Business Analytics

UNIVERSITETI I EJL
ЈИЕ УНИВЕРЗИТЕТ
SEE UNIVERSITY

Probability

Faton Berisha

Chapter 3

Probability

Probability

Probability Concepts

- An *experiment* is any process of observation with an uncertain outcome
- The possible outcomes for an experiment are called the *experimental outcomes***.**
    - Each individual repetition of the experiment results with exactly one experimental outcome.
- *Probability* is a measure of the chance that an experimental outcome will occur when an experiment is carried out

Probability

- If E is an experimental outcome, then $P(E)$ denotes the probability that E will occur and:

Conditions

1. $0 \leq P(E) \leq 1$

   such that:

   ❖If E can never occur, then $P(E) = 0$

   ❖If E is certain to occur, then $P(E) = 1$

2. The probabilities of all the experimental outcomes must sum to 1

Assigning Probabilities to
Experimental Outcomes

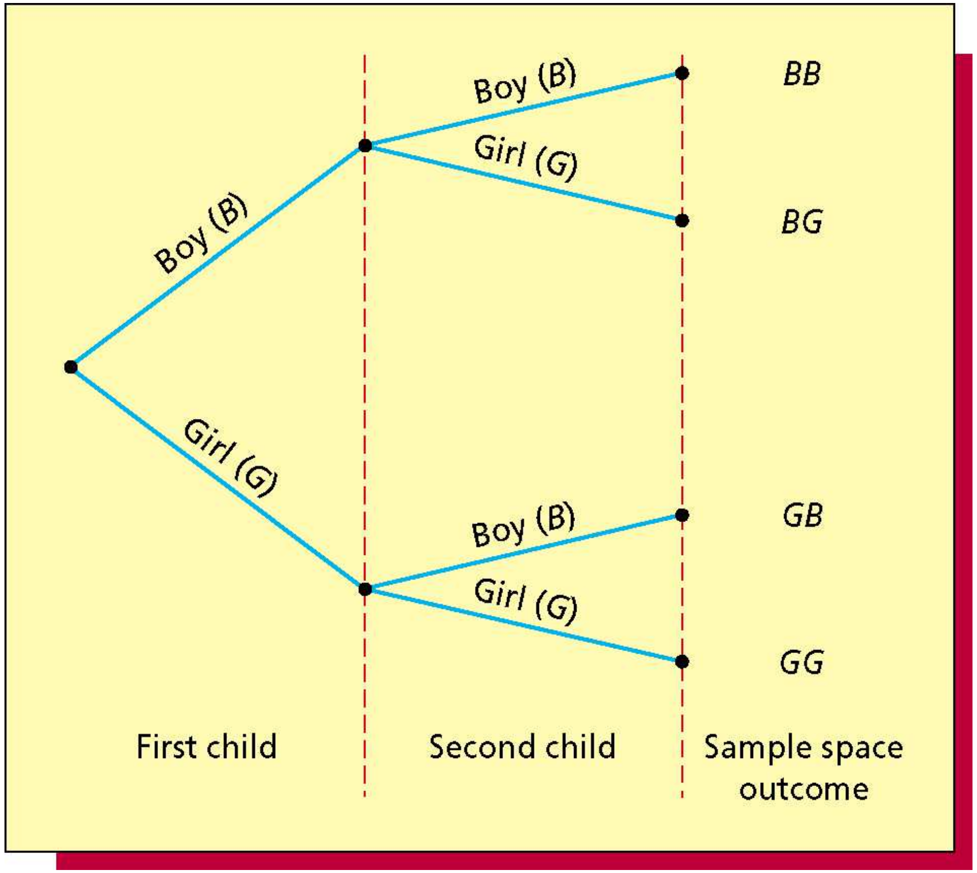
- Classical Method
  - For equally likely outcomes
- Relative frequency
  - In the long run
- Subjective
  - Assessment based on experience, expertise or intuition

