculate probabilities using the rules of addition and multiplication.
  - Use a tree diagram to organize and compute probabilities.
  - Determine the number of outcomes in an experiment using the multiplication, permutation and combination formulas.

Defining probability

- If samples are drawn at random, their characteristics (such as the sample mean) depend upon chance.
- Hence to understand how to interpret sample evidence, we need to understand chance, or probability.
- Simply put, **probability is the chance or likelihood that something or an event will happen**.

Defining probability

- The probability of an event *A* may be defined in two ways:
  - The **objective** (or **frequentist**) **view**: the proportion of trials in which the event occurs, calculated as the number of trials approaches infinity.
  - The **subjective view**: someone's degree of belief about the likelihood of an event occurring.

Defining probability

- The objective view of probability in turn may be divided into (1) classical probability and (2) empirical probability.

Defining probability

- **Classical view** assumes that all outcomes of an experiment are equally likely and mutually exclusive, and that the probability of an event occurring is the ratio of the number of outcomes to the sample space.
  - By mutually exclusive, we mean the occurrence of any one outcome precludes the occurrence of any other in the same trial.
  - In the classical approach, the probability of an event is known "a priori."
  - **Example:** for a fair die, the probability of any number appearing in a single toss is $^{1}/_{6}$.

Defining probability

- In the **empirical** (or relative frequency) approach, the probability is determined on the basis of the proportion of times that a particular event occurs in a set of trials.
- This approach is called empirical because it is based on the collection and analysis of data.
- The probability value obtained in both classical and empirical approaches indicate the long run rate of occurrence of the event (that is, when the experiment is performed a large number of times).

Some terminologies

- **Experiment:** an activity such as tossing a coin, which has a range of possible observations or outcomes.
- **Outcome:** a particular result of an experiment.
- **Trial:** a single performance of the experiment.
- **Sample space:** all possible outcomes of the experiment. For a single toss of a coin the sample space is {Heads, Tails}.
- **Event:** a collection of one or more outcomes of an experiment.

Calculating probability

- Let *A* be the event we want to calculate a probability for, and *S*, the sample space.
- The probability of the event A occurring is given by

