Lecture Notes EMSE 4765: DATA ANALYSIS - Probability Review

Chapter 2: Outcomes, Events, and Probability

Version: 01/08/2021

THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

Text Book: A Modern Introduction to Probability and Statistics, Understanding Why and How

By: F.M. Dekking. C. Kraaikamp, H.P.Lopuhaä and L.E. Meester

2 Outcomes, Events, and Probability

2.1 Sample Spaces...

- **Sample spaces are sets** whose **elements describe the possible outcomes of an experiment** in which we are interested. Notation:

$$\Omega \equiv \text{Sample Space.}$$

- **Experiment 1: The tossing of a coin** :

$$\Omega = \{H, T\}, H = \text{"Heads"}, T = \text{"Tails"}$$

- **Experiment 2: The tossing of a die** :

$$\Omega = \{\ 1, 2, 3, 4, 5, 6\}$$

- **Experiment 3: Time to failure of a part *t*** :

$$\Omega = [0, \infty)$$

- **Experiment 4: One step log-differences of interest rates** :

$$Ln(i_{k+1}) - Ln(i_k) \in (-\infty, \infty) \Rightarrow \Omega = (-\infty, \infty), \text{ Why?}$$

- **Events are combinations of outcomes, i.e. subsets** of the sample space Ω. **Event *A*** *occurs* if the outcome is an element of the **subset *A*.**

Die Experiment: Those outcomes that are even : $E = \{2, 4, 6\}$.

Die Experiment: Those outcomes that are odd : $O = \{1, 3, 5\}$.

Die Experiment: Those outcomes that are prime: $P = \{1, 2, 3, 5\}$

Die Experiment: Those outcomes that are prime and even: $P \cap E = \{2\}$

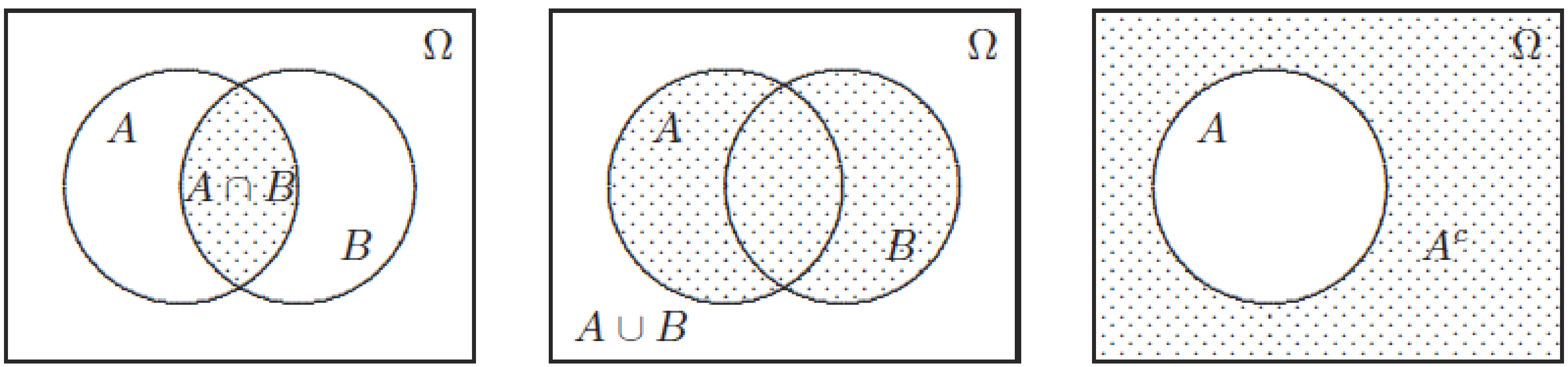
Die Experiment: Those outcomes that are prime or odd:

$$P \cup O = \{1, 2, 3, 5\}.$$

- Thus, **Events can be combined:** **The intersection $A \cap B$** is a combination of events A and B and we say that for $A \cap B$ to *occur*, **both events A and B have to *occur*.**

- **The union $A \cup B$:** The union event $A \cup B$ *occurs*, **when either A or B *occurs* or both.**

- **The complement A^c:** The complement event A^c *occurs*, **when A does not *occur*.**