Sample Spaces and Events

- *Sample space* of an experiment is the set of all possible experimental outcomes (*sample space outcomes, elementary events*).
  - Example 3.1. Choosing a CEO in a company
  - Example 3.2. Genders of two children
  - Example 3.3. Aswering three questions in a pop-up quiz

Example: Gjinitë e fëmijëve

Example 3.2: Genders of two children



Event and Probability of an Event

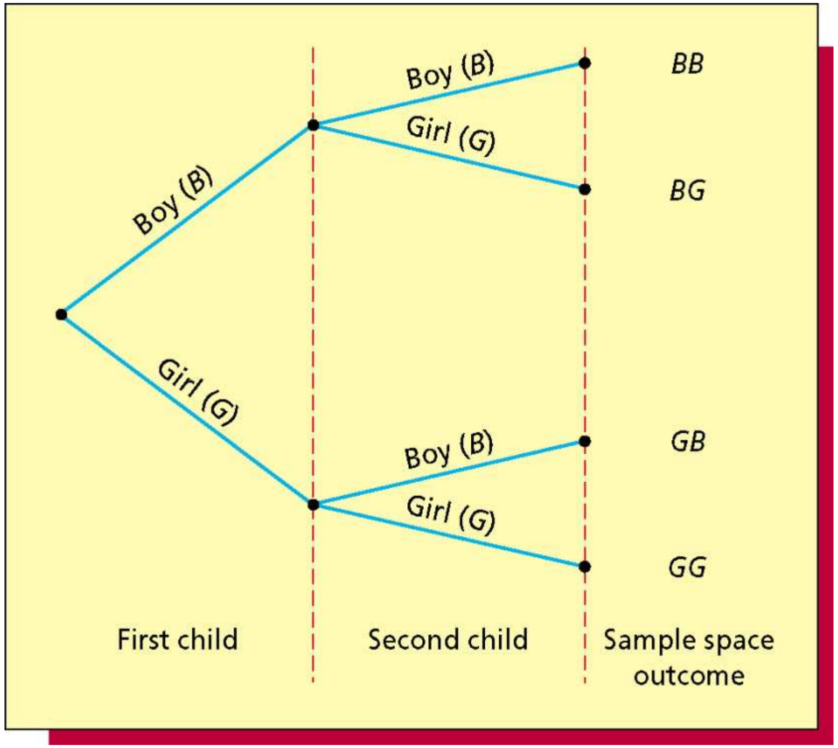
- An *event* is a set (or collection) of sample space outcomes.
- The *probability* of an event is the sum of the probabilities of the experimental outcomes that belong to the event.
- If A is an event, then

$$0 \le P(A) \le 1.$$

  - If A can never occur, then $P(A) = 0$ (*impossible event*).
  - If A is certain to occur, then $P(A) = 1$ (*certain event*).

Example of Events: Gender of Children

Example 3.4. Genders of two children



Events:

P(Two boys) = $P(BB)$ = ¼

P(One boy and one girl)

= $P(BG) + P(GB)$ = ¼ + ¼ = ½

P(Two girls) = $P(GG)$ = ¼

P(At least one girl)

= $P(BG)$ + P(GB) + P(GG)