$$\mathrm{P}(A) = \frac{n(A)}{n(S)}$$

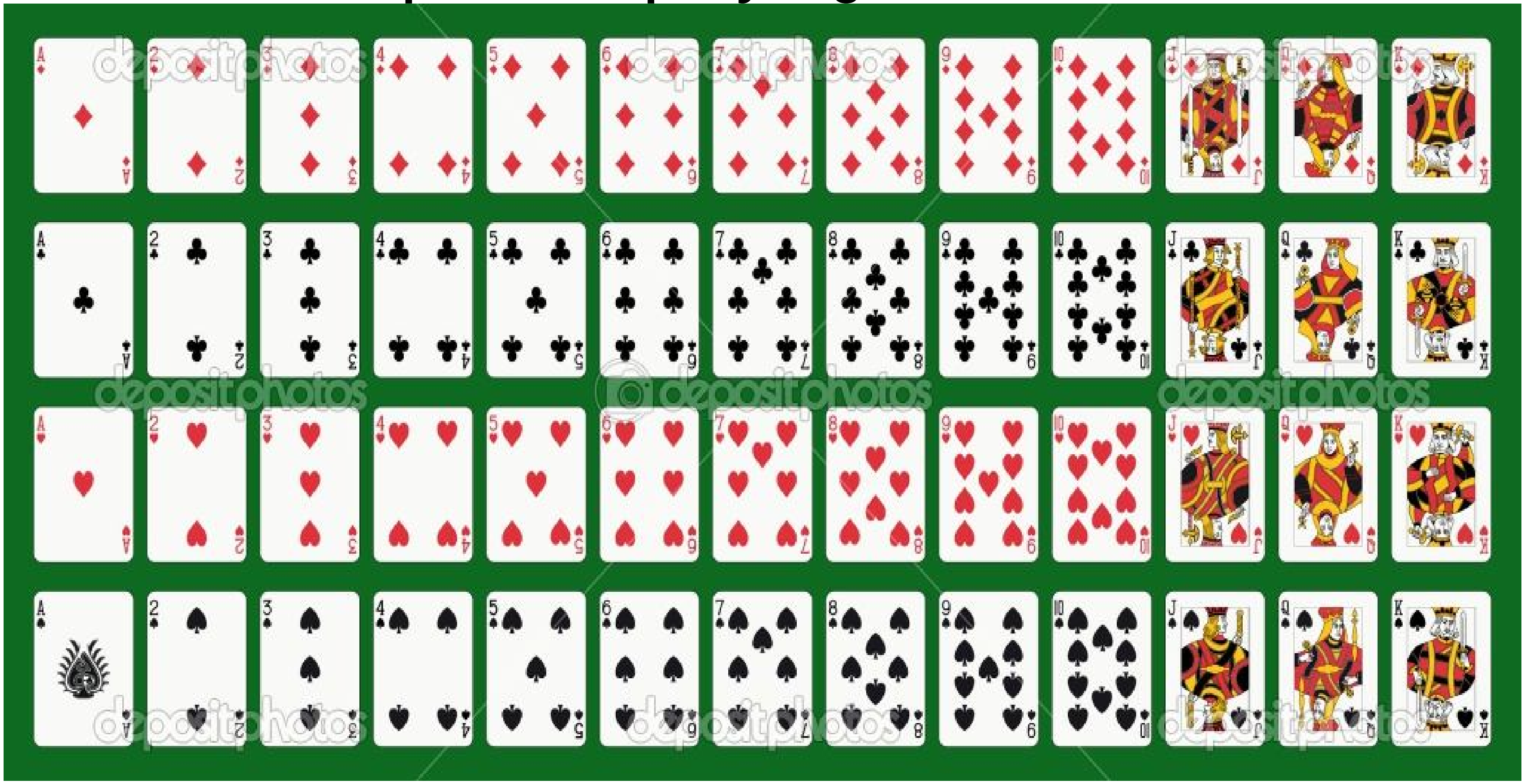
- Where n(A) is the number of times the event ***A*** occurs and n(S) is the sample space.
- Example: On a single toss of a coin, P (Heads) = ½ ; P (Tails) = ½. Why?

Rules of probabilities

- The probability of an event lies between zero and 1, i.e. $0 \leq P(A) \leq 1$
- $\sum P(A_i) = 1$; i.e. summed over all outcomes
- P (not *A*) = 1 - P(*A*)
- P (not *A*) is called the complement of *A*.

Illustration

- Consider a pack of playing cards



Illustration

- The probability of picking any one card from a pack (e.g. King of Spades) is 1/52.
- This is the same for each card since there are 52 cards in all.
- Summing over all cards: 1/52 + 1/52 + ... +1/52= 1
- P (not King of Spades) = 51/52 = 1-P(King of Spades).

Compound events

- Often we want to calculate more complicated probabilities:
  - what is the probability of drawing any Spade?
  - what is the probability of throwing a ‘double six’ with two dice?
  - what is the probability that a randomly chosen student from this class obtained a grade B or better in ECON 213 last semester?
- These are **compound events** because they involve more than one outcome.
- A student making a grade B or better must either make a B, B+ or an A.

Mutually exclusive and non-exclusive events

- **Mutually exclusive**: two or more events that cannot occur together.
  - Example: rolling a die and finding the probability of 4 or 5 showing up.
- **Non-exclusive**: two or more events that can occur together.
  - Example: picking a king of hearts from a pack of cards.

Addition rule for probabilities

- The addition rule is used when we wish to determine the probability of either one event or another (or both).
- The addition rule: the **'or'** rule
  - P(*A* or *B*) = P(*A*) + P(*B*)
    The probability of rolling a five or six on a single roll of a die is

1	2	3	4	5	6

  - P(5 or 6) = P(5) + P(6) = 1/6 + 1/6 = 1/3
  - This is the (**special**) addition rule for mutually exclusive events.

Example 1

- The Ghana Civil Aviation Authority recently supplied the following information on flights to the Kotoka International Airport:

Arrival	Frequency
Early	100
On Time	800
Late	75
Canceled	25
Total	1000

- If A is the event that a flight arrives early, and B is the event that a flight arrives late, find the probability that a flight is either early or late.