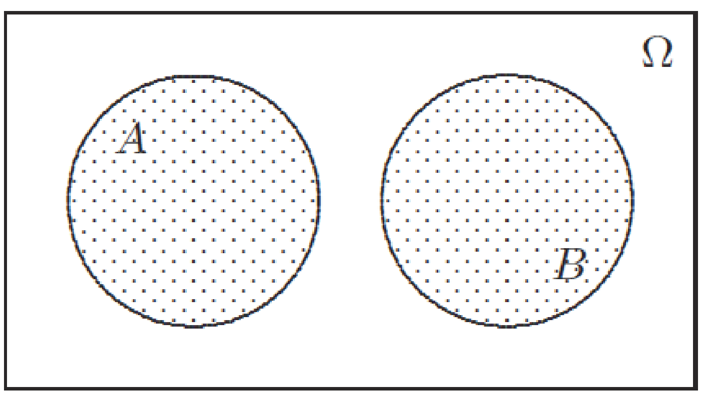


Intersection $A \cap B$ Union $A \cup B$ Complement A^c

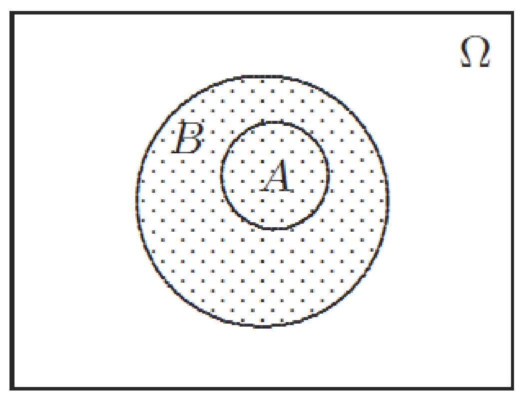
2 Outcomes, Events, and Probability

2.2 Events...

- **Empty Set:** The complement of the total event Ω is denoted the empty set $\emptyset$.
- **Disjoint or mutually exclusive events A and B:** Events A and B cannot occur at the same time. **In set notation: $A \cap B = \emptyset$.**
- **Event A implies event B** when all the outcomes of event A also outcomes of the event B. **In set notation: $A \subset B$.**



Disjoint sets A and B



A subset of B

2 Outcomes, Events, and Probability

2.2 Events. . .

Die Experiment: $O \cap E = \emptyset,\ P \cup O = \{1, 2, 3, 5\} = P \Rightarrow O \subset P$

Die Experiment: $P = \{1, 2, 3, 5\}, E = \{2, 4, 6\} \Rightarrow P \cap E = \{2\} \Rightarrow$

$$(P \cap E)^c = \{1, 3, 4, 5, 6\} = \{4, 6\} \cup \{1, 3, 5\} = P^c \cap E^c$$

DeMorgan's Law: For any two events A and B we have

$$(A \cap B)^c = A^c \cup B^c \text{ and } (A \cup B)^c = A^c \cap B^c$$

Question: Are the descriptions of the following events equivalent?

Event 1: "John or Mary is to blame, or both."

Event 2: "It is certainly not true that neither John nor Mary is to blame."

Solution Approach: $J = \{\text{John is to blame}\}$, $M = \{\text{Mary is to blame}\}$. Next, express the two statements above in terms of the events J, M, J^c and M^c. **Check the equivalence of the statements by means of the DeMorgan's laws.**

Answer:

"John or Mary is to blame, or both." $= J \cup M$

"Neither John nor Mary is to blame." $= J^c \cap M^c$

"It is not true that neither John nor Mary is to blame." $= (J^c \cap M^c)^c$

According to DeMorgan's Laws:

$$(J \cup M)^c = J^c \cap M^c \Leftrightarrow [(J \cup M)^c]^c = (J^c \cap M^c)^c \Leftrightarrow J \cup M = (J^c \cap M^c)^c$$

- We want to express **how likely it is that an event occurs**. To do this **one assigns a probability to each event**. Since **each event** has to be assigned a probability, we speak of **a probability function**.

Definition: A probability function $Pr(\cdot)$ on **a finite sample space Ω** assigns to each event A in Ω a number $Pr(A)$ in $[0,1]$ such that
(i) $Pr(\Omega)=1$, and
(ii) $Pr(A\cup B)=Pr(A)+Pr(B)$ if A and B are disjoint, i.e cannot occur at the same time. The number $Pr(A)$ is called the probability that A occurs.