= ¼ + ¼ + ¼ = ¾

Note. All outcomes equally likely: $P(BB)$ = ... = $P(GG)$ = ¼

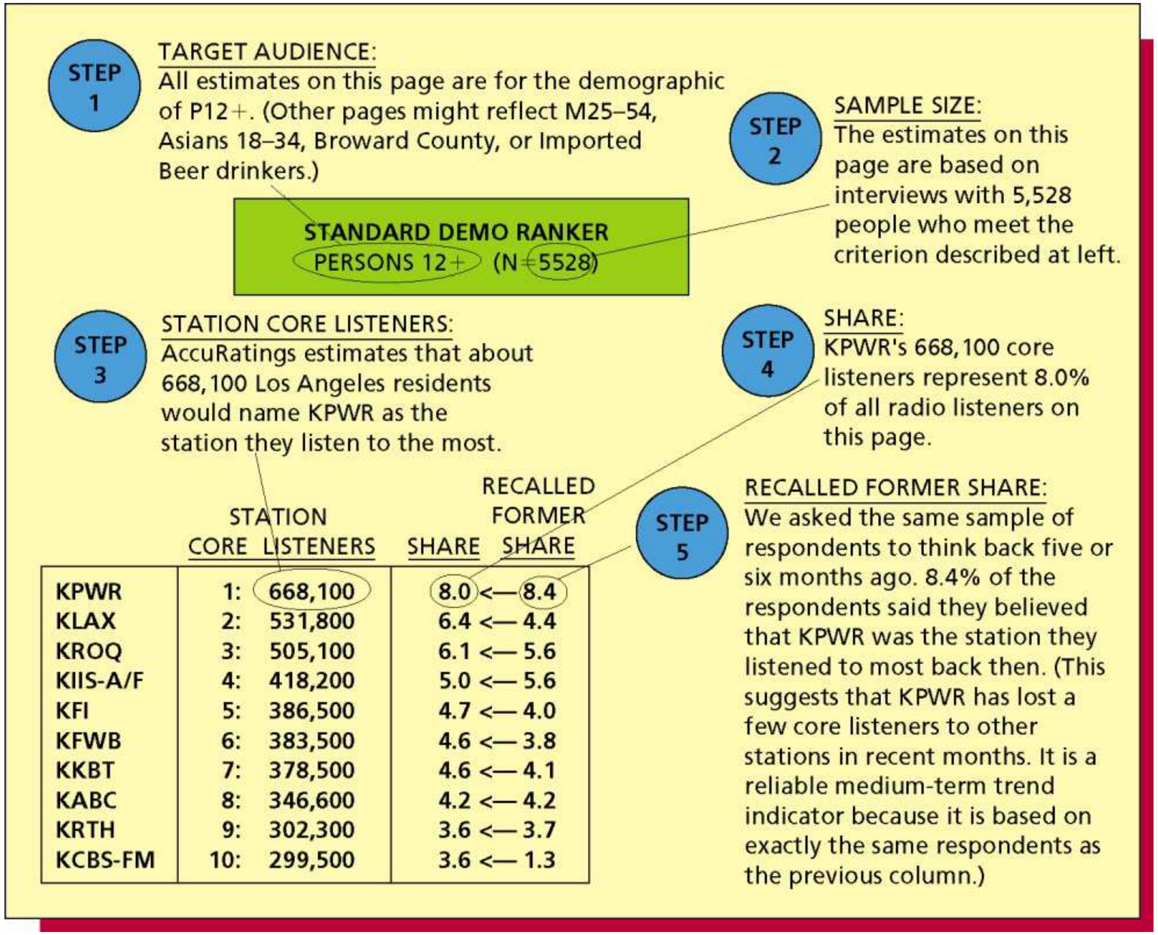
Probability

- If the sample space outcomes (or experimental outcomes) are all equally likely, then the probability that an event will occur is equal to the ratio

$$\frac{\text{Number of sample space outcomes that correspond to the event}}{\text{Total number of sample space outcomes}}$$

Example: AccuRatings Case

Example 3.7. Case of media ratings



STEP 1 — TARGET AUDIENCE:
All estimates on this page are for the demographic of P12+. (Other pages might reflect M25–54, Asians 18–34, Broward County, or Imported Beer drinkers.)

STANDARD DEMO RANKER
PERSONS 12+ (N=5528)

STEP 2 — SAMPLE SIZE:
The estimates on this page are based on interviews with 5,528 people who meet the criterion described at left.

STEP 3 — STATION CORE LISTENERS:
AccuRatings estimates that about 668,100 Los Angeles residents would name KPWR as the station they listen to the most.

STEP 4 — SHARE:
KPWR's 668,100 core listeners represent 8.0% of all radio listeners on this page.