A Note on Complement Rule

- The complement rule is used to determine the probability of an event occurring by subtracting the probability of the event *not* occurring from 1.
- If P(A) is the probability of event A and P(A') is the probability of the complement of A, then

$$P(A) + P(A') = 1 \quad \text{OR} \quad P(A) = 1 - P(A').$$

- In previous example, if C is the event that a flight arrives on time and D is the event that a flight is canceled, we can use the complement rule to calculate the probability of an early (A) or a late (B) flight.

Addition rule for probabilities

- What if events *A* and *B* can simultaneously occur (i.e., they are non-exclusive)?
- Then the previous addition formula gives the wrong answer...
  - P (King or Heart) = 4/52 + 13/52 = 17/52 ✖
  - This double counts the King of Hearts; 16 dots highlighted

	A	K	Q	J	10	9	8	7	6	5	4	3	2
Spades	•	•	•	•	•	•	•	•	•	•	•	•	•
Hearts	•	•	•	•	•	•	•	•	•	•	•	•	•
Diamonds	•	•	•	•	•	•	•	•	•	•	•	•	•
Clubs	•	•	•	•	•	•	•	•	•	•	•	•	•

Addition rule for probabilities

- We therefore subtract the King of Hearts:

  So P(King or Heart) = 4/52 + 13/52 - 1/52 = 16/52
- The formula is therefore
  - P(A or B) = P(A) + P(B) - P(A and B)
- This is the **general** rule for the addition of probabilities (whether mutually exclusive or non-exclusive events).
- When A and B cannot occur simultaneously (or are mutually exclusive), then P(A and B) = 0, and we obtain the special rule discussed earlier.

Example 2

- In a sample of 500 University of Ghana students, 320 said they had a stereo, 175 said they had a TV, and 100 said they had both. If a student is selected at random, what is the probability that the student has either a stereo or a TV in his or her room?

Dependent and Independent Events

- **Independent events**: when the occurrence (or non-occurrence) of one event has no effect on the probability of occurrence of the other.
  – Example: tossing a coin twice and obtaining two heads.
- **Dependent events**: when the occurrence (or non-occurrence) of one event does affect the probability of occurrence of the other.
  – Example: drawing two aces from a pack of cards, one after the other, without replacement.

The multiplication rule

- Multiplication rule is used when we want the probability that all of several events will occur.
- That is, P(*A* and *B*).
- P(*A* and *B*) = P(*A*) $\times$ P(*B*)
- This is the (**special**) multiplication rule for **independent** events.
- Example: probability of obtaining a double-six when rolling two dice is: P (6 and 6) = P (6) $\times$ P (6) = $^1/_6 \times {}^1/_6 = {}^1/_{36}$.

Example 3

- Kwame owns stocks in companies A and B which are independent of each other. The probability that stock A increases in value next year is 0.5. The probability that stock B will increase in value next year is 0.7. What is the probability that both stocks increase in value next year?

Multiplication rule

- There is a slight complication when events are dependent.
- For instance, P(drawing two Aces from a pack of cards, without replacement)...
- If the first card drawn is an Ace ($P = {}^{4}/_{52}$), that leaves 51 cards, of which 3 are Aces.
- The probability of drawing the second Ace is ${}^{3}/_{51}$, different from the probability of drawing the first Ace.
- They are not **independent events**; the probability changes.
- Thus P(two Aces) = ${}^{4}/_{52} \times {}^{3}/_{51} = {}^{1}/_{221}$

Conditional probability

- $^3/_{51}$ is the probability of drawing a second Ace **given** that an Ace was drawn as the first card.
- This is the **conditional probability** and is written

  P (Second Ace | Ace on first draw)

  – That is, the probability of a second Ace given that an Ace was drawn first.
- In general, it is written as **P(B|A)**
- That is, the probability of event B occurring, given that A has occurred.

Conditional probability

- Consider P (A2| not-A1)...
- That is, 'not-Ace' is drawn first, leaving 51 cards of which 4 are Aces.
- Here, we assumed that the first card drawn is **not** an Ace.
- Hence P (A2| not-A1) = $^{4}/_{51}$
- So P (A2| not-A1) $\neq$ P (A2|A1)
- They are **not** independent events.