Coin Toss Experiment: $H\equiv$ "Heads", $T\equiv$ "Tails".

$$Pr(H)=Pr(T)=\frac{1}{2}$$

Die Experiment: $D\in\{1,2,3,4,5,6\}$

$$Pr(D=d)=\frac{1}{6},\ d=1,\ldots,6$$

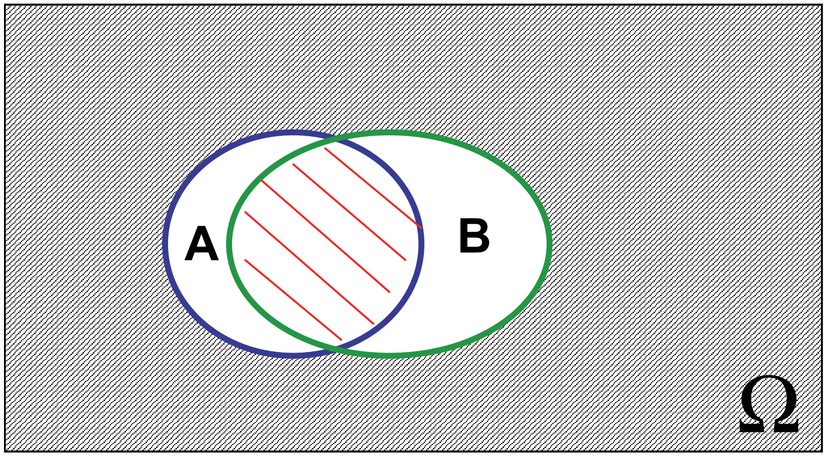
2.3 Probability Rules...

- **Complement Rule :** A and A^c are disjoint by definition $\Rightarrow$

  $$\Omega = A \cup A^c \Rightarrow Pr(\Omega) = Pr(A) + Pr(A^c) \Leftrightarrow Pr(A^c) = 1 - Pr(A)$$

- **Calculating probabilities of events $\boldsymbol{A}$ and $\boldsymbol{B}$, that are not disjoint.**

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



Example: **Recall the die experiment**

$$P = \{1, 2, 3, 5\}, E = \{2, 4, 6\}, P \cup E = \Omega,\ P \cap E = \{2\}$$

$$Pr(P \cup E) = 1,\ Pr(P) = \tfrac{4}{6},\ Pr(E) = \tfrac{1}{2},\ Pr(P \cap E) = \tfrac{1}{6}$$

Note:

1. $Pr(P \cup E) = Pr(P) + Pr(E) - Pr(P \cap E)$?

$$1 = \frac{4}{6} + \frac{1}{2} - \frac{1}{6} = \frac{4}{6} + \frac{3}{6} - \frac{1}{6} = 1? \Rightarrow \textbf{Answer: Yes!}$$

2. $Pr(P \cap E) = \frac{1}{6} \neq Pr(P) \times Pr(E) = \frac{4}{6} \times \frac{1}{2} = \frac{1}{3}$. Thus, we need to investigate **the probability of an intersection** more later.

Definition: **A probability function $Pr(\cdot)$** on **an infinite sample space Ω** assigns to each event A in Ω a number $Pr(A)$ in $[0,1]$ such that
(*i*) $Pr(\Omega) = 1$, and
(*ii*) $Pr(A_1 \cup A_2 \cup A_3 \cup \dots) = Pr(A_1) + Pr(A_2) + Pr(A_3) + \dots$
if $A_1, A_2, A_3, \dots$ are disjoint events, i.e. they cannot occur at the same time

- **We toss a coin repeatedly until the first "Heads" turns up.** The outcome of the experiment is **the number of tosses it takes for that to happen**. Thus:

$$\Omega = \{1, 2, 3, \dots\}$$

What is **the probability function $Pr(\cdot)$** for this experiment?

Suppose $Pr(H) = p \Rightarrow Pr(1) = p$, $Pr(2) = Pr(TH) = (1-p)p$,
$Pr(3) = Pr(TTH) = (1-p)^2 p$ and $Pr(n) = (1-p)^{n-1} p$

- **Does this define a probability mass function on $\Omega = \{1, 2, 3, \dots\}$?**

$$\begin{aligned} Pr(\Omega) &= Pr(1) + Pr(2) + Pr(3) + \dots \\ &= p + (1-p)p + (1-p)^2 p + (1-p)^3 p + \dots \\ &= p[1 + (1-p) + (1-p)^2 + (1-p)^3 + \dots] = 1? \end{aligned}$$

Yes, since

$$[1 + (1-p) + (1-p)^2 + (1-p)^3 + \dots] = 1/p.$$

This an example of **a geometric series** and it is well know that for $0 < x < 1$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

Hence, substitution of $x = \mathbf{1 - p}$ in theabove yields:

$$1 + (\mathbf{1-p}) + (\mathbf{1-p})^2 + (\mathbf{1-p})^3 + \dots = \frac{1}{1 - (\mathbf{1-p})} = \frac{1}{p}$$