STEP 5 — RECALLED FORMER SHARE:
We asked the same sample of respondents to think back five or six months ago. 8.4% of the respondents said they believed that KPWR was the station they listened to most back then. (This suggests that KPWR has lost a few core listeners to other stations in recent months. It is a reliable medium-term trend indicator because it is based on exactly the same respondents as the previous column.)

		STATION CORE LISTENERS	SHARE	RECALLED FORMER SHARE
KPWR	1:	668,100	8.0	8.4
KLAX	2:	531,800	6.4	4.4
KROQ	3:	505,100	6.1	5.6
KIIS-A/F	4:	418,200	5.0	5.6
KFI	5:	386,500	4.7	4.0
KFWB	6:	383,500	4.6	3.8
KKBT	7:	378,500	4.6	4.1
KABC	8:	346,600	4.2	4.2
KRTH	9:	302,300	3.6	3.7
KCBS-FM	10:	299,500	3.6	1.3

Estimated probability (percentage):

$P(KPWR) = 445/5528$

$= 0.8050$

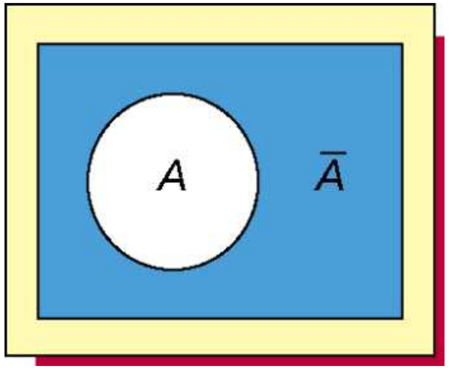
Assumed 8,300,000 inhabitants in LA of age 12+

Estimated number of listeners:

$Listeners = N \cdot P(KPWR)$

$= 8{,}300{,}000 \cdot 0.08 = 668{,}100$

Some Elementary Probability Rules



The *complement* $\overline{A}$ of an event A is the set of all sample space outcomes not in A

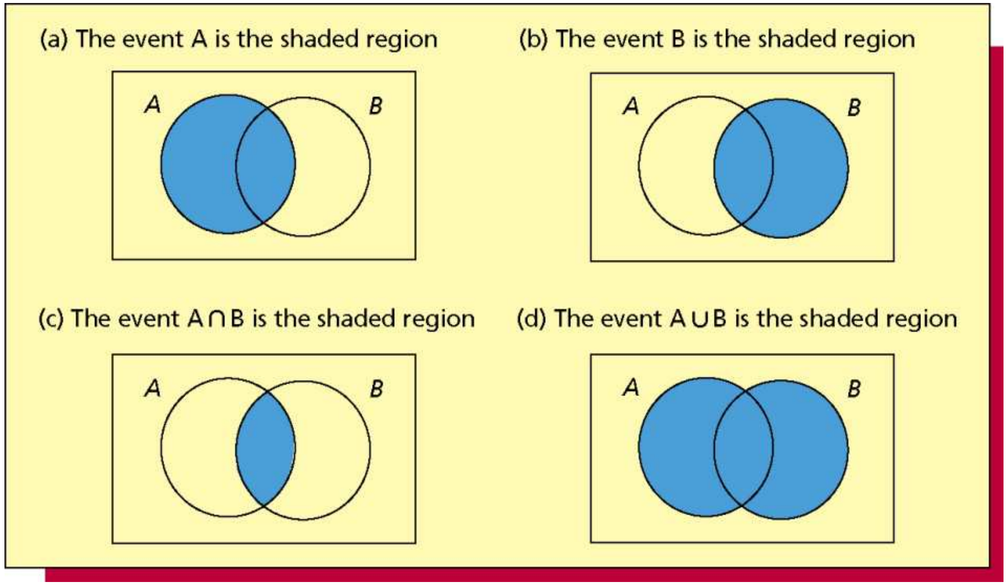
$$\boxed{P(\overline{A}) = 1 - P(A)}$$

Union of *A* and *B***,** $A \cup B$

Elementary events that belong to either *A* **or** *B* (or both)