Conditional probability

- The general rule for multiplication is:

  $P(A \text{ and } B) = P(A) \times P(B|A)$

  So that P(B|A) = P(A and B)/P(A)

- For independent events

  $P(B|A) = P(B|\text{ not-}A) = P(B)$

- And so

  $P(A \text{ and } B) = P(A) \times P(B)$

Example 4

- The Dean of UGBS presented the following information on undergraduate students during a School Board meeting. (1) If a student is selected at random, what is the probability that the student is a female accounting major? (2) Given that the student is a female, what is the probability that she is an accounting major?

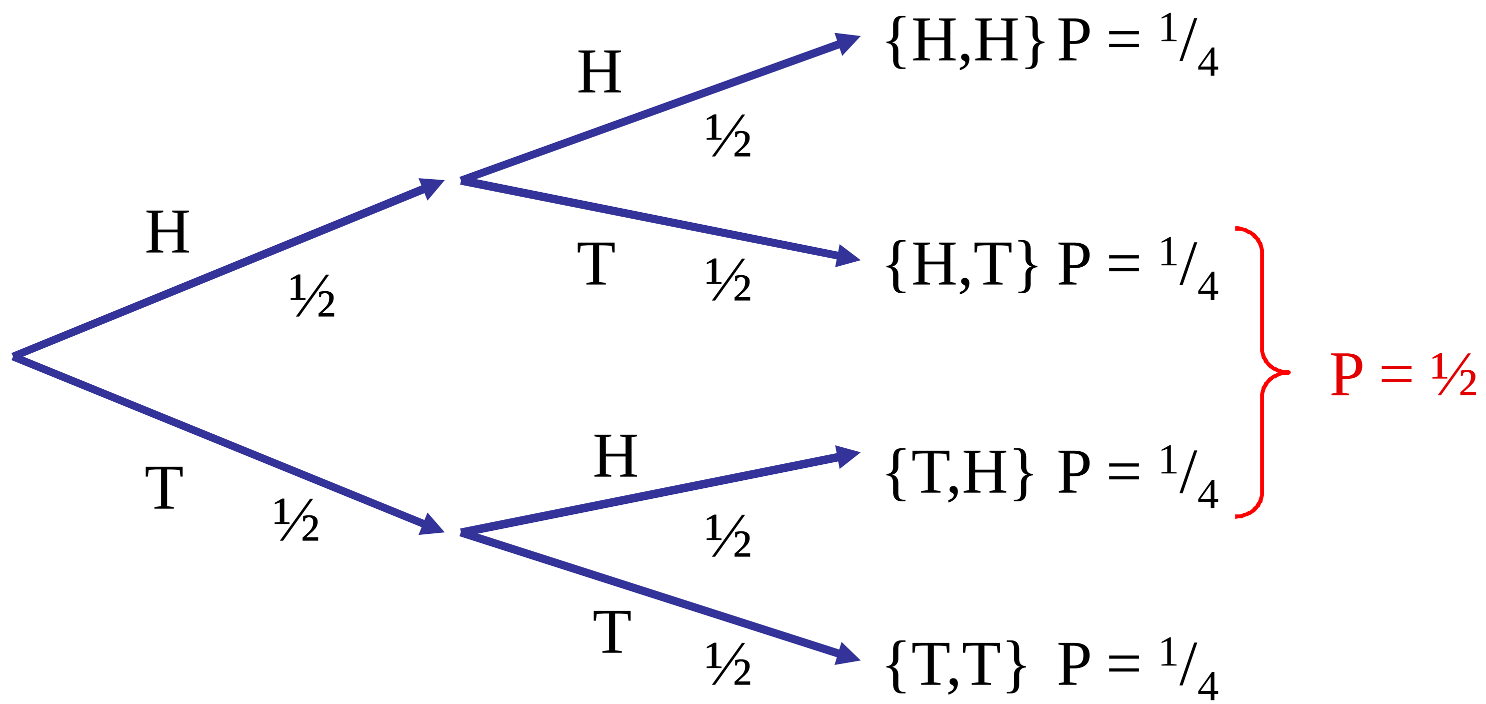
MAJOR	Male	Female	Total
Accounting	170	110	280
Finance	120	100	220
Marketing	160	70	230
Management	150	120	270
Total	600	400	1000

Combining the rules

- Consider that we wish to find the probability of one head in two tosses of a coin.
- We can solve this by combining the addition and multiplication rules
- P (1 Head in two tosses)...
- ...= P **(** [H and T] or [T and H] **)**
  = P **(** [H and T] **)** + P **(** [T and H] **)**
  = $[1/2 \times 1/2] + [1/2 \times 1/2]$
  = $1/4 + 1/4 = 1/2$

Tree diagram

- We can also solve the problem using a tree diagram



Gets complicated with larger number of outcomes

- What about calculating
  - P (3 Heads in 5 tosses)?
  - P (30 Heads in 50 tosses)?
- How many routes through the tree diagram?
  - Drawing takes too much time, so we need a formula...
- In other words, if the number of possible outcomes in an experiment is large, it is difficult or cumbersome to list the total number of outcomes in either the event set or sample space.
- There are techniques for determining the number of outcomes.

Multiplication formula

- The multiplication formula is used to find the total number of outcomes for two or more groups of objects.
- If there are n_1 objects of one kind and n_2 of another, then there are $n_1 \times n_2$ ways of selecting both.
- That is, total number of outcomes = $n_1 \times n_2$
- In general if there are *k* groups of objects, and there are n_1 items in the first group, n_2 items in the second, …, and n_k items in the k^{th} group, the number of ways we can select one item each from the *k* groups is: $n_1 \times n_2 \times \ldots\ldots\ldots \times n_k$
- If $n_1 = n_2 = \ldots\ldots\ldots = n_k$, then $n_1 \times n_2 \times \ldots\ldots\ldots \times n_k = n^k$

Example 5

- Dr. Asempa has 10 shirts and 8 ties. How many shirt/tie outfits does he have?
- Answer = 10x8 = 80.

Permutation formula

- The permutation formula is applied to find the number of outcomes when there is only one group.
- We are interested in how many different subsets that can be obtained from a given set of objects.
- Example - the number of ways of having 3 girls in a family of 5 children.
- If ***r*** items are selected from a set on ***n*** objects (where $r \leq n$), any particular sequence of these r items is called a permutation.

Permutation formula

- The formula is

$$nPr = \frac{n!}{(n-r)!}$$

- Where $n! = n \times (n-1) \times (n-2) \times \ldots \times 1$

- n = total number of objects
- r = number of objects selected at a time
- n! is called “n factorial” and it is the product of all the positive integers up to and including n.

Permutation

- So the number of different ways of having 3 girls in a family of 5 children is

$$5P3 = \frac{5\times 4\times 3\times 2\times 1}{2\times 1} = 60$$

- In permutation the <u>order</u> of arrangement of the objects is <u>important</u>!
- For example, the arrangement (Kofi, Ama) is a different permutation from (Ama, Kofi) even though it is the same two individuals.

Combination formula

- Just like permutation, the **combinatorial formula** gives the number of ways in which a particular event may occur, but **<u>without</u>** regard to order.
- The formula for combination is

$$nCr = \frac{n!}{r!(n-r)!}$$

- Again $n! = n \times (n-1) \times (n-2) \times \ldots \times 1$

Combination formula

- So the number of ways of having 3 girls in a family of 5 children (without regard to order of arrange) is

$$5C3 = \frac{5\times 4\times 3\times 2\times 1}{(3\times 2\times 1)\times(2\times 1)} = 10$$

Applying Combination Formula in Calculating Probabilities

- We can write the probability of 1 Head in 2 tosses as the probability of a head and a tail (in that order) times the number of possible orderings (# of times that event occurs).
- P (1 Head) = $\frac{1}{2} \times \frac{1}{2} \times 2C1 = \frac{1}{4} \times 2 = \frac{1}{2}$ (as before).
- We can formalise this in the Binomial distribution ….. soon!