Intersection of *A* and *B***,** $A \cap B$

Elementary events that belong to both *A* **and** *B*



These figures are "Venn diagrams"

The Addition Rule

The probability that *A* or *B* (the union of A and B) will occur is

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

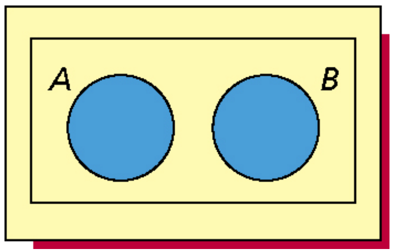
where $P(A \cap B)$ is the "joint" probability of A and B both occurring together

A and *B* are *mutually exclusive* if they have no sample space outcomes in common, or equivalently, i.e.

$$P(A \cap B) = 0$$

If A and B are mutually exclusive, then

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



Example 3.12. AccuRating case

Conditional Probability

The probability of an event *A*, given that the event *B* has occurred, is called the *conditional probability* of *A* given *B* and is denoted as $P(A \mid B)$

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

Note: $P(B) \neq 0$

Interpretation: Restrict the sample space to just event B. The conditional probability $P(A|B)$ is the chance of event A occurring in this new sample space

- *If* A occurred, *then* what is the chance of B occurring?

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

Independence of Events

Two events A and B are said to be *independent* if and only if:

$$P(A \mid B) = P(A)$$

or, equivalently,

$$P(B \mid A) = P(B)$$

The Multiplication Rule

The joint probability that *A* and *B* (the intersection of *A* and *B*) will occur is

$$P(A \cap B) = P(A) \cdot P(B \mid A)$$
$$= P(B) \cdot P(A \mid B)$$

If A and B are *independent*, then the probability that *A* and *B* (the intersection of *A* and *B*) will occur is

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

Examples 3.13, 3.14

Contingency Tables

$P(R_1 \cap C_1)$

$P(R_1)$

	C_1	C2	Total
R_1	.4	.2	.6
R_2	.1	.3	.4
Total	.5	.5	1.00

$P(R_2 \cap C_2)$

$P(C_2)$

Bayes' Theorem

- Sometimes, we have a *prior* probability that an event will occur
- Then, based on new information, we revise the prior probability
  - This is called a *posterior* probability
- This revision is done using Bayes' Theorem

Bayes' Theorem
Example 3.17: AIDS Testing

Question: Suppose that a person selected randomly for testing, tests positive for AIDS. The test is known to be highly accurate (99.9% for people who have AIDS, 99% for people who do not.) What is the probability that the person actually has AIDS?

Answer: Surprisingly, much lower than most of us would guess!

The Facts :

AIDS Incidence Rate : 6 cases per 1000 Americans

$$P(AIDS) = 0.006 \qquad P(\overline{AIDS}) = 0.994$$

Testing Accuracy :

$$P(POS|AIDS) = 0.999 \qquad P(POS|\overline{AIDS}) = 0.01$$

Solution : $P(AIDS|POS)$

Example 3.17: AIDS Testing Continued

$$P(AIDS|POS) = \frac{P(AIDS \cap POS)}{P(POS)} = \frac{P(AIDS \cap POS)}{P(AIDS \cap POS) + P(\overline{AIDS} \cap POS)}$$

$$= \frac{P(AIDS)P(POS|AIDS)}{P(AIDS)P(POS|AIDS) + P(\overline{AIDS})P(POS|\overline{AIDS})} \qquad (Bayes'Theorem)$$

$$= \frac{(0.006)(0.999)}{(0.006)(0.999) + (0.994)(0.01)} = \frac{0.005994}{0.005994 + 0.00994}$$

$$= \frac{0.005994}{0.015934}$$

$$= \mathbf{0.3}8